Парадигма развития науки

Методологическое обеспечение

А.Е. Кононюк

ДИСКРЕТНО-НЕПРЕРЫВНАЯ МАТЕМАТИКА

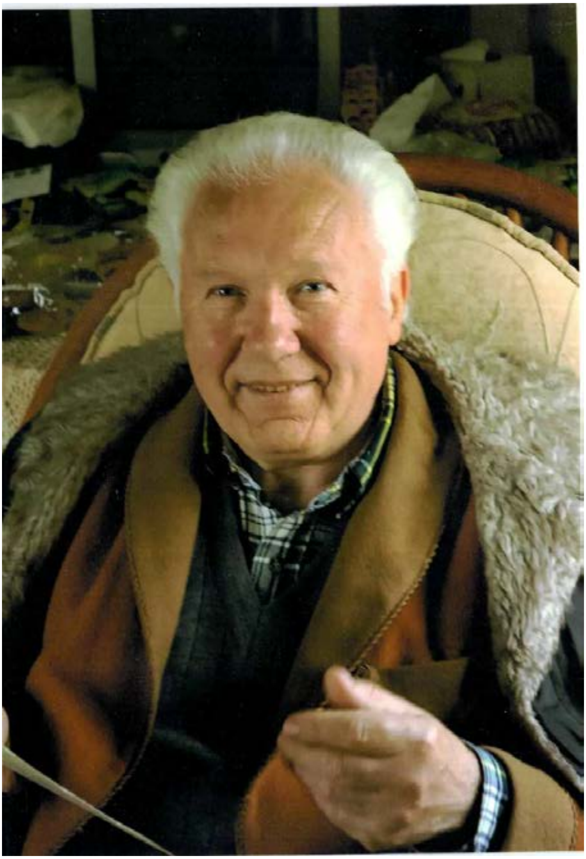
Книга 3

Отношения

Часть 1

Четкие

Киев
«Освіта України»
2013



Кононюк Анатолий Ефимович


Структурная схема развивающейся панмедийной системы наук
I. Группа междисциплинарных наук
ПН Теория понятий
ЛГ Теория лингвистики
ПЗ Теория познания
СЗ Теория созидания
ПТ Общая теория понятоведения
II. Группа наук познания (фундаментальных наук)
ИН Теория информации
СТ Теория систем
РС Теория распознавания
КС Теория консалтинга (консультирования)
ПР Теория принятия решения
РР Теория реализации принятых решений
III. Группа наук созидания (прикладных наук)
СС Теория социальных систем
ЖО Теория жизнеобеспечения
КМ Теория коммуникаций и транспортирования
ОБ Теория образования (обучения)
СР Теория сохранения (защиты и охраны)
РЗ Теория развития
IV. Группа общедисциплинарных наук (обеспечивающих наук)
ДМ Теория дискретно-непрерывной математики (четкая и нечеткая)
НИ Теория научных исследований
ММ Теория моделей и моделирования
ОМ Теория оптимизации

УДК 51 (075.8)
ББК В161.я7
К65

Рецензенты:
В.В.Довгай - к-т физ.-мат. наук, доц. (Национальный технический университет „КПІ”);
В.В.Гавриленко - д-р физ.-мат. наук, проф., *О.П.Будя* - к-т техн. наук, доц. (Киевский университет экономики, туризма и права); *Н.К.Печурин* - д-р техн. наук, проф. (Национальный авиационный университет).

Кононюк А. Е.
К213 Дискретно-непрерывная математика. (Отношения (четкие)). — В 12-и кн. Кн. 3, Ч. 1— К.:Освіта України. 2013.— 506 с.

ISBN 978-966-373-693-8 (многотомное издание)
ISBN 978-966-373-694-5 (книга 3, часть 1)

Многотомная работа содержит систематическое изложение математических дисциплин, исспользуемых при моделировании и исследованиях математических моделей систем и сетей.

В работе излагаются основы теории множеств, отношений, поверхностей, пространств, алгебраических систем, матриц, графов, математической логики, теории формальных грамматик и автоматов, теории алгоритмов, которые в совокупности образуют единную методолгически взамосвязанную математическую систему «Дискретно-непрерывная математика».

Для бакалавров, специалистов, магистров, аспирантов, докторантов всех специальностей.

УДК 51 (075.8)
ББК В161.я7

ISBN 978-966-373-693-8 (многотомное издание)
ISBN 978-966-373-694-5 (книга 3, часть 1)

Оглавление

Введение

Мы будем все время иметь дело с простыми категориями, которыми мы повседневно пользуемся, называя определенным образом те или иные ситуации.

Основная трудность (в данном случае — вполне преодолимая) состоит в том, чтобы эти совершенно обыденные категории перевести в ранг точных математических понятии. Подобный перевод весьма типичен для математики. Он даже имеет специальное название. **Когда мы переходим от расплывчатого и привычного понятия к точно формулируемому, то это последнее называется <u>экспликацией исходного.</u>**

Так, например, математическое понятие «алгоритм» есть экспликация такого обычного понятия как «метод решения задачи».

Возьмем еще пример, требующий большей математической эрудиции: понятие «производная», лежащее в основе дифференциального исчисления, есть не что иное, как экспликация интуитивно ясного понятия «скорость изменения данной величины».

Довольно очевидно, что, поскольку исходное понятие всегда бывает достаточно расплывчатым, оно допускает не одну экспликацию.

В сущности эта книга посвящена экспликации одного существенного понятия, а именно, «отношения», к его основных разновидностей. Что такое отношение, проще всего пояснить примерами. Следующие суждения в действительности выражают отношения между некоторыми объектами:

«Иван — брат Петра»,

«Иван — сосед Петра»,

«Железо тяжелее воды»,

«Киев южнее Москвы»,

«Ночь и день имеют одинаковое количество букв.Эти пять предложений выражают отношения разного типа. Однако можно заметить родство в характере отношений, утверждаемых первым, вторым и пятым предложениями. Все они говорят о том, что некие два объекта принадлежат общему классу: сыновей общих родителей, жителей одного дома или поселки, слов с фиксированным числом букв. Третье и четвертое отношения имеют то общее, что выражают относительный порядок объектов в системе. Когда мы говорим, что железо тяжелее воды, мы не предполагаем, что вещества делятся на категории легких и тяжелых. И не утверждаем, что железо тяжелое, а вода легкая. Свинец еще тяжелей железа, а водород гораздо легче воды. Точно так же, деление городов на южные и северные отнюдь не обязательно, чтобы четвертое предложение было справедливым. С точки зрения жителей Мурманска, Москва — это весьма и весьма южный город с черными ночами и созревающими фруктами, а для тбилисцев Киев имеет все основания считаться северным. Даже если бы мы предложили условное деление городов на южные и северные, то в каждой из групп можно было бы найти опять-таки более южных и более северных представителей.

В дальнейшем мы сможем четко определить эту интуитивно ощущаемую разницу между отношениями того и другого типа. Мы увидим, что первый, второй и пятый примеры — это отношения типа эквивалентности, определяющие разбиения объектов на классы подобных друг другу. Остальные два примера — это отношения типа порядка, устанавливающие относительное расположение объектов в системе.

Важно обратить внимание на следующее обстоятельство. Во всех пяти примерах четко выделяются названия объектов и названия отношений. Если вместо названия объекта подставить в предложение название другого объекта, то возможны следующие ситуации:

1) отношение опять будет выполнено;
2) отношение перестанет выполняться;
3) отношение потеряет смысл.

Так, если в третье предложение вместо слова «железо» мы подставим слово «медь», то суждение останется справедливым. Если в четвертое предложение вместо слова «Москва» подставить «Ташкент», то оно перестанет быть верным. Если же в четвертое предложение вместо Москвы поставить «железо», то суждение обратится в бессмыслицу. Аналогично, подставив в первое суждение объекты из четвертого, мы получим предложение «Киев — брат Москвы». Можно, конечно, понимать его в переносном смысле, но ясна, что слово «брат» тогда уже не будет значить «сын общих родителей». (Ср. выражение «Киев — мать городов русских».)

Любопытно, что в пятое предложение можно подставлять, казалось бы, любые объекты, поскольку для любого слова имеет смысл говорить о числе букв. Это объясняется тем, что слова «ночь» и «день» в этом предложении употреблены не как имена соответствующих явлений, а как имена самих себя. Более точно это предложение должно было бы звучать так:

«Слово „ночь" и слово „день" имеют одинаковое количество букв».

В таком виде уже ясно, что сама форма суждения ограничивает класс объектов —объектами отношения здесь могут быть только сами слова.

Итак, мы видим, что говорить об отношении можно только тогда, когда мы умеем выделять множество объектов, на которых это

отношение определено. Значит, прежде чем пытаться формализовать понятие отношения, нужно научиться формально говорить о множествах и их свойствах. Трудность состоит в том, что понятие множества является в математике «первичным»: его обычно не считают нужным определять через другие понятия. Более того, в полной теории множеств имеются парадоксы.

Мы не будем здесь излагать теорию множеств. Будем надеяться, что первоначальные понятия теории множеств читателю уже знакомы.

1. Введение в отношения

1.1. Основные понятия

Фундаментальным понятием дискретной математики является понятие *отношения*, которое используется для обозначения связи между объектами или понятиями.

В математике, как и в жизни, разные объекты могут иметь какое-то отношение к другим объктам или не иметь. Например, родственные отношения, дружеские отношения, дипломатические отношения, равноправные отношения.

Отношение можно задать в виде неполных предложений - предикатов, например, „менше чем ...”, „больше чем ...”, „эквивалентно”, „конгруэнтно” и т.п.

Смотря на многочисленные примеры вокруг, мы делаем вывод, что **отношения отличаются от соответствий тем, что определяются на одном множестве**.

Не имеет смысла говорить об отношениях между студентами и оценками. О дипломатических, родственных или любых других отношениях между должностью и зарплатой. Для определения конкретного отношения необходимо определить множество, и пары, для которых имеет место данное отношение. Например, на множестве людей отношения "быть братом", "учиться в одной группе" или "быть выше ростом".

Естественно, что при соответствующем выборе отношения его аргументы могут быть связаны довольно просто. Они не обязательно должны быть связаны какой-нибудь простой или очевидной формулой, хотя в ситуациях, когда нужно осуществить некоторые вычисления, иногда можно найти удачное описание отношений.

Перед тем как подойти к этому вопросу с математической позиции, рассмотрим несколько случаев, которые возникают из рассмотрения следующей простой ситуации (которая также приводит к возникновению понятий отношения). Предположим, что для некоторой конечной вычислительной машины мы имеем множество программ P, конечное множество значений данных D и множество результатов R. Если мы выберем конкретное значение из D, то оно может использоваться в некоторых программах из P, и для каждой программы из P существует совокупность значений их D, которые в ней используются. Таким образом, мы имеем соответствие между значениями данных и программами, и, следовательно, существуют элементы в $D\times P$, которые представляют интерес. Аналогично, если мы сведем рассмотрение к $p\in P$, то p связывает соответствующие значения данных из D с результатами из R. Можно рассмотреть данные, приводящие p к остановке, или результаты, которые не могут быть получены из p. Следовательно, мы приходим к подмножеству $D\times R$. (При переработке данных от D к R возникают некоторые ассоциации, которые могут оказаться полезными для запоминания терминологии.)

Задать отношение — это значит указать: между какими объектами оно выполняется. Например, отношение «быть братом» полностью определено, если мы составим список всех пар людей таких, что один из них — брат второго.

Заметим, что мы здесь заранее выбрали множество объектов, между которыми определяется отношение. Именно, отношение «быть братом» мы полагаем заданным на множестве людей. Рассмотрим несколько простых примеров. Предположим, что Татьяна, Александр и Михаил — дети одних и тех же родителей, перечисленные в порядке старшинства. Тогда на этом множестве из трех людей отношение «быть братом» выполнено для следующих пар:

«Александр — брат Татьяны», «Александр — брат Михаила»,

«Михаил — брат Татьяны» «Михаил — брат Александра».

В первом и третьем суждении объекты нельзя поменять местами. Это означает, что отношение «быть братом», вообще говоря, не симметрично. Если «х — брат *у»,* то «у — брат *х»* только в том случае, когда *у* — мужчина. Полезно обратить внимание, что отношение «Александр — брат Александра» не выполнено, т. е. данное отношение, как принято говорить, не рефлексивно. По этому поводу можно напомнить старую загадку: «Сын моего отца, а мне не брат. Кто

это такой?».Ответ теперь ясен: «Я сам». Отношение «быть старше» выполнено на том же множестве для следующих пар:
«Татьяна старше Александра»,
«Татьяна старше Михаила», «Александр старше Михаила».
Следующий пример показывает, что отношения можно устанавливать и между объектами разных множеств. Рассмотрим множество M_1 учащихся некоторой школы и множество M_2 учителей той же школы. Тогда существует естественное отношение «x — ученик y», где x — один нз учащихся (элемент множества M_1, а y— один из учителей (элемент множества M_2). Ясно, что для одного и того же ученика x это отношение может выполняться при разных y. И, наоборот, один и тот же учитель y имеет разных учеников.
Отношение может быть определено не только для пар объектов *(бинарные* отношения), но и для троек, четверок и т. д. Например, отношение «образовывать футбольную команду» выполняется для некоторых групп из 11 людей. Оно задается списками основных составов футболистов, участвующих в различных матчах. Это отношение не следует путать с бинарным отношением «входить в одну футбольную команду». Действительно, два игрока из одной команды еще не образуют команды. Команду может образовывать только комплект из 11 игроков.
Хороший пример трехместных (или, как любят говорить математики, *тернарных)* отношений доставляют алгебраические операции. Например, отношение «образовывать сумму» имеет смысл для троек чисел (x,y,z) и выполняется в том случае, когда

$$x + y = z.$$

Пропорциональность чисел x, y, z, u:

$$\frac{x}{y}=\frac{z}{u}$$

есть отношение, выполненное для некоторых четверок чисел (x,y,z,u).

Мы будем изучать в основном бинарные отношения, т. е. отношения, которые выполняются (или не выполняются) между двумя объектами.

Перейдем теперь к формальным описаниям отношений.

1. Бинарные отношения. Многие задачи математики, техники и других областей человеческой деятельности получают удобную интерпретацию языком теории отношений. **Все арифметические операции - это собственно говоря некоторые отношения между числовыми множествами.** Множество деталей остается складским имуществом до тех пор, пока между ними не реализуются определенные отношения, которые превращают эти детали в какой-нибудь механизм или устройство (телевизор, станок, дом, мост и т.п.).

Разнообразные отношения складываются между людьми — родители и дети, начальники и подчиненые, учителя и ученики.

Отношение между элементами двух множеств, т.е. *бинарные отношения* устанавливают соответствие элементов одного множества X элементам другого множества Y. Ясно, что такое отношение может быть задано некоторой совокупностью ***упорядоченных пар*** *(х, у),* которые являются элементами множества $X \times Y$. Это совсем не значит, что всегда нужно перечислять все такие пары. Часто отношение задается некоторым свойством, выраженным в словесной или символической форме.

Если A — отношение, то *соотношение* xAy можно записать также в виде $(x, y) \in A$, где $A \subset X \times Y$. Например, выражение $3 < 7$ и $(3,7) \in <$ означает одно и тоже, но первое из них привычнее. В то же время $(7,3) \in <$ означало бы $7<3$, что неверно. Таким образом, в общем случае переставлять элементы в паре (x, y) нельзя, что и подчеркивается названием этой пары — *упорядоченной*. Элемент x называют *первой координатой,* а элемент y - второй *координатой* упорядоченной пары.

2. Области определения и значений. Множество первых координат x является *областью определения (левой областью)* $D_o(A)$, а множество вторых координат — *областью значений (правой областью)* $D_З(A)$ отношения A. Если $x \in X$ и $y \in Y$, то $D_o(A) \subset X$ и $D_З(A) \subset Y$. В таких случаях говорят, что A есть *отношение от* X *к* Y. Его называют также *соответствием* и обозначают $X \rightarrow Y$. Если $Y=X$, то любое отношение xAy является подмножеством множества $X \times X$ и называется *отношением в* X.

Пусть, например, $X = \{2, 3\}$ и $Y = \{3, 4, 5, 6\}$. Произведение этих множеств $X \times Y = \{(2, 3), (2, 4), (2, 5), (2, 6), (3, 3), (3, 4), (3, 5), (3, 6)\}$. Отношение «быть делителем» есть множество $A = \{(2, 4), (2, 6), (3, 3), (3, 6)\}$, отношение = есть множество $Y = \{(3, 3)\}$, а отношение > есть пустое множество Ø. Области определения и значений отношения A — это соответственно множества $D_o(A)=\{2,3\}=X$ и $D_З(A) = \{3, 4, 6\} \subset Y$.

Если область определения отношения совпадает с некоторым множеством X, то говорят, что отношение *определено на* X. Подобный случай имеет место в приведенном выше примере отношения A «быть делителем». Очевидно, для отношения включения $\subset$ подмножеств универсума U областью определения и областью значений служит множество подмножеств $P(U)$ этого универсума.

Заслуживают внимания три частичных случая отношений в X:

1) *полное (универсальное) отношение* $P=X\times X$, которое имеет место для каждой пары (x_1, x_2) элементов из X (например, отношение «работать в одном отделе» на множестве сотрудников данного отдела);

2) *тождественное (диагональное) отношение* E, равносильное $x=x$ (например, равенство на множестве действительных чисел);

3) *пустое отношение,* которому не удовлетворяет ни одна пара элементов из X (например, отношение «быть братом» на множестве женщин). Очевидно, для любого отношения A в X справедливо $\varnothing\subset A\subset P$.

3. Сечение. Рассмотрим отношение $A\subset X\times Y$; если $x_i\in X$, то *сечение* по x_i отношения A, обозначаемое $A(x_i)$, есть множество $y\in Y$ таких, что $(x_i, y)\in A$. Множество всех сечений отношения A *называют фактором-множеством* Y по отношению A и обозначают Y/A. Оно целиком определяет отношение A.

Пусть например, $X = \{x_1, x_2, x_3, x_4, x_5\}$, $Y = \{y_1, y_2, y_3, y_4\}$ и $A = \{(x_1\ y_1), (x_1, y_3), (x_2, y_1), (x_2\ y_3), (x_2, y_4), (x_3, y_1, (x_3, y_2), (x_3, y_2), (x_4, y_3), (x_5,y_2), (x_5, y_4)\}$. Очевидно, $A(x_1)= \{y_1, y_3\}$; $A(x_2) = \{y_l, y_3, y_4\}$ и т.п. Если записать под каждым элементом из X соответствующее сечение отношения A, то элементы второй строки образуют фактор-множество Y/A:

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ \{y_1, y_3\} & \{y_1, y_3, y_4\} & \{y_1, y_2, y_4\} & \{y_3\} & \{y_2, y_4\} \end{pmatrix}.$$

Объединение сечений по элементам некоторого подмножества $B\subset X$ является сечением $A(Y)$ этого подмножества, т.е. $A(B)=\square_{x\in B} A(x)$. Так, $A = (x_2, x_3)= A(x_2)\cup A(x_3)= \{y_1, y_2, y_3, y_4\}$.

4. Графические представления.

При решении задачи на первом этапе часто полезно начертить «рисунок» для того, чтобы более ясно увидеть компоненты задачи. Особенно это полезно для описания отношений, так как записанные в виде множества упорядоченных пар отношения нелегко расшифровываются.

Отношение — это множества, обладающие определенной структурой; их элементы имеют несколько компонент, и поэтому, в принципе, мы можем использовать диаграммы Венна для их изображения. Хотя этим методом и можно пользоваться, особенно при описании некоторых больших множеств чисел, но существуют методы, которые более эффективны в общих ситуациях (которые включают, в частности, бинарные отношения на небольших множествах). В этом пункте мы коротко рассмотрим некоторые из них. Для описания этих методов используем множество

$$X = \{a, b, c, d\}$$

и отношения E_x, P_x и A, где

$$A = \{(a, b), (a, c), (b, d), (c, e), (e, b)\}.$$

Вначале рассмотрим метод, который относится к традиционной аналитической геометрии. Начертим пару взаимно перпендикулярных осей (OX — горизонтальная ось, а OY - вертикальная ось) и на каждой отметим точки, представляющие элементы множества X (рис. 1.1, *а*).

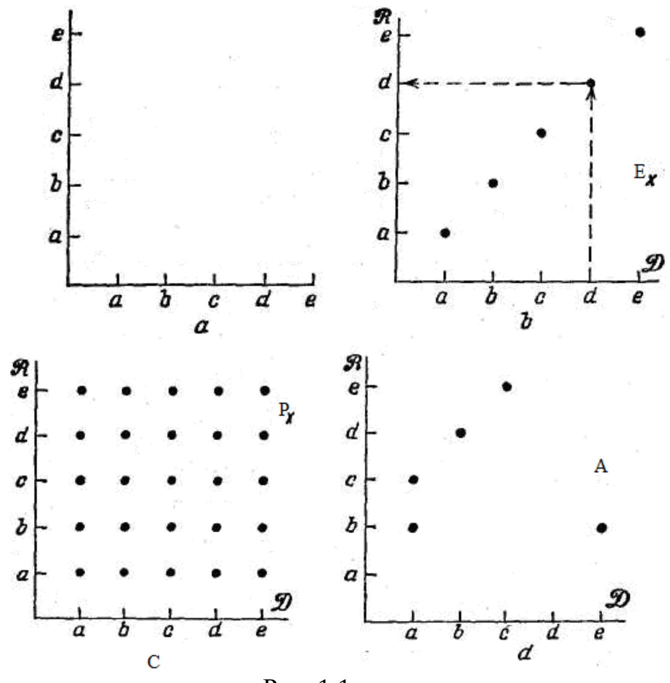


Рис. 1.1.

Теперь в правом верхнем координатном углу отметим точки с координатами (x, y), в которых $x \in X$, и $y \in Y$.

Множества, соответствующие E_x, P_x и A, изображены на рис. 1.1, b, c и d.

Основной недостаток этого метода заключается в том, что при увеличении $|X|$ трудно увидеть элементы в области и установить соответствие с точками, которые обозначают отношения. Чтобы преодолеть этот недостаток, можно опустить точки и соединить стрелкой $x \in \mathfrak{D}$, и $y \in \mathfrak{R}$, когда *(x, y)* принадлежит отношению (рис.1.2).

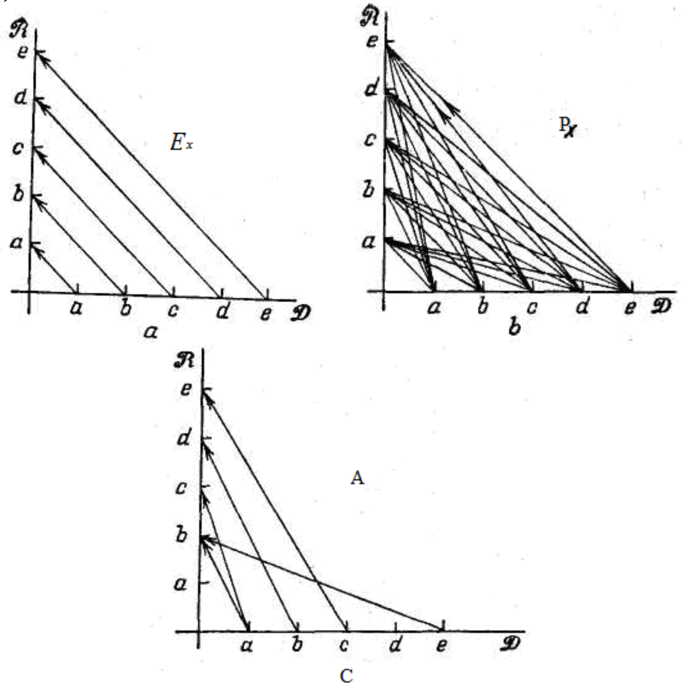
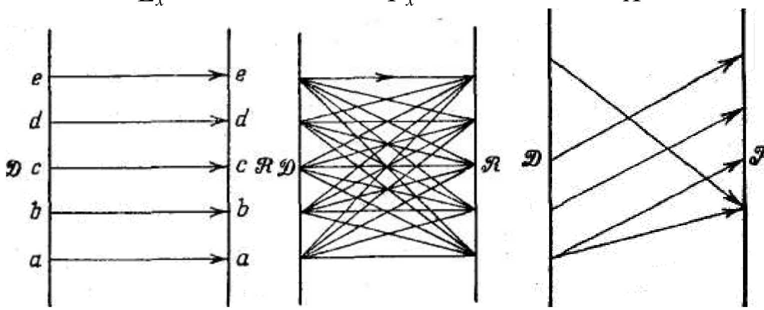


Рис. 1.2.

Диаграмма, которая представляет P_x, получилась довольно запутанной, но это естественно, поскольку число элементов в P_x увеличилось.

С другой стороны, отношение E_x и A представлены наглядно, и легко увидеть их области определения и значений. Диаграмма для P_x наиболее неудобна в месте пересечения осей. Теперь, когда не используются координаты в областях определения и значений для расстановки элементов, отношение (как в первом методе) можно начертить параллельными. Поэтому, используя параллельные вертикальные линии и двигаясь слева направо (линия слева является областью определения), мы получаем диаграмму, которая изображена на рис. 1.3. Здесь стрелки не нужны, так как мы знаем, что отношения идут от области определения к области значений.

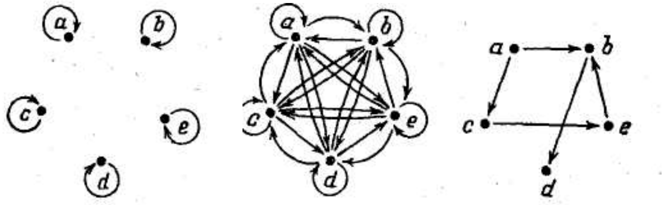
E_x P_x A



a *b* *c*

Рис. 1.3.

Это приводит к двух возможностям: мы можем или заменить стрелки прямыми линиями, или заменить две линии, которые изображают области определения и значений, простой совокупностью точек. (Например, точка *c* в области определения есть той же самой, что и точка, которая представляет *c* в области значений). Это показано на диаграмме, которая изображена на рис. 1.4.



E_x P_x A

a *b* *c*

Рис. 1.4.

Итак, нами обозначены наиболее важные методы графического изображения бинарных отношений.

5. Матрица и граф отношения. Из предыдущего ясно, что конечное отношение можно представить с помощью фактор-множества. Другой способ — *матричный* — основанный на представлении отношения соответствующей ему прямоугольной таблицей (матрицей). Ее столбцы соответствуют первым координатам, а строки — вторым координатам. На пересечении i-го столбца и j-й строки ставится единица, если выполнено соотношения x_iAy_j, и нуль, если это соотношение не выполняется (нулевые клетки можно оставлять пустыми). Эта матрица содержит всю информацию об отношении A. Например, для отношения, рассмотренного в (3), матрица имеет вид:

	x_1	x_2	x_3	x_4	x_5
y_1	1	1	1		
y_2			1		1
y_3	1	1		1	
y_4		1	1		1

Ненулевые элементы i-го столбца указывают на совокупность элементов $y \in Y$, представляющую собой сечение $A(x_i)$, например,

$$A(x_5) = \{y_2, y_4\}.$$

Полному отношению соответствует матрица, все клетки которой заполненные единицами, тождественному — единичная матрица, а пустому — нулевая матрица.

Отношения можно также задавать (или изображать) с помощью *ориентированного графа.* Вершины графа соответстуют элементам множеств X и Y, а дуга, направленная из вершины x_i к y_j, означает, что x_iAy_j. Граф отношения из (3) показан на рис. 1.5.

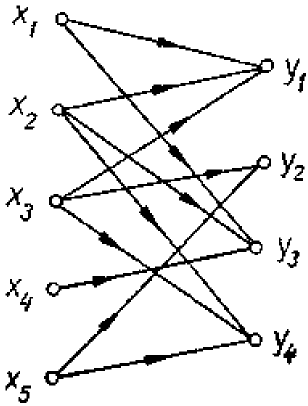


Рис. 1.5. Граф отношения от X к Y (биграф)

Отношение в X отображается графом с вершинами, соответствующими элементам этого множества. При этом, если имеют место соотношения x_iAx_j, и x_jAx_i, то вершины связываются двумя противоположно направленными дугами, которые можно условно заменять одной ненаправленной дугой (или указывать противоположные направления на одной дуге). Соотношению x_iAx_i соответствует петля, выходящая из x_i и входящая в эту же вершину. На рис. 1.6 показан граф отношения в X, заданного множеством $A = \{(x_1, x_2), (x_1, x_5), (x_2, x_1), (x_2, x_3), (x_3, x_1), (x_3, x_4), (x_4, x_3), (x_4, x_5), (x_5, x_5), (x_6, x_2)\}$.

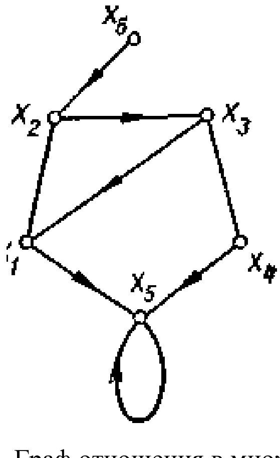


Рис 1.6. Граф отношения в множестве X.

Графы полного, тождественного и пустого отношений изображены на рис. 1.7 (для пустого отношения граф состоит из изолированных вершин).

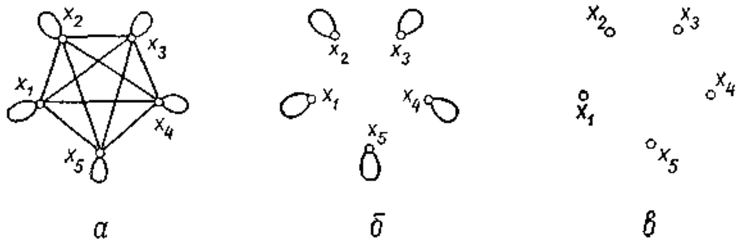


Рис. 1.7. Граф полного (*а*), тождественного *(б)* и пустого (*в*) отношений

6. Симметризация отношения. Так как отношение — это множества, то над ними можно выполнять все теоретико-множественные операции. Кроме этого, определяются специфические для отношений операции: *обращение (симметризация)* и *композиция.*

Отношения, *симметричное (обратное)* некоторому отношению $A\subset X\times Y$, обозначается через A^{-1} и представляет собой подмножество множества $Y\times X$, образованное теми парами $(y,\ x)\in Y\times X$, для которых $(x,\ y)\in A$. Переход от A к A^{-1} осуществляется взаимной перестановкой координат каждой упорядоченной пары. Так, обратное отношение для «*x* есть делитель *y*» будет «*y* делится на *x*» и для приведенного в (2) примера выражается множеством $\{(4, 2), (6, 2), (3, 3), (6, 3)\}$.

При переходе от A до A^{-1} область определения становится областью значений, и наоборот. Матрица обратного отношения получается транспонированием исходной матрицы.

7. Композиция отношений. Пусть даны три множества X, Y, Z и два отношения $A\subset X\times Y$ и $B\subset Y\times Z$. *Композиция отношений* A и B есть отношения C, состоящее из всех тех пар $(x,\ z)\subset X\times Z$, для которых существует такое $y\in Y$, что $(x,\ y)\in A$ и $(y,\ z)\in B$.

Сечение отношения C по x совпадает с сечением отношения B по подмножеству $A(x)\subset Y$, т.е. $C(x)=B(A(x))$.

Рассмотрим, например, два отношения: A из примера (3) и

$$B=\{(y_1, z_2), (y_2, z_1), (y_2, z_2), (y_3, z_3), (y_4, z_3)\}\ .$$

Очевидно,

$$C=\{(x_1, z_2), (x_1, z_3), (x_2, z_2), (x_2, z_3), (x_3, z_1), (x_3, z_2), (x_3, z_3), (x_4, z_4), (x_5, z_1), (x_5, z_3)\}$$

Сечение

$$C(x_3)=\{z_1, z_2, z_3\}.$$

С другой стороны

$$B(A(x_3))=B(\{y_1, y_2, y_4\})=\{ z_2\} \square \{ z_1, z_2\} \cup \{ z_3\}=\{ z_1, z_2, z_3\}.$$

Теперь нужно решить вопрос, как записать композицию отношений A и B. Если исходить из соотношений xAy и yBz, то естественно записать $xCz=xABz$, т.е. $C=AB$. Но при этом соотношение $C(x)=Y(A(x))$ приняло бы неудобную форму $(AB)(x)=Y(A(x))$. Поэтому композицию C отношений A и B обычно записывают как $C=BA$ (или $C=B^{o}A$) тогда $(BA)(x)= B(A(x))$.

Как увидим дальше, такая запись имеет и другие преимущества.

Композиция отношений обладает ассоциативным законом, т.е.

$$D(BA)=(DB)A=DBA,$$

но не коммутативна

$$(BA\neq AB).$$

Можно также показать, что

$$(BA)^{-1} =A^{-1}B^{-1}.$$

8. Представление композиции отношений матрицами и графами.

Композиция отношений

$$A\subset X\times Y \text{ и } B\subset Y\times Z.$$

наглядно представляется с помощью графов. Прежде всего необходимо к графу отношения A достроить граф отношения B.

Граф отношения $C=BA$ получим, исключив вершины, соответствующие элементам множества Y. При исключении вершины y, каждый проходящий через нее путь от вершин x к вершинам z заменяется одной дугой с тем же направлением. Параллельные ветви с одинаковыми направлениями соответствуют одинаковым парам в C и рассматриваются как одна ветвь. Граф композиции отношений из (7) показан на рис. 1.8.

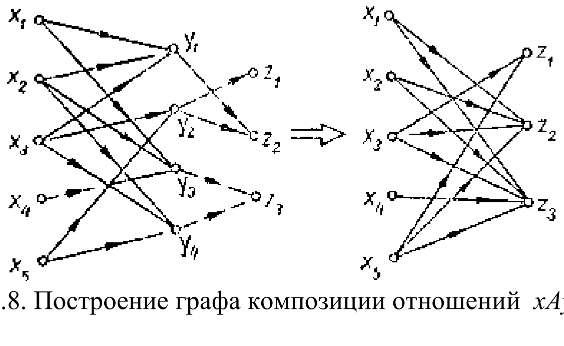


Рис 1.8. Построение графа композиции отношений xAy и yBz

Матрица композиции C=BA представляется как произведение матриц отношений B и A (в порядке их следования), которое выполняется по обычному правилу умножения прямоугольных матриц с последующей заменой отличного от нуля элемента результирующей матрицы единицей. Так, для рассматриваемого примера имеем:

$$\begin{array}{c|cccc|} & y_1 & y_2 & y_3 & y_4 \\ \hline z_1 & & 1 & & \\ z_2 & 1 & 1 & & \\ z_3 & & & 1 & 1 \\ \hline \end{array} \; \begin{array}{c|ccccc|} & x_1 & x_2 & x_3 & x_4 & x_5 \\ \hline y_1 & 1 & 1 & 1 & & \\ y_2 & & & 1 & & 1 \\ y_3 & 1 & 1 & & 1 & \\ y_4 & & 1 & 1 & & 1 \\ \hline \end{array} = \begin{array}{c|ccccc|} & x_1 & x_2 & x_3 & x_4 & x_5 \\ \hline z_1 & & & 1 & & 1 \\ z_2 & 1 & 1 & 2 & & 1 \\ z_3 & 1 & 2 & 1 & 1 & 1 \\ \hline \end{array} \rightarrow$$

$$\rightarrow \begin{array}{c|ccccc|} & x_1 & x_2 & x_3 & x_4 & x_5 \\ \hline z_1 & & & 1 & & 1 \\ z_2 & 1 & 1 & 1 & & 1 \\ z_3 & 1 & 1 & 1 & 1 & 1 \\ \hline \end{array}$$

Приведенное выше правило легко доказывется на основе выражения для элемента

$$c_{ik}=b_{i1}a_{1k}+ b_{i2}a_{2k}+\ldots+b_{in}a_{nk}=\sum_{i=1}^{n} b_{ij}a_{jk}$$

произведения матриц, соответствующих отношениям B и A. В этом выражении слагаемое $b_{ij}a_{jk}$ равно единице при условии, что $a_{jk}=b_{jk}=1$, а это возможно только, если имеют место соотношения x_iAy_j и y_jBz_k, т.е. x_iBAz_k. Если в выражении для c_{ik} не одно, а несколько единичных слагаемых, то каждое из них соответствует одному и тому же соотношению x_iBAz_k и их сумма должна быть заменена единицей (это соответствует замене нескольких одинаково направленных дуг графа одной дугой).

Если A и B — отношения в X, то графы этих и граф композиции $C = BA$ строится на множестве вершин, соответствующих $x_i \in X$, по тому же правилу, что и в общем случае. В приведенном на рис. 1.9 примере сплошные дуги соответствуют отношению A, а пунктирные — отношению B (справа показан граф композиции $C = BA$).

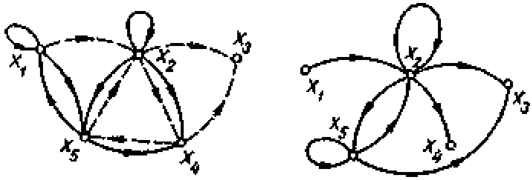


Рис. 1.9. Граф композиции отношений в множестве X

1.2. Общие свойства отношений.

Рассмотрим некоторые из основных свойств, которыми могут быть наделены отношения. Говорят, что **свойство имеет место, если выполняется соответствующее условие**.

Пусть A — бинарное отношение в множестве X. Определим общие свойства таких отношений, которые должны выполняться для всех $(x_i, x_j) \in A$. Говорят, что $A \subset X \times X$:

1) *рефлексивно,* если $A \supset E$ (E — тождественное отношение), т.е. оно всегда выполняется между объектом и им самим: xAx (равенство, самообслуживание);

2) *антирефлексивно,* если $A \cap E = \varnothing$, т.е. может выполняться только для несовпадающих объектов: из $x_i \, A \, x_j$ следует $x_i \neq x_j$ (строгое неравенство, «быть старше»);

3) *симметрично,* если $A = A^{-1}$, т.е. при выполнении соотношения x_iAx_j выполняется и соотношение x_jAx_i (расстояние между двумя точками, «быть братом»);

4) *асимметричено,* если $A \cap A^{-1} = \varnothing$, т.е. из двух соотношений x_iAx_j и x_jAx_i по меньшей мере одно не выполняется (строгое включение, «быть отцом»); если отношение асимметрично, то оно и антирефлексивно;

5) *антисимметрично,* если $A \cap A^{-1} \subset E$, т.е. оба соотношения x_iAx_j и x_jAx_i выполняются одновременно только тогда, когда $x_i = x_j$ (нестрогое неравенство $\leq$, включение);

6) *транзитивно,* если $AA \subset A$, т.е. из x_iAx_j и $x_j \, A \, x_k$ следует x_iAx_k («быть делителем», «быть родственником»).

Для рефлексивного отношения все элементы матрицы на главной диагонали — единицы, а для антирефлексивного — нули. Симметричность отношения влечет и симметричность матрицы, асимметречность отношения — несимметричность матрицы с нулевыми элементами на главной диагонали, антисимметричность отношения — только несимметричность матриц. В матрице транзитивного отношения для каждой пары единичных элементов, один из которых расположен в i-м столбце и j-й строке, а другой в j-м столбце и k-й строке, обязательно существует единичный элемент, расположенный в клетке на пересечении i-го столбца и k-й строки (наличие единичных элементов на главной диагонали не нарушает транзитивности) .

Граф рефлексивного отношения содержит петли у всех вершин, антирефлексивного — не содержит ни одной петли. Для симметричного отношения вершины графа могут быть связаны только парами противоположно направленных дуг (ненаправленными ребрами). В графе асимметричного отношения петли отсутствуют, а вершины могут быть связаны только одной направленной дугой. В случае антисимметричного отношения могут быть петли, но связь между вершинами, если она имеется, также отображается только одной направленной дугой. Наглядно проявляется на графе и свойство транзитивности: если через некоторую совокупность вершин графа проходит путь, то существуют дуги, соединяющие любую пару вершин из этой совокупности (рис. 1.10, *а).*

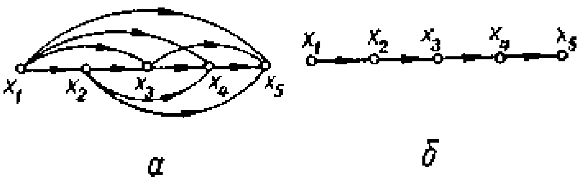


Рис.1.10. Граф транзитивного отношения (*а*) и его упрощенное изображение (б)

Обычно в графе транзитивного отношения изображают только этот путь, а обусловленные транзитивностью дуги опускают (рис. 1.10, *б).* Такой упрощенный граф называют *графом редукции* .

В более общей ситуации мы можем интерпретировать рассмотренные выше характеристики отношения путем построения диаграммы:

а) отношение рефлексивно тогда и только тогда, когда для каждого узла (точки) на диаграмме существует стрелка, которая начинается и заканчивается на этом узле;

б) отношение симметрично тогда и только тогда, когда для каждой стрелки, которая соединяет два узла, существует также стрелка, которая соединяет эти узлы в обратном направлении;

в) отношение транзитивно тогда и только тогда, когда для каждой пары узлов x и y, связанных последовательностью стрелок от x к a_1, от a_1 к a_2, ..., от a_{n-1} к a_n, от a_n к y, существует также стрелка от x к y;

г) отношение антисимметрично тогда и только тогда, когда не существует двух различных узлов, связанных парой стрелок.

Существует много других свойств отношений, которые можно было бы рассмотреть. Однако рассмотренные выше свойства являются наиболее важными и будут часто использоваться в дальнейшем.

Многоместные отношения. Отношение может быть определено не только для пар объектов, но и для троек, четверок и т.д.

Отношение n объектов *(n-местное отношение)* определяется как множество n-мерных векторов $(x_1, x_2, ..., x_n)$, являющееся подмножеством произведения

$$X_1 \times X_2 \times ... \times X_n,$$

причем

$$x_1 \in X_1, x_2 \in X_2, ..., x_n \in X_n.$$

Многомерные векторы можно определить в терминах упорядоченных пар, например тройка (x_1, x_2, x_3) рассматривается как упорядоченная пара $((x_1, x_2), x_3)$, где первая координата (x_1, x_2) сама является упорядоченной парой, причем

$$(x_1, x_2) \in X_1 \times X_2.$$

Вообще, n-мерный вектор выражается как упорядоченная пара через $((x_1, x_2, ..., x_{n-1}), x_n)$, если определено $(x_1, x_2, ..., x_{n-1})$.

Примером трехместных (тернарних) отношений являются: арифметические операции над числами, отношение между родителями и детьми (отец, мать, ребенок) и т.п. Пропорция $x : y = z : u$ иллюстрирует четырехместное отношение. Множество, которое задается с помощью общего свойства элементов из некоторого универсума, можно рассматривать как *одноместное (унарне)* отношение в этом универсуме.

1.3. Функции как отношения

Частным случаем отношений можно считать и функции. Пусть отношение A на множестве M таково, что для всякого $x \in M$ существует ровно один элемент $y \in M$, для которого справедливо соотношение xAy. Тем самым каждому элементу $x \in M$ сопоставляется некоторый элемент $y \in M$, определенный этим условием. Такое отношение называется *функцией,* или *отображением* (или *однозначным соответствием),* а сам элемент $y \in M$, соответствующий элементу $x \in M$, называется *значением функции А на элементе х.* Эта зависимость между x и y выражается обозначением

$$y = A(x)$$

Множество A тех пар (x, y), для которых выполнено соотношение xAy, называется *графиком* функции.

Например, если M — числовая прямая, а отношение A есть отношение равенства $y = x$, то график состоит из всех точек вида (x, x) и является биссектрисой координатного угла, т. е. обычным графиком функции $y = x$. Если отношение A выполнено для тех пар, для которых $y = \sin x$ (ясно, что для каждого x существует единственное число y, обладающее этим свойством), то график этой функции есть обычная синусоида.

Итак, наше определение графика является обобщением обычного графика числовых функций.

В данном случае очень полезно рассмотреть отношения на таких парах (x, y), где x принадлежит множеству M, а y — другому множеству L. Отношение α этого вида мы снова будем называть *функцией,* или *отображением,* если для каждого $x \in M$, существует единственный элемент $y \in L$, для которого выполнено соотношение $x\alpha y$. Такую функцию α мы будем символически записывать как α: $M \to L$; здесь M называется *областью отправления* функции α, а L — ее *областью прибытия.* Отображение α: $M \to L$ называется также *отображением множества М в множество L.* Элемент множества L, который при этом соответствует элементу x из M, обозначается $\alpha(x)$ и называется *образом* элемента x. Сам элемент x называется *прообразом* элемента $\alpha(x)$. Из определения отображения α: $M \to L$ явствует, что каждый элемент $y \in M$ имеет ровно один образ. Но не всякий элемент $y \in L$ обязан иметь прообраз. Если же такой прообраз существует, то он не обязан быть единственным.

Пример 1. Пусть M — множество людей, а L — множество натуральных чисел. Пусть α: $M \to L$ — отображение, которое каждому человеку ставит в соответствие его рост, выраженный в сантиметрах (округленный, как принято, до целочисленного значения). Ясно, что каждому человеку соответствует определенный рост, но значение роста: 400 *см* — не соответствует никакому человеку. С другой стороны, существует масса людей, у которых рост— 172 *см.*

Пример 2. Пусть M — множество ныне живущих людей, L — множество всех людей, а отображение α: $M \to L$ ставит в соответствие каждому человеку его отца. Ясно, что у каждого $x \in M$ имеется единственный образ. Но не у всякого $y \in L$ есть прообраз, так как далеко не всякий человек является отцом другого. Например, если *y*— женщина. Кроме того, несколько человек могут иметь общего отца.

Отображение α: *M*→*L* называется *сюръективным,* если любой элемент *y* из *L* имеет прообраз. В этом случае говорят еще, что *M* отображается *на L*.

Например, пусть *M* —множество всех русских слов, *L* — множество частей речи русского языка, а отображение α: *M*→*L* сопоставляет каждому слову часть речи, к которой оно принадлежит. Ясно, что каждая часть речи соответствует по крайней мере одному слову— примеру на эту часть речи. (Мы здесь считаем, что грамматические омонимы каким-то образом уже различены, т. е. про слово «печь» известно, глагол это или существительное.)

Отображение α: *M*→*L* называется *инъективным,* если для каждого элемента $y \in L$ существует не более одного прообраза.

Например, пусть *M* — множество людей, стоящих в некоторой очереди, *L* — множество натуральных чисел, а отображение α: *M*→*L* сопоставляет каждому находящемуся в очереди его порядковый номер. Ясно, что каждый номер может быть присвоен лишь одному человеку. С другой стороны, это отображение не сюръ-ектнвно, так как существуют номера, не присвоенные никому.

Если отображение α: *M*→*L* одновременно сюръективно и инъективно, то оно называется *биективным.* Множества *M* и *L*, для которых существует биективное отображение α: *M*→*L* называются *равномощными.* Легко убедиться, что если *M* конечно, а *M* н *L* равномощны, то *M* содержит такое же число элементов, как *L*. Для этого достаточно пересчитать все элементы из *M;* если элементу $x \in M$ приписан номер *n(x),* то тот же номер следует приписать его образу α*(x).* Так как отображение сюръективно, то все элементы множества *L* получат номера. Так как отображение инъективно, то каждый элемент из *L* получит единственный номер. Тем самым для пересчета элементов множества *L* требуется ровно столько же номеров, сколько для пересчета элементов множества *M.* Легко сообразить, что количество элементов не зависит от способа нх пересчета.

Для бесконечных множеств равномощпость естественно принять в качестве обобщения понятия «иметь одинаковое количество элементов».

Полезно ввести еще такие понятия.

Пусть α: *M*→*L* и M_1 — подмножество множества *M.* *Образом множества* M_1 (обозначается через α(M_1)) мы назовем множество всех образов $\{\alpha(x)\}$, где $x \in M_1$. В частности, α*(M)* есть образ всего *M.*

Легко видеть, что $\alpha : M \to \alpha(M)$ есть сюръективное отображение. Аналогично, если $L_1 \subseteq L$, то *полным прообразом множества* L_1 (обозначается через $\alpha^{-1}(L_1)$) называется объединение прообразов всех элементов, входящих в L).

Определим теперь так называемое *единичное отображение* множества *M:*

$$\varepsilon_M : M \to M,$$

которое каждому элементу $x \in M$ сопоставляет этот самый элемент. (Легко видеть, что единичное отображение ε — это то же самое, что и диагональное отношение *E.)*

Пусть α: *M*→*L* отображение β: *M*→*L* называется *обратным* к α, если $\alpha\beta = \varepsilon_M$ и $\beta\alpha = \varepsilon_L$, т. е. если отображение β переводит любой образ α*(x)* в *x,* а отображение α переводит любой образ β(*y*) в *y*. В этом случае мы будем писать: $\beta = \alpha^{-1}$. Читатель легко убедится в том, что для существования отображения, обратного к α, необходимо и достаточно, чтобы α было биективным.

Иногда бывает удобно рассматривать функции α: *M*→*L,* которые определены не на всем *M*, а только на некотором его подмножестве M_1 которое тогда называется *областью определения* функции. Тогда удобно пополнить множество L до множества $L_\# = L \cup \{\#$, т. е. прибавить к L элемент $\#$, не входивший ранее в *L*. Элемент $\#$ играет роль как бы *пустого* элемента. В таком случае мы считаем (опять-таки, по определению), что отображение $\alpha : M \to L_\#$ ставит в соответствие любому элементу из $M \setminus M_1$ пустой элемент $\#$. Часто бывает удобно рассуждать так, как будто пустой элемент $\#$ заранее содержится в любом множестве. Тогда не нужно различать *L* и $L_\#$.

Пример 1. *M* — некоторое множество людей в данный момент времени, а *L* — множество их головных уборов. Функция α: *M*→*L* сопоставляет каждому человеку надетый на нем головной убор. Ясно, что функция α определена лишь на подмножестве множества *M,* состоящем из людей, у которых что-то надето на голову. Остальным, простоволосым, сопоставляется пустой головной убор.

Пример 2. *M* — множество русских словоформ, а *L* — множество русских окончаний. Функция α сопоставляет каждой словоформе ее окончание:

бежать — ать
окно — о
столом — ом

Словоформам «стол», «пальто», «вместе» соответствуют нулевые (пустые) окончания. (Иногда букву «о» в слове «пальто» неграмотные люди воспринимают как окончание именительного падежа среднего рода и пытаются склонять это слово по падежам. Но это вовсе не русское окончание, а часть французской основы «Paletot».)

Иногда и для произвольного отношения (A, M, L) об элементах y таких, что xAy, удобно говорить как об элементах, сопоставленных, или соответствующих, элементу x. В этих случаях отношение (A, M, L), которое приобретает, так сказать, функциональный характер, мы будем называть *соответствием.* Таким образом, соответствие — это «многозначная» функция. Запись ψ: $M \rightarrow L$ для произвольного отношения $\psi = \langle A, M, L \rangle$ и будет как раз означать, что мы рассматриваем отношение ψ как соответствие.

Можно было бы, разумеется, вместо соответствия ψ: $M \rightarrow L$ рассматривать функцию

$$\alpha\colon\ M \rightarrow 2^L,$$

которая каждому элементу $x \in M$ сопоставляет множество $L_x \subseteq L$ всех тех y, для которых $\langle x, y \rangle \in A$ (L_x может быть, в частности, пустым); однако язык соответствий (= «многозначных функций») часто бывает более удобным.

Как и в случае «однозначных» функций, можно ввести понятия *всюду определенного соответствия* (для любого $x \in M$ множество L_x непусто), *инъективного соответствия* (для любых $x \neq y$ $L_x \cap L_y = \varnothing$) и *сюръекгивного соответствия* (для всякого $y \in L$ существует $x \in M$, для которого $y \in L_x$).

1.4. Операции над отношениями

Исходя из операций над множествами, мы можем определить ряд полезных операций над отношениями. В этом параграфе мы будем считать, что все отношения заданы на одном и том же множестве M.

Итак, возьмем два отношения A и B. Каждому из них соответствует некоторое множество пар (подмножества

$$A \subseteq M \times M \text{ и } B \subseteq M \times M).$$

Пересечением отношений $A \cap B$ мы назовем отношение, определяемое пересечением соответствующих подмножеств. Ясно, что соотношение $xA \cap By$ выполнено тогда и только тогда, когда одновременно выполнены xAy и xBy.

Пример. Пусть M — множество вещественных чисел, A — отношение «быть не меньше», B — отношение «быть не равным». Тогда, $A \cap B$ есть отношение «быть строго больше». В самом деле, xAy равносильно тому, что $x \geq y$; xBy равносильно тому, что $x \neq y$. Но эти неравенства выполняются одновременно тогда и только тогда, когда $x > y$.

Аналогично, *объединением* отношений $A \cup B$ мы назовем отношение, определяемое объединением соответствующих множеств. Соотношение $x A \cup B y$ выполнено из соотношений *xAy* или *xBy.*

Например, если *A* — отношение «больше» на множестве чисел, а *B* — отношение «равно», то $A \cup B$ — это отношение $\geq$.

Для отношений можно определить понятие включения. Мы будем писать $A \subseteq B,$ если множество пар, для которых выполнено первое отношение, содержится во множестве тех пар, для которых выполнено второе отношение. Соответственно, мы будем писать: $A \subset B,$ если множество пар *A* является подмножеством множества *B,* причем $A \neq B$.

Например, выполнено включение

$$< \subset \leqslant.$$

В самом деле, если $x < y$, то заведомо $x \leq y$. Но существуют такие пары, что $x \leq y$, но неверно соотношение $x < y$. Это будет в том случае, когда $x = y$.

Очень важно отметить такое (вполне очевидное из определения) свойство включения: если $A \subseteq B,$ то из *xAy* следует *xBy.* И обратно: если из *xAy* следует *xBy,* то $A \subseteq B.$

Отсюда видно, что для любого отношения *A*

$$\varnothing \subseteq A \subseteq U,$$

где Ø — пустое, а *U* — полное отношения.

Теперь мы введем некоторые операции, не сводящиеся непосредственно к теоретико-множественным.

Простейшая из них — переход к обратному отношению. Если *A* — отношение на множестве *M*, то *обратное отношение* A^{-1} определяется условием: $xA^{-1}y$ равносильно *yAx.*

Например, если *A* —отношение >, то A^{-1} есть отношение <. В самом деле, запись $x < y$ равносильна записи $y > x$. Еще пример: если *A* означает «быть мужем», то A^{-1} — «быть женой».

Очень важную роль играет операция, обозначаемая *AB* — *произведение* отношений. Эта операция определяется следующим образом: соотношение *xABy* равносильно тому, что существует такое $z \in M,$ для которого выполнены соотношения *xAz* и *zAy*.

Пусть *A* — отношение «быть женой», а *B* — «быть отцом». Что означает в этом случае соотношение *xABy?* По определению существует такое *z*, что «*x*—жена *z*» и «*z*— отец *y*». Иначе, «*x* есть жена отца *y*», т. е. «*x* — мать или мачеха *y*».

Пусть *A* — отношение «быть братом», а *B* — отношение «быть родителем». Тогда произведение *AB* есть отношение «быть братом одного из родителей», т. е. «быть дядей»

Возьмем теперь хорошо знакомые отношения «меньше» (обозначим его *A)* и «больше» (обозначим его *B)* на множестве целых чисел. Соотношение $xABy$ выполнено, если существует z такое, что $x < z$ и $z > y$. Ясно, что такое z существует всегда — можно взять, скажем, $z = x + y + 1$. Таким образом, $A\ B$ есть, в данном случае, полное отношение.

Далее мы убедимся, что произведение отношений обладает рядом хороших алгебраических свойств, роднящих его с обычным произведением чисел. А пока попробуйте определить, что за отношение будет AA? А что это за отношение, если A (отношение «меньше») задать на множестве всех вещественных чисел? А какие отношения выражают AB и BA, если A и B — те же отношения неравенства < и > на множестве М, состоящем только из чисел 0, 1, 2, 3, 4, 5, 6, 7, 8, 9?

Определим еще одну важную операцию, которая называется транзитивным замыканием отношения A и будет обозначаться через $\hat{A}$. Смысл этого названия будет ясен из теоремы 1.5.

Если A — некоторое отношение на множестве M, то его *транзитивное замыкание* определяется следующим образом. Соотношение $x\hat{A}y$ считается выполненным, если существует цепочка элементов из M: $z_0 = x, z_1, \ldots, z_n = y$ такая, что между соседями в этой цепочке выполнено отношение A:

$$z_0Az_1,\quad z_1Az_2,\ \ldots,\ z_{n-1}Az_n.$$

В частности, эта цепочка может состоять только из двух элементов (n=l): $z_0 = x$ и $z_1 = y$. Значит, если выполнено xAy, т. е. z_0Az_1, то выполнено и соотношение $x\hat{A}y$. Этот факт можно записать в виде соотношения

$$A \subseteq \hat{A}. \tag{1.1}$$

Если цепочка состоит из трех элементов ($n = 2$), то это значит, что xAz и zAy. Иначе говоря, $xAAy$. Если цепочка состоит из четырех элементов, то $xAAAy$. Продолжая это рассуждение, мы заключаем, что $x\hat{A}y$ в том и только в том случае, когда выполнено хотя бы одно соотношение вида $xAA \ldots Ay$. Или, сокращенно, xA^ny. Используя операцию объединения, этот факт можно записать в виде равенства

$$\hat{A} = A \cup A^2 \cup A^3 \cup \ldots \cup A^n \cup \ldots \tag{1.2}$$

Итак, мы доказали, что транзитивное замыкание отношения A есть объединение всех степеней этого отношения.

Теперь выясним, как введенные операции можно выразить с помощью операций над матрицами и графами. Поскольку матрицы, которые нам нужны, состоят только из нулей и единиц, нам будет полезно ввести специальную (так называемую *булеву)* арифметику на множестве из нуля и единицы. Эта арифметика задается следующими двумя таблицами —сложения и умножения:

$$\begin{array}{ll} 0+0=0 & 0\cdot 0=0 \\ 0+1=1 & 0\cdot 1=0 \\ 1+0=1 & 1\cdot 0=0 \\ 1+1=1\ (!) & 1\cdot 1=1 \end{array}$$

Как видим, эта арифметика отличается от обычной только тем, что сложение двух единиц дает в результате единицу. Зато, оперируя над числами 0 и 1, мы не выходим за пределы этих чисел. Легко убедиться, что в этой арифметике мы можем выполнять привычные преобразования, но не можем пользоваться вычитанием: 1 — 1 может равняться и нулю и единице.

Теперь мы можем определить операции над матрицами и графами, соответствующие операциям над отношениями.

Далее мы условимся, что нумерация на множестве *М* уже выбрана и что матрицы, соответствующие (при данной нумерации) отношениям *А* и *В,* обозначены

$$\|a_{ik}\| \text{ и } \|b_{ik}\|.$$

Очевидно, величина

$$c_{ik}=a_{ik}b_{ik} \tag{1.3}$$

равна единице в том и только в том случае, когда выполнены оба соотношения x_iAx_k и x_iBx_k, т. е. когда выполнено соотношение $x_iA\cap Bx_k$. Значит, матрица $\|c_{ik}\|$, определенная по (1.3), представляет отношение $C=A\cap B$. Этому факту можно придать несколько иное выражение. Назовем *пересечением матриц* $\|a_{ik}\|\cap\|b_{ik}\|$ матрицу $\|c_{ik}\|$, полученную почленным перемножением элементов исходных матриц (согласно (1.3)). Тогда пересечение отношений представляется пересечением матриц.

Например, пусть *А* и *В* представляются матрицами четвертого порядка *(М* содержит четыре элемента):

$$\|a_{ik}\|=\begin{Vmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{Vmatrix}, \quad \|b_{ik}\|=\begin{Vmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 \end{Vmatrix}.$$

Тогда пересечение $A\cap B$ представляется матрицей

$$\|c_{ik}\| = \begin{Vmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{Vmatrix}.$$

В терминах графов пересечение определяется так. Нарисуем множество вершин M и изобразим отношение A пунктирными стрелками, а отношение B — штриховыми стрелками. Теперь соединим простыми стрелками те и только те вершины, которые соединены обоими типами стрелок. Очевидно, что этот граф изображает пересечение отношений $A \cap B$ (рис. 1.11).

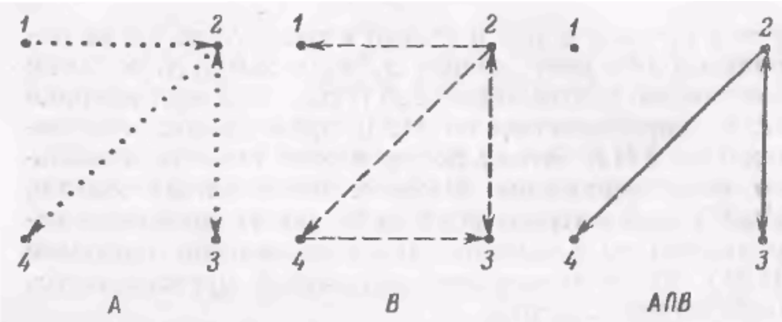


Рис. 1.11. Пересечение отношений.

Объединение отношений $A \cup B$, представляемых матрицами $\|a_{ik}\|$ и $\|b_{ik}\|$, может быть аналогично выражено с помощью операции *объединения* (сложения) *матриц.* А именно, обозначим через $\|c_{ik}\| = \|a_{ik}\| + \|b_{ik}\|$ матрицу, элементы которой определяются ус» ловием

$$c_{ik} = a_{ik} + b_{ik}. \tag{1.4}$$

В формуле (1.4) сложение понимается в смысле булевой арифметики. Посмотрев на таблицу сложения в этой арифметике, мы легко убеждаемся, что c_{ik} = 1 в том и только том случае, когда хотя бы одно из слагаемых a_{ik}, или b_{ik} равно единице. Значит, c_{ik} = 1 равносильно тому, что $x_i A \cup B x_k$.

Для предыдущего примера матриц четвертого порядка их объединение представляется матрицей

$$\|c_{ik}\| = \begin{Vmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{Vmatrix}.$$

Граф объединения строится путем проведения стрелок между всеми вершинами, которые соединены стрелкой хотя бы одного типа. Взяв графы *А* и *В* из рис. 1.11, мы получим граф объединения таким, как он изображен на рис. 1.12.

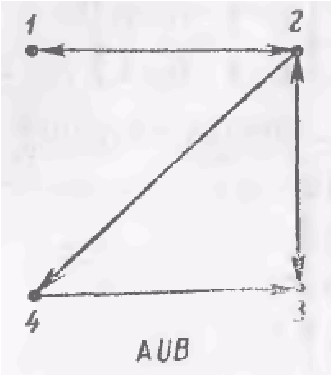


Рис. 1.12. Объединение отношений.

Произведение отношений *АВ* представляется так называемым *произведением матриц.* Эта, играющая большую роль в алгебре, операция над матрицами,определяется следующим правилом:

$$c_{ik} = a_{i1}b_{1k} + a_{i2}b_{2k} + \ldots + a_{in}b_{nk},$$

или, используя сокращенное обозначение для суммы,

$$c_{ik} = \sum_{j=1}^{n} a_{ij}b_{jk}. \qquad (1.5)$$

Число *n* обозначает здесь порядок матрицы — количество элементов множества *М.* Несмотря на простоту утверждаемой связи между произведениями отношений и матриц, проведем необходимое доказательство.

Пусть выполнено соотношение x_iABx_k. Покажем, что величина c_{ik}, вычисленная согласно (1.5), равна единице. Действительно, по определению произведения отношений существует элемент $x_j \in M$ такой, что x_iAx_j и x_jBx_k. Это значит, что $a_{ij} = b_{jk} = 1$. Значит, $a_{ij}b_{jk} = 1$. Но по правилам булевой арифметики, если одно из слагаемых равно единице, то сумма (1.5) заведомо равна единице, т. е. c_{ik}=1. Обратно, пусть c_{ik} = 1. Тогда среди слагаемых в (1.5) хотя бы одно равно единице. Пусть это будет $a_{ij}b_{jk}$. Но произведение $a_{ij}b_{jk}$ равно 1 только тогда, когда $a_{ij} = b_{jk} = 1$. А это означает, что x_iAx_j и x_jBx_k, т. е. x_iABx_k.

Итак, мы доказали, что произведению отношений соответствует произведение матриц.

Графовая интерпретация произведения такова. Пусть опять отношение *А* изображается пунктирными стрелками, а *В* — штриховыми. Соединим вершины x_i и x_k простой стрелкой, если можно пройти из x_i в x_k так: сначала из x_i , но пунктирной стрелке в некоторое x_j, а затем из x_j по штриховой стрелке в x_k (рис. 1.13).

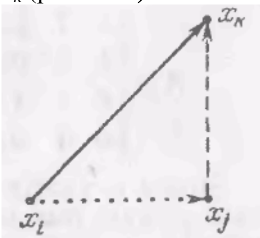


Рис. 1.13. Произведение отношений

Эти новые стрелки изображают произведение отношений *АВ*.

На рис. 1.13 видно, что способ построения графа для произведения отношений напоминает правило параллелограмма для сложения скоростей или сил. Сходство это не случайно. Пусть *М* — множество точек на плоскости, и соотношение *хАу* означает, что из точки *х* можно попасть в точку *у* за единицу времени, двигаясь со скоростью *a,* а соотношение *хВу* означает, что, двигаясь со скоростью *b,* можно за единицу времени попасть из *х* в *у.* Тогда *хАВу* означает, что, двигаясь со скоростью *a* + *b,* можно за единицу времени попасть из *х* в *у.*

Операция A^{-1} в матричной форме выражается весьма просто. Если *А* представляется матрицей $\|a_{ik}\|$, то A^{-1} изображается матрицей $\|\alpha_{ik}\|$, у которой поменялись ролями строки и столбцы: $\alpha_{ik} = a_{ki}$. Иначе говоря, матрица для A^{-1} получается из исходной матрицы симметричным отражением относительно главной диагонали. Действительно, если $a_{ik} = 1$, то x_iAx_k и $x_kA^{-1}x_i$, т. е. $\alpha_{ki} = 1$. Если же $a_{ik} = 0$, то $\alpha_{ki} = 0$.

Пример.

$$A \rightarrow \begin{Vmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{Vmatrix}, \quad A^{-1} \rightarrow \begin{Vmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{Vmatrix}.$$

Чтобы из графа, изображающего отношение *A,* получить граф, изображающий отношение A^{-1}, надо все стрелки поменять на противоположно направленные, а все петли оставить на месте.

Операция транзитивного замыкания $\hat{A}$ в матричной форме выражается через объединение степеней матрицы *A* согласно формуле (1.2). Более наглядным является переход от графа, изображающего отношение *A,* к графу, изображающему отношение $\hat{A}$. В самом деле, из определения транзитивного замыкания следует, что в новом графе стрелка соединяет вершины x_i и x_k, если в исходном графе существует путь, ведущий из x_i в x_k по направлению стрелок. На рис. 1.14 изображен граф отношения *A*.

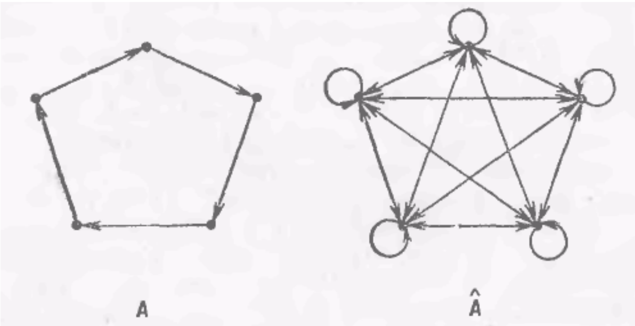


Рис. 1.14. Транзитивное замыкание отношения.

Очевидно, что из любой его вершины есть путь, ведущий в любую вершину, в том числе и в ее самое. Таким образом, отношению $\hat{A}$ соответствует в данном случае полный граф.

1.5. Алгебраические свойства операций

Поскольку операции пересечения и объединения отношений возникли из теоретико-множественных операций пересечения и объединения, то все их свойства в точности таковы, как у теоретико-множественных операций.

Рассмотрим теперь алгебраические свойства остальных операций.

У операции обращения A^{-1} есть важное свойство. Оно выражается равенством

$$(A^{-1})^{-1} = A. \tag{1.6}$$

Действительно, $x\,(A^{-1})^{-1}y$ равносильно тому, что $yA^{-1}x$.

А последнее равносильно тому, что xAy.

Операция умножения AB, в отличие от умножения обычных чисел, не перестановочна: в общем случае, $AB \neq BA$. Это можно увидеть на простом примере, когда отношения представляются следующими матрицами:

$$A \rightarrow \begin{Vmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{Vmatrix}, \quad B \rightarrow \begin{Vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{Vmatrix}.$$

В этом случае

$$AB \rightarrow \begin{Vmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{Vmatrix}, \quad BA \rightarrow \begin{Vmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{Vmatrix}.$$

Выкладки предоставляем читателю. Впрочем, они хорошо понятны в графовом представлении отношений (рис. 1.15, где отношение A изображено пунктиром, а отношение B — штрихами. При каждой вершине подразумеваются две — пунктирная и штриховая — петли.).

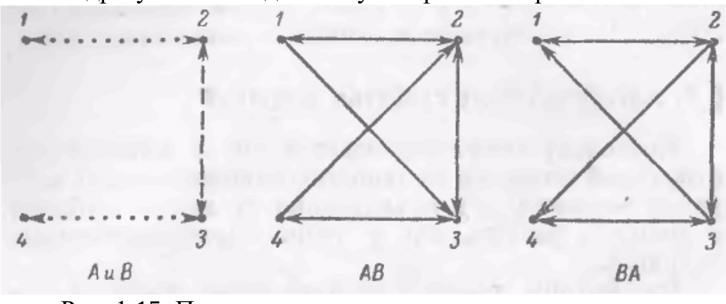


Рис. 1.15. Пример некоммутативности произведения.

В случае, когда произведение отношений не зависит от порядка: $AB = BA$, говорят, что A и B *коммутируют*.

Легко проверить, что диагональное отношение E играет роль единицы:

$$AE = EA = A \tag{1.7}$$

для любого отношения *A*.
Аналогично, для пустого отношения имеем

$$A\varnothing=\varnothing A=\varnothing. \tag{1.8}$$

В самом деле, $x\varnothing Ay$ не может выполняться ни для какой пары, так как $x\varnothing z$ никогда не выполнено. Равенство (1.8) означает, что пустое отношение Ø ведет себя относительно умножения отношений как нуль при обычном умножении чисел.
Ассоциативный (сочетательный) закон оказывается справедливым для произведения отношений:

$$(AB)C=A(BC). \tag{1.9}$$

В самом деле. Если $x(AB)Cy$, то существует такое z, что $xABz$ и zCy. Из $xABz$ вытекает существование такого w, что xAw и wBz. Из wBz и zCy следует $wBCy$. Из xAw и $wBCy$ получаем $xA(BC)y$. Аналогично, из $xA(BC)y$ легко вывести $x(AB)Cy$. Итак, (1.9) доказано.
Ассоциативный закон позволяет отказаться от расстановки скобок в произведениях и писать просто: *ABC*, *ABCD* и т. п. Вместо произведений типа *AAA, AAAA* мы будем писать степени A^3, A^4, ….
Теперь рассмотрим свойства, связывающие различные операции.
Простейшее из этих свойств — правило обращения произведения:

$$(AB)^{-1}=B^{-1}A^{-1}. \tag{1.10}$$

Действительно, $x(AB)^{-1}y$ означает, что $yABx$, т. е. существует z, для которого yAz и zBx. Но это значит, что $xB^{-1}z$ и $zA^{-1}y$, т. е. $xB^{-1}A^{-1}y$.
Другое свойство, связывающее операции обращения и произведения, состоит в следующем: если для всякого x существует такое z, что xAz, то

$$AA^{-1}\supseteq E. \tag{1.11}$$

Действительно, из xAz следует $zA^{-1}x$, т. е. $xAA^{-1}x$. Но xEy означает, что $x=y$. Значит, из xEy следует $xAA^{-1}y$.
Аналогично, если для всякого x существует такое z, что zAx, то

$$A^{-1}A\supseteq E. \tag{1.12}$$

Доказанные свойства означают, что для отношении, которые выполняются не слишком редко (каждый элемент x хоть с кем-то да находится в отношении *A),* операция обращения похожа на числовую операцию перехода от a к a^{-1}: включения (1.11) и (I.I2) близки к числовому равенству $aa^{-1}=1$, так как *E,* как уже говорилось, играет у нас роль единицы.
Следующие два свойства связывают операцию произведения с пересечением и объединением. Они похожи на распределительный (дистрибутивный) закон умножения относительно сложения. Первый из этих «распределительных законов» имеет вид

$$(A \cup B)C = (AC) \cup (BC). \qquad (1.13)$$

Доказывается он таким образом. Сначала предположим, что выполнено соотношение $x(A \cup B)Cy$. Это означает существование такого z, что выполнено по крайней мере одно из соотношений: xAz или xBz — и соотношение zCy. Тогда выполнено либо $xACy$, либо $xBCy$. Значит, выполнено соотношение $x(AC) \cup (BC)y$. Обратно, пусть выполнено $x(AC) \cup (BC)y$. Это значит, что либо $xACy$, либо $xBCy$. То есть либо существует z_1 для которого xAz_1 и z_1Cy, либо существует z_2, для которого xBz_2 и z_2Cy. Но так как $A \subseteq A \cup B$ и $B \subseteq A \cup B$, то в первом случае имеем $x(A \cup B)z_1$ и z_1Cy, т. е. $x(A \cup B)Cy$. Во втором случае: $x(A \cup B)z_2$ и z_2Cy, т. е. опять-таки $x(A \cup B)Cy$. Итак, из выполнения правой части (1.13) следует выполнение левой части и обратно. Тем самым равенство (1.13) доказано.

Второй «распределительный закон» имеет более слабую форму включения:

$$(A \cap B)C \subseteq (AC) \cap (BC). \qquad (1.14)$$

Предположим, что выполнено соотношение $x(A \cap B)Cy$. Это означает, что существует z такое, что одновременно выполнены соотношения xAz, xBz и zCy. Значит, одновременно выполнены пары соотношений: xAz и zCy; xBz и zCy. Или $xACy$ и $xBCy$, т. е. $x(AC) \cap (BC)y$, что и требовалось доказать. Но заменить в (1.14) включение равенством нельзя.

Возьмем конкретные отношения A, B, C на множестве из четырех элементов $M = \{x_1, x_2, x_3, x_4\}$, так что выполнены только следующие соотношения (рис. 1.16): x_1Ax_2, x_1Bx_3, x_2Cx_4, x_3Cx_4. Ясно, что $A \cap B = \varnothing$.

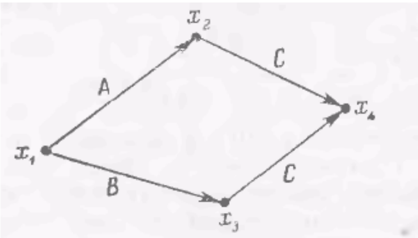


Рис 1.16.

Таким образом, согласно (1.8), $(A \cap B)C = \varnothing$. С другой стороны, x_1ACx_4 и x_1BCx_4. Следовательно, $x_1(AC) \cap (BC)x_4$, т. е. $(AC) \cap (BC) \neq \varnothing$. В этом случае имеем строгое включение

$(A \cap B)C \subset (AC) \cap (BC)$, что показывает невозможность замены в (1.14) включения на равенство.

Предоставляем читателю проверить следующие простые свойства операций:

$$(A \cup B)^{-1} = A^{-1} \cup B^{-1}, \tag{1.15}$$

$$(A \cap B)^{-1} = A^{-1} \cap B^{-1}. \tag{1.16}$$

Для операции транзитивного замыкания справедливо следующее важное свойство:

$$\text{если } A \subseteq B, \quad \text{то} \quad \hat{A} \subseteq \hat{B}. \tag{1.17}$$

Доказательство мы предоставляем читателю. Аналогично, подобная «монотонность» выполнена для других операций, а именно:

$$1) \text{ если } A \subseteq B, \text{ то } A^{-1} \subseteq B^{-1}; \tag{1.18}$$

$$2) \text{ если } A \subseteq B, \text{ то } AC \subseteq BC \text{ и } CA \subseteq CB. \tag{1.19}$$

Наконец, очевидно следующее свойство:

$$\hat{\hat{A}} = \hat{A}. \tag{1.20}$$

По-видимому, этим исчерпываются основные свойства операций, справедливые для любых отношений. В следующих главах мы изучим алгебраические свойства этих операций для некоторых специальных классов отношений.

В качестве заготовки на дальнейшее мы определим некоторые операции, выражающиеся через исходные:

1) симметризованное произведение —

$$A \circ B = AB \cup BA;$$

2) транзитивное замыкание объединения —

$$A \hat{\cup} B = \widehat{A \cup B};$$

3) транзитивное замыкание симметризованного произведения —

$$A \hat{\circ} B = \widehat{A \circ B}.$$

Из определения ясно, что эти три операции коммутативны.

Однако ассоциативный закон для симметризованного произведения уже не обязан быть справедливым в общем случае. В самом деле, пользуясь доказанным ранее распределительным законом, сосчитаем два тройных произведения:

$$(A \circ B) \circ C = (AB \cup BA)C \cup C(AB \cup BA) = \\ = ABC \cup BAC \cup CAB \cup CBA; \tag{1.21}$$

$$A \circ (B \circ C) = A(BC \cup CB) \cup (BC \cup CB)A = \\ = ABC \cup ACB \cup BCA \cup CBA. \tag{1.22}$$

Если A и C коммутируют, то

$$BAC \cup CAB = BCA \cup ACB.$$

Сравнивая это равенство с (1.21) и (1.22), получаем, что при *AC* = *C A*

$$(A \circ B) \circ C = A \circ (B \circ C).$$

В частности, ассоциативный закон верен, когда все три отношения коммутируют. Тогда

$$(A \circ B) \circ C = A \circ (B \circ C) = ABC.$$

Для читателя будет полезным упражнением фактически построить пример таких трех отношений, для которых ассоциативный закон не имеет места.

1.6. Свойства отношений

Здесь мы укажем некоторые доплнительные сведения о важных свойствах отношений, которые позволят нам в дальнейшем выделить существенные классы отношений.

Определение 1.1. Отношение *А* называется *рефлексивным,* если $E \subseteq A.$ Иначе говоря, рефлексивное отношение всегда выполнено между объектом и им самим: *хАх.*

Содержательные примеры рефлексивных отношений: «быть похожим на», «иметь общий признак с» (если каждый объект имеет хоть один признак), «быть не старше». С другой стороны, отношения типа «быть братом», «быть старше» заведомо не рефлексивны.

Рефлексивные отношения всегда представляются матрицей, у которой на главной диагонали стоят единицы. В графе, изображающем рефлексивное отношение, каждая вершина имеет петлю. Именно поэтому, имея дело с заведомо рефлексивным отношением, мы не будем эти петли изображать на чертеже.

Определение 1.2. Отношение *А* называется *антирефлексивным,* если из *хАу* следует *х≠ у,* т. е., в алгебраической записи, $A \cap E = \varnothing.$ Иначе говоря, $A \subseteq \neq,$ т. е. отношение *А* может выполняться лишь для несовпадающих объектов.

Отношения, приведенные выше в качестве примеров нерефлексивных отношений, являются антирефлексивными. Отношение «быть эталоном для», вообще говоря, не будет ни рефлексивным, ни антирефлексивным.

Матрица, представляющая антирефлексивное отношение, имеет на главном диагонали нули, а в соответствующем графе петли непременно отсутствуют.

Определение 1.3. Отношение A называется *симметричным,* если $A \subseteq A^{-1}$. Иначе говоря, если выполнено соотношение xAy, то выполнено и соотношение yAx.

Содержательными примерами таких отношении служат «быть похожим на», «быть одинаковым с», «быть родственником».

В матрице, представляющей симметричное отношение, элементы, симметрично расположенные относительно главной диагонали, равны между собой:

$$a_{ik} = a_{ki}.$$

В соответствующем графе вместе с каждой стрелкой, идущей из вершины x_i в вершину x_k, существует и противоположно направленная стрелка. Поэтому в таком графе можно вообще не обозначать стрелки, а рисовать только петли и отрезки, соединяющие разные вершины. Иначе говоря, симметричное отношение естественно изображается неориентированным графом. Так мы и будем впредь изображать графы заведомо симметричных отношений.

Теорема 1.1. *Отношение A тогда и только тогда симметрично, когда*

$$A = A^{-1}$$

Доказательство. По определению $A \subseteq A^{-1}$, но, в силу (1.18), имеем

$$A^{-1} \subseteq (A^{-1})^{-1}.$$

Отсюда, согласно (1.20), получается

$$A^{-1} \subseteq A.$$

Сравнивая это включение с исходным, приходим к выводу, что $A = A^{-1}$. Обратное утверждение очевидно.

Определение 1.4. Отношение A называется *асимметричным,* если $A \cap A^{-1} = \varnothing$. Это означает, что из двух соотношений xAy и yAx по меньшей мере одно не выполнено. Для матричных элементов это приводит к равенству:

$$a_{ik}a_{ki} = 0. \tag{1.23}$$

В соответствующем графе не может быть стрелок, соединяющих две вершины в противоположном направлении, т. е. направление стрелки всегда существенно.

Теорема 1.2. *Если отношение A асимметрично, то оно антирефлексивно.*

Доказательство. Предположим, что для какого-то x выполнено xAx. Тогда верно было бы и $xA^{-1}x$, т. е. $xA \cap A^{-1}x$. Но тогда отношение $A \cap A^{-1}$ не было бы пустым.

Этот факт можно было бы вывести и из уравнений для матричных элементов: подставляя в (1.23) $i = k$, получаем $a^2_{kk} = 0$, т. е. $a_{kk} = 0$.

Из теоремы 1.2 вытекает, что граф асимметричного отношения не может иметь петель.

Определение 1.5. Отношение *А* называется *антисимметричным,* если $A \cap A^{-1} \subseteq E.$ Это означает, что оба соотношения *хАу* и *уАх* выполняются одновременно только тогда, когда $x = y$.

Для матричных элементов это приводит к утверждению:

$$a_{ik}a_{ki} = 0, \quad \text{если} \quad i \neq k.$$

Определение 1.6. Отношение *А* называется *транзитивным,* если $A^2 \subseteq A.$

Раскрывая алгебраическое условие, приходим к следующему: если *xAz* и *zAy,* то выполнено и *xAy.* По индукции отсюда следует такое свойство: если $xAz_1,\ z_1Az_2,\ \ldots,\ z_{n-1}Ay,$ то *xAy.*

Это свойство хорошо интерпретируется на графе, изображающем отношение *А*. Именно, если точки *х* и *у* соединены путем, проходимым по направлению стрелок, то существует стрелка, непосредственно идущая из вершины *х* в вершину *у.*

Замечание. Нетрудно показать, что для рефлексивного отношения *А* транзитивность эквивалентна равенству $A^2 = A$.

Теорема 1.3. *Если А транзитивно, то* $A = \hat{A}.$ Иными словами, *транзитивное замыкание транзитивного отношения совпадает с ним самим.*

Доказательство. Сначала докажем, что для транзитивного отношения *А* верно включение:

$$A^n \subseteq A. \tag{1.24}$$

В самом деле, при *n*=2 это есть определение транзитивности отношения. Предположим, что (1.24**)** уже доказано для некоторого *n.* Тогда, по ассоциативному закону $A^{n+1} = A^nA;$ учитывая предположение индукции (1.24) и (1.19), имеем

$$A^{n+1} = A^nA \subseteq AA \subseteq A.$$

Итак, индукция у нас прошла благополучно. Теперь обратимся к формуле (1.2). определяющей транзитивное замыкание $\hat{A},$ и заменим каждый член объединения на больший согласно (1.24). Получаем

$$\hat{A} = A \cup A^2 \cup A^3 \cup \ldots \cup A^n \cup \ldots$$
$$\ldots \subseteq A \cup A \cup \ldots \cup A \cup \ldots = A.$$

Итак, $\hat{A} \subseteq A.$ Но, с другой стороны, согласно (1.1), всегда $A \subseteq \hat{A}.$ Значит, $A = \hat{A}.$ Теорема доказана.

Легко видеть, что имеет место и обратная

Теорема 1.4. *Если* $A = \hat{A},$ *то А транзитивно.*

Доказательство. Из формулы (1.2) следует, что $A^2 \subseteq \hat{A}$. Так как $\hat{A} = A$, то и $A^2 \subseteq A$.

Теорема 1.5. *Для любого отношения А транзитивное замыкание* $\hat{A}$ *равно пересечению* $\bigcap B$ *всех транзитивных отношений В, содержащих А.*

Доказательство. Поскольку $\hat{\hat{A}} = \hat{A}$, из теоремы 1.4 вытекает, что отношение $\hat{A}$ всегда транзитивно. Кроме того, $A \subseteq \hat{A}$. Значит, $\hat{A}$ — одно из *В,* о которых говорилось в условии. Следовательно, $\hat{A} \supseteq \bigcap B$. Для доказательства обратного включения предположим, что *В* — произвольное транзитивное отношение, содержащее *А*. Итак, $A \subseteq B$. По **(**1.17**)** $\hat{A} \subseteq \hat{B}$. Но по теореме 1.3 $\hat{B} = B$. Следовательно, $\hat{A} \subseteq B$. Теорема доказана.

Если отношение $\langle A, M \rangle$ есть сужение отношения $\langle A_1, M_1 \rangle$, на него автоматически переносятся все введенные выше свойства последнего отношения. Так, рефлексивность отношения $\langle A_1, M_1 \rangle$ влечет рефлексивность сужения $\langle A, M \rangle$. Действительно, если для любого $x \in M_1$ верно xA_1x, то при $x \in M$ будет также выполнено *хАх.* Симметричность отношения $\langle A_1, M_1 \rangle$ влечет симметричность сужения, поскольку при $x \in M$ и $y \in M$ из *хАу* следует *уАх.* Проверку для остальных свойств предоставляем читателю.

1.7. Инвариантность свойств отношений

В этом параграфе мы изучим случаи, когда те или иные свойства результата операции над отношениями определяются аналогичными свойствами операндов

Лемма 1.1. *Если отношения А и В рефлексивны, то рефлексивны и следующие отношения:*

$$A \bigcup B, \quad A \bigcap B, \quad A^{-1}, \quad AB, \quad \hat{A}.$$

Доказательство непосредственно вытекает из соответствующих определений. Например, из того, что для всякого *х* выполнены соотношения *хАх* и *хВх,* следует, что выполнено $x A \bigcap B x$ и $x A \bigcup B x$. Несколько сложнее обстоит дело со свойством антирефлексивности. В этом случае справедлива

Лемма 1.2. *Если отношения А и В антирефлексивны, то антирефлексивны и следующие отношении:*

$$A \bigcup B, \ A \bigcap B, \ A^{-1}.$$

Доказательство этих утверждений проводится также, как и в предыдущем случае.

Что касается произведения *AB* и транзитивного замыкания $\hat{A}$ антирефлексивных отношений, то они уже могут не не быть антирефлексивными (легко видеть, что произведение *AB* отношений *A* и *B* тогда и только тогда антирефлексивно, когда $A \cap B^{-1} = \varnothing$). Примером может послужить отношение *A,* задаваемое на множестве *M* из двух элементов матрицей

$$A \to \begin{Vmatrix} 0 & 1 \\ 1 & 0 \end{Vmatrix}.$$

Легко видеть, что квадрат этой матрицы

$$A^2 \to \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}$$

задает уже рефлексивное отношение, а транзитивное замыкание

$$\hat{A} \to \begin{Vmatrix} 1 & 1 \\ 1 & 1 \end{Vmatrix}$$

как полное отношение, тоже рефлексивно. Советуем читателю нарисовать соответствующие графы.

Теперь рассмотрим, как ведет себя при различных операциях свойство симметричности отношений.

Лемма 1.3. *Если отношения A и B симметричны, то симметричны и следующие отношения:* $A \cup B$, $A \cap B$, A^{-1}.

Доказательство. В силу (1.6) и теоремы 1.1 имеем $(A^{-1})^{-1} = A = A^{-1}$, т. е. отношение A^{-1} также симметрично. Из равенства (1.15) получаем

$$(A \cup B)^{-1} = A^{-1} \cup B^{-1} = A \cup B,$$

т. е. объединение $A \cup B$ симметрично. Из равенства (1.16) имеем

$$(A \cap B)^{-1} = A^{-1} \cap B^{-1} = A \cap B,$$

следовательно, симметричность пересечения доказана.

Что касается симметричности произведения, то полный ответ дает здесь

Лемма 1.4. *Чтобы произведение AB симметричных отношений A и B было симметрично, необходимо и достаточно, чтобы отношения A и B коммутировали.*

Доказательство. Пусть *AB = B A.* Тогда, согласно (1.10), имеем

$$(AB)^{-1} = (BA)^{-1} = A^{-1}B^{-1} = AB,$$

т. е. произведение AB симметрично. Обратно, если AB симметрично, то по теореме 1.1 $AB = (AB)^{-1}$. Но тогда по (1.10) получаем $AB = (AB)^{-1} = B^{-1}A^{-1} = BA$, т. е. $AB = BA$. Лемма доказана.

Знакомые с линейной алгеброй читатели наверняка уже догадались, что эта теорема является просто вариантом известной теоремы о том, что произведение симметричных матриц симметрично в том и только том случае, когда эти матрицы перестановочны (коммутируют).

Следствие. *Транзитивное замыкание* $\hat{A}$ *симметричного отношения A есть симметричное отношение.* Действительно, из леммы 1.4 и (1.9) легко вывести что отношения $A^2, A^3, \ldots, A^n, \ldots$ симметричны. Но тогда, согласно (1.2) и лемме 1.3, транзитивное замыкание

$$\hat{A} = A \cup A^2 \cup A^3 \cup \ldots$$

также симметрично.

Советуем читателю доказать это утверждение без леммы 1.4, исходя непосредственно из определения транзитивного замыкания.

Для свойства асимметричности справедлива

Лемма 1.5. 1) *Если отношение A асимметрично, то пересечение* $A \cap B$ *асимметрично при любом B.*

2) *Если отношение A асимметрично, то асимметрично также и отношение* A^{-1}.

Доказательство. 1) По определению 1.4 $A \cap A^{-1} = \varnothing$. Тогда, согласно (1.16), имеем

$$(A \cap B) \cap (A \cap B)^{-1} = A \cap B \cap A^{-1} \cap B^{-1} = \\ = A \cap A^{-1} \cap B \cap B^{-1} = \varnothing \cap B \cap B^{-1} = \varnothing,$$

т. е. $A \cap B$ асимметрично.

2) Аналогично, учитывая (1.6), имеем также

$$A^{-1} \cap (A^{-1})^{-1} = A^{-1} \cap A = A \cap A^{-1} = \varnothing,$$

что означает сохранение асимметричности у обратного отношения.

Объединение асимметричных отношений может уже не быть асимметричным (объединение $A \cup B$ асимметричных отношений A и B тогда и только тогда асимметрично, когда $A \cap B^{-1} = \varnothing$.). Также не обязаны быть асимметричными произведение и транзитивное замыкание асимметричных отношений.

Лемма 1.6. *Если отношения A и B антисимметричны, то антисимметричны также и следующие отношения:* $A \cap B$, A^{-1}.

Доказательство. Для обратного отношения фактически воспроизводится предыдущее рассуждение. Для пересечения доказательство почти совпадает с доказательством леммы 1.5:

$$(A \cap B) \cap (A \cap B)^{-1} = (A \cap A^{-1}) \cap (B \cap B^{-1}) \subseteq E \cap E = E.$$

Антисимметричность может не сохраняться при объединении (объединение $A \cup B$ антисимметричных отношений A и B тогда и только тогда антисимметрично, когда $A \cap B^{-1} \subseteq E.$), произведении и транзитивном замыкании отношений.

Про свойство транзитивности можно утверждать следующее:

Лемма 1.7. *Если отношения А и В транзитивны, то транзитивны также следующие отношения:*

$$A \cap B, \quad A^{-1}, \quad \hat{A}.$$

Доказательство. Пусть справедливы соотношения $x A \cap B y$ и $y A \cap B z.$ Тогда справедливы и соотношения xAy, yAz, xBy, yBz. Но в силу транзитивности отношений A и B имеем отсюда $x A \cap B z,$ т. е. $A \cap B$ траизитивио. Если верно $xA^{-1}y$ и $yA^{-1}z$, то по определению обратного отношения имеем zAy и yAx, т. е. zAx и $xA^{-1}z$. Это и означает, что A^{-1} транзнтивно. Наконец, транзитивность отношения $\hat{A}$ вытекает из (1.20) и теоремы 1.4.

Примеры решения типовых задач

Основные положения

Наиболее часто встречающимися и хорошо изученными являются двухместные, или *бинарные,* отношения. Если a, b находятся в отношении R, это часто, как вы знаете, записывается как aRb. Рассмотрим ряд примеров.

Пример 1.

1) Отношение на N:

а) отношение $\leq$ выполняется для пар (7, 9) и (7, 7), но не выполняется для пары (9, 7);

б) отношение «иметь общий делитель, отличный от единицы» выполняется для пар (6, 9), (4, 2), (2, 4), (4, 4), но не выполняется для пар (7, 9) и (9, 7);

в) отношение «быть делителем» (т.е. aRb означает «a - делитель b») выполняется для пар (2, 4) и (4,4), но не выполняется для пар (4, 2) и (7, 9).

2) Отношение на множестве точек действительной плоскости:

а) отношение «находиться на одинаковом расстоянии от начала координат» выполняется для пары точек (3, 4) и ($-2, \sqrt{21}$), но не выполняется для пары точек (3, 4) и (1, 6); это отношение совпадает с отношением «находиться на одной и той же окружности с центром в начале координат»;

б) отношение «находиться на разном расстоянии от начала координат» выполняется для тех и только тех пар точек, для которых не выполняется предыдущее отношение;

в) отношение «быть симметричным относительно оси *x»* выполняется для всех пар точек (x_1, y_1) и (x_2, y_2), которые удовлетворяют условию $x_1 = x_2, y_1 = y_2$.

3) Отношение на множестве людей:

а) «жить в одном городе»;

б) «быть моложе»;

в) «быть сыном»;

г) «быть знакомым».

Пример 2.

Для задання бинарных отношений можно использовать любые способы задання множеств (например, список пар, для которых данное отношение выполняется). Для отношений на конечных множествах обычно используется матричный способ задачи. Матрица бинарного отношения на множестве $M = \{a_1 ..., a_m\}$ — это квадратная матрица C порядка m, в которой элемент c_{ij}, стоящий на пересечении i-й строки и
j-го столбца, определяется следующим образом:

$$c_{ij} = \begin{cases} 1, & если \quad a_i R a_j; \\ 0 & в\ противном\ случае. \end{cases}$$

Например, для конечного множества {1, 2, 3, 4, 5, 6} матрицы отношений «а», «б», «в» из примера 1, п. 1 приведены в таблице матриц соответственно:

	1	2	3	4	5	6
1	1	1	1	1	1	1
2	0	1	1	1	1	1
3	0	0	1	1	1	1
4	0	0	0	1	1	1
5	0	0	0	0	1	1
6	0	0	0	0	0	1

а)

	1	2	3	4	5	6
1	0	0	0	0	0	0
2	0	1	0	1	0	1
3	0	0	1	0	0	1
4	0	1	0	1	0	1
5	0	0	0	0	1	0
6	0	1	1	1	0	1

б)

	1	2	3	4	5	6
1	1	1	1	1	1	1
2	0	1	0	1	0	1
3	0	0	1	0	0	1
4	0	0	0	1	0	0
5	0	0	0	0	1	0
6	0	0	0	0	0	1

в)

Для любого множества M отношение E, заданное матрицей, в которой по главной диагонали стоят единицы, а в других местах — нули, называется *отношением равенства* на M. Поскольку отношение на M задаются подмножествами M^2, для них можно определить те же операции, которые и над множествами. Например, отношение «б» из примера 1, п. 2 является дополнением отношения «а», отношение $\leq$ является объединением отношений $<$ и равенства.

Пример 3. Пусть

$$A = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}.$$

Тогда $R = \{(x, y)\colon x, y \in A, x$ — делитель y и $x \leq 5\}$ может быть записано в явном виде:

$$R = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (1,7)\backslash\ (1,8), (1,9), (1,10), (2,2), (2,4), (2,6), (2,8), (2,10), (3,3), (3,6), (3,9), (4,4), (4,8), (5,5), (5,10)\}.$$

Пример 4 (шахматы). Как и выше, пусть

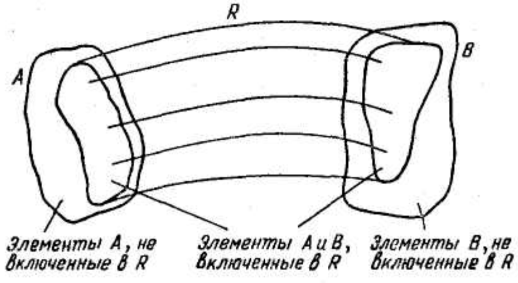
$$F = \{a, b, c, d, e, f, g, h\}, R = \{1, 2, 3, 4, 5, 6, 7, 8\}$$

и пусть $S = F\times R$.

Таким образом, S — множество всех клеток, которые обозначаются парами (x, y), где $x\in F$, $y\in = R$. Определим бинарное отношение C (для ладьи!) на S так, что $(s, t)\in C$ тогда и только тогда, когда s и t — элементы S и ладья может пройти от s к t одним ходом на пустой доске. (Напомним, что ладья может изменять либо горизонтальную координату, либо вертикальную, но не обе координаты одновременно). Поэтому $C= S\times S$ и

$$C=\{((f_s,r_s), (f_i,r_i))\colon (f_s= f_i \text{ и } r_s\neq r_i \text{ или } f_s\neq f_i \text{ и } r_s= r_i)\}.$$

Пример 5. Пусть отношение R определено в соответствии с изображением на рисунке.



Необходимо сконцентрировать наше внимание на том, что происходит на концах R, т.е. рассмотреть элементы A и B, которые принадлежат R. Они являются элементами подмножеств A и B соответственно и имеют специальные названия.

Определение. Свяжем с каждым бинарным отношением R между A и B два множеств — *область определения* $D(R)$ и *область значений* $\mathcal{R}(R)$. Они определяются следующим образом:

$$D(R)=\{x: (x, y\in R\},\ \mathcal{R}(R)=\{y: (x,y)\in R\}.$$

Пусть отношение R такое же, как и в примере 3. Тогда

$$D(R)=\{1, 2, 3, 4, 5\},\ \mathcal{R}(R)=A.$$

Пример 6. Предположим, что мы имеем некоторую программу. Она читает два числа из множества $A=\{1, 2, 3, 4, 5\}$, обозначаемых x и y, и, если $x<y$, печатает число z (также из A) такое, что $x \le z< y$. В любом случае программа останавливается после считывания всех чисел из A.

Задача определяет отношение P, $P\subseteq A\times A$ такое, что

$$P=\{((x, y), z): x< y, x\le z< y\}.$$

Не все входные данные приводят к выдаче результата. Поэтому область определения P не совпадает с A^2. Ясно, что

$$\begin{aligned}P=\{&((1,2), 1), ((1,3), 1), ((1,3), 2),\\ &((1,4), 1), ((1,4), 2), ((1,4), 3),\\ &((1,5), 1), ((1,5), 2), ((1, 5), 3), ((1,5), 4),\\ &((2,3), 2), ((2,4), 2), ((2,4), 3),\\ &((2,5), 2), ((2,5), 3), ((2,5), 4),\\ &((3,4), 3), ((3,5), 3), ((3,5), 4),\\ &((4,5), 4)\};\end{aligned}$$

$$\begin{aligned}\mathcal{R}(P)=\{&(1,2), (1,3), (1,4), (1,5),\\ &(2,3), (2,4), (2,5),\\ &(3,4), (3,5),\\ &(4,5));\end{aligned}$$

$$\mathcal{R}(P)=\{1,2,3,4\}.$$

Свойства отношений

Пример 7. Пусть

$\rho=\{(x, y): x,y\in N$ и x- делитель $y\}$,

$\sigma=\{(x, y): x,y\in N$ и $x\le y\}$,

$\tau=\{(x, y): x,y\in N\setminus(1)$ и x и y имеют общий делитель$\}$.

Тогда ρ:

а) рефлексивно, так как $x/x=1$ для всех $x\in N$;

б) несимметрично, поскольку 2 — делитель 4, но 4 не является делителем 2;

в) транзитивно, так как если $y/x\in N$ и $z/y\in N$, то

$$z/x=(y/x)*(z/y)\in N;$$

г) антисимметрично, так как если $x \le y \in N$ и $y \le x \in N$, то $x = y$.

Аналогично σ:

а) рефлексивно, так как $x \le x$ для всех $x \in N$;

б) несиимметрично, так как $2 \le 3$, но $3 \nleq 2$;

в) траизитивно;

г) антисиетрично, так как если $x \le y$ и $y \le x$, то $x=y$.

Наконец, τ рефлексивно и симметрично, но не транзитивно или антисимметрично.

Пример 8. Пусть P — множество всех людей, а A и S определяются следующим образом:

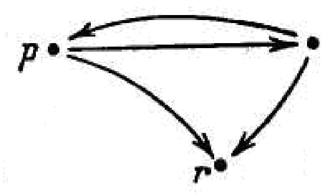
$A = \{(x, y): x, y \in P$ и x — предок $y\}$,

$S = \{(x, y) \in P$ и x и y имеют одних и тех же родителей$\}$.

Очевидно, что A транзитивно, а S рефлексивно, симметрично и транзитивно.

Заметим, что свойства симметричности и антисимметричности не являются взаимоисключающими. Для любого множества X отношение E_x есть симметричным и антисиметричним. (Проверьте!) Мы можем также иметь отношения, которые не являются ни симметричными, ни антисимметричными.

П р и м е р 9. Пусть опять P є множество всех людей. Определим отношение B такое, что xBy тогда и только тогда, когда x является братом y. В семье, которая состоит из двух братьев p и q и сестры r, мы имеем ситуацию, которая изображена на рисунке.



Отношение B не симметрично, так как pBr, но не rBp. Это отношение также не является антисимметричным, так как pBq и qBp, хотя p и q различны.

В более общей ситуации мы можем интерпретировать рассмотренные выше характеристики отношения путем построения диаграммы:

а) отношение рефлексивно тогда и только тогда, когда для каждого узла (точки) на диаграмме существует стрелка, которая начинается и заканчивается на этом узле;

б) отношение симметрично тогда и только тогда, когда для каждой стрелки, которая соединяет два узла, существует также стрелка, которая соединяет эти узлы в обратном направлении;

в) отношение транзитивно тогда и только тогда, когда для каждой пары узлов x и y, связанных последовательностью стрелок от x к a_1, от a_1 к a_2, ..., от a_{n-1} к a_n, от a_n к y, существует также стрелка от x к y;

г) отношение антисимметрично тогда и только тогда, когда не существует двух различных узлов, связанных парой стрелок.

Индивидуальные тестовые задачи

1. Даны два множества

$$X=\{x_1, x_2, x_3, x_4, x_5, x_6\} \text{ и } Y=\{y_1, y_2, y_3, y_4\}$$

и определено бинарное отношение

$$A=\{(x_1, y_2),(x_2, y_1),(x_2, y_2),(x_4, y_2),(x_4, y_3),(x_5, y_1),(x_5, y_3)\}$$

Для данного отношения A:

а) записать область определения и область значений;

б) определить сечение по каждому элементу из X;

в) определить сечение по подмножествам $X' =\{x_1, x_4\}$ и $X''=\{x_2, x_3, x_5\}$ множества X;

г) записать матрицу;

д) определить симметричное отношение A^{-1}.

2. Пусть X — множество студентов; Y — множество дисциплин и соотношение xAy, где $x\in X$ и $y\in Y$, означает «студент x изучает дисциплину y». Дайте словесное описание областей определения и значений, сечений и обратного отношения, полученных в задаче 1.

3. По результатам задачи 1 определите множества

$$A(x_2) \square A(x_4),\ A(x_2)\backslash A(x_4) \text{ и } A(x_2)+ A(x_4).$$

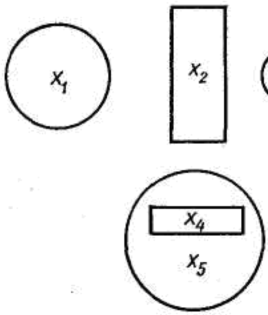
Дайте им словесное описание согласно условию задачи 2.

4. Дайте описание геометрического образа (см. рисунок), представляющего собой множество $X=\{x_1, x_2, x_3, x_4, x_5\}$ фигур на плоскости посредством отношений:

а) унарных: «x_i — окружность», «x_j — прямоугольник»;

б) бинарных: «x_i больше x_j», «x_i правее x_j», «x_i над x_j», «x_i внутри x_j», « x_i шире x_j »;

в) тернарных: «x_i между x_j и x_k», «x_i ближе к x_j , чем к x_k».



Геометрический образ к задаче 4

5. Приведите примеры рефлексивных, антирефлексивных, симметричных, асимметричных, антисимметричных и транзитивных отношений.

6. Какими общими свойствами обладают бинарные отношения, заданные в некотором множестве людей X и выражающиеся соотношениями (x_i , x_j $\in X$):

а) «x_i похожий на x_j »;

б) «x_i знакомый с x_j »;

в) «x_i родственник x_j»;

г) « x_i сосед x_j »;

д) «x_i живет в одном доме с x_j »;

е) «x_i весит больше, чем x_j»;

ж) «x_i подчиненный $x_{j;}$».

7. Записать композицию $C = BA$ отношений A ={(1, 2), (1, 3), (2, 1), (2,4), (3,3)} и B = {(1, 1),(1,3),(2, 2), (2, 3), (3, 1), (4, 2), (4,3)}. Проверить результат с помощью операций над матрицами заданных отношений.

8. Пусть отношение A задано в множестве действительных чисел R. Тогда на плоскости каждой упорядоченной паре будет соответствовать точка с координатами x и y, если $(x, y)\in A$. Отношение A изобразится графиком, представляющим собой подмножество точек плоскости (области, линии или отдельные точки). При этом отношение записывается как

$$A=\{(x, y) \in R\times R|\ P=\{(x, y)\},$$

где $P(x, y)$ - определяющее свойство отношения A, выражаемое обычно алгебраическими уравнениями и неравенствами.

Постройте графики для следующих отношений (в тех случаях, когда график является частью плоскости, эта часть штрихуется):

а) $\{(x, y) \in R\times R|\ x= y\}$;

б) $\{(x, y) \in R\times R|\ y\geq x\}$;

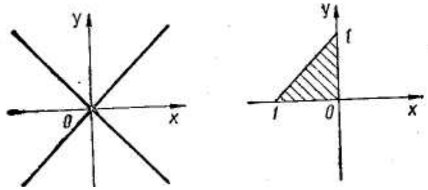
в) $\{(x, y) \in R\times R| \ y\geq x$ и $x\geq 0\}$;

г) $\{(x, y) \in R\times R| \ x^2+ y^{2+}1\}$;

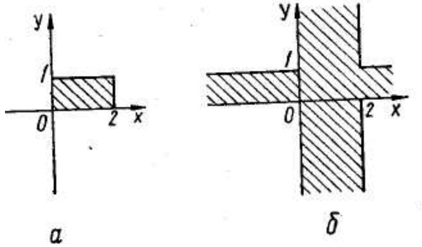
д) $\{(x, y) \in R\times R| \ |x|+2|y|=1\}$;

е) $\{(x, y) \in R\times R| \ y\geq 0$ и $y\leq x$ и $x+y\leq 1\}$;

9. Запишите отношение, графики которых показаны на рисунке.



10. Покажите, что график, приведенный на рисунке *а* соответствует пересечению, а на рисунке *б* — объединению бинарных отношений $\{(x, y) \in R\times R| \ 0\leq x\leq 2\}$ и $\{(x, y) \in R\times R| \ 0\leq y\leq 1\}$.



а) график пересечения, б) график объединения бинарних. отношений.

Запишите отношение для каждого из этих графиков одним соотношением.

11. Докажите следующие свойства симметризации и композиции отношений:

а) $(AB)^{-1} = B^{-1}A^{-1}$;

б) $(A \square B)^{-1} = A^{-1} \cup B^{-1}$;

в) $(A \mathrm{I} \ B)^{-1} = A^{-1} \mathrm{I} \ B^{-1}$;

г) $(A^{-1})^{-1} = A$.

12. Докажите, что

$$(X\times Y)^{-1} = Y\times X.$$

13. Покажите, что при симметризации отношения все его свойства сохраняются (маются на внимании свойства, перечисленные в задаче 5).

14. Докажите следующие утверждения:

а) если отношения *A* и *B* рефлексивны, то рефлексивны и отношения $A \cup B$, $A\cap B$, AB;

б) если отношения A и B антирефлексивны, то антирефлексивны и отношения $A \cup B$, $A \cap B$;

в) если отношения A и B симметричны, то симметричны и отношения $A \cup B$, $A \cap B$,

г) если отношение A асимметрично, то пересечение $A \cap B$ асимметрично при любом B;

д) если отношения A и B антисимметричны, то антисимметричны и отношения $A \cap B$;

е) если отношения A и B транзитивны, то транзитивны и отношения $A \cap B$.

15. Покажите, что:

а) композиция BA антирефлексивных отношений A и B тогда и только тогда антирефлексивна, когда $A \cap B = \varnothing$;

б) объединение $A \cup B$ асимметричных отношений A и B тогда и только тогда асимметричесно, когда $A \cap B^{-1} = \varnothing$;

в) объединение $A \cup B$ антисимметричных отношений A и B тогда и только тогда антисимметрично, когда $A \cap B^{-1} \subset E$ (E — тождественное отношение).

2. Виды и типы отношений

2.1. Одинаковость и эквивалентность

2.1.1. От одинаковости к эквивалентности

В обыденной речи мы часто говорим об одинаковости (о равенстве) каких-то объектов (предметов, множеств, абстрактных категорий), не заботясь о надлежащем уточнении смысла, который мы вкладываем в слово «одинаковый». Попробуем выявить этот смысл, проанализировав различные ситуации, когда мы уверенно считаем некоторые объекты одинаковыми.

Возьмем стандартный комплект шахматных фигур. С точки зрения шахматного игрока все белые пешки в нем одинаковы. Расставляя их на шахматной доске, шахматист будет выбирать их из коробки в произвольном порядке. В начальной позиции все они будут поставлены на вторую горизонталь и шахматист не будет размышлять над вопросом, куда ему лучше поставить выбранную наугад пешку. Точно так же любая из черных ладей при расстановке фигур перед игрой может с равным успехом попасть на королевский или ферзевый фланг. Эти ладьи одинаковы.

Но представим себе другую ситуацию: этот же комплект шахмат отдан ребенку, который играет в солдатики. Для него отдельные пешки могут приобрести индивидуальность, получить имена и метки. Однако в тот момент, когда этот же мальчик начнет использовать шахматы по прямому назначению, пешки одного цвета опять станут одинаковыми.
Возьмем еще одну ситуацию: шахматные фигуры в процессе игры. Предположим, что шахматист стоит перед выбором: отдать ли противнику пешку, проникшую уже на седьмую горизонталь и грозящую вот-вот превратиться в ферзя, или пешку, мирно стоящую в начальной позиции. Ясно, что (при прочих равных условиях) первая пешка гораздо ценней и шахматист уже не считает обе свои пешки одинаковыми. Правда, в этой ситуации объектами являются не сами по себе деревянные фигурки, а «пешки в данной позиции». В позиции этюдного характера каждая пешка играет свою индивидуальную роль, и они, разумеется, не одинаковы для хорошего шахматиста.
Разница здесь того же характера, как между словом русского языка и словом в данном контексте. Например, слова «пешка» и *«пешка»,* хотя и напечатаны разным шрифтом, одинаковы, как слова русского языка. Но в контекстах «Гроссмейстер эффектно пожертвовал пешку» и «Он был только пешкой в чужих руках» это слово имеет разные значения. Иначе говоря, слова одинаковы, а значения различаются.
Аналогично, об одинаковости людей мы можем говорить в различном смысле. С профессиональной точки зрения продавца готового платья люди, имеющие один и тот же пол, рост и размер, неразличимы. Они одинаковы в том смысле, что им нужно демонстрировать одни и те же вещи. Впрочем, хороший продавец различает покупателей по их вкусам, а хороший портной понимает, что кроме роста и размера есть индивидуальные особенности фигуры. Но для работника склада, который выдает форму (скажем, штормовые костюмы для альпинистов), существен только размер. Для профессора анатомии малосущественно, на чьем трупе он будет демонстрировать студентам устройство человеческих органов. Но уже для профессора психиатрии нет одинаковых больных.
С точки зрения инспектора по кадрам люди с тождественными анкетными данными одинаковы. Но для научного руководителя лаборатории нет одинаковых и взаимозаменимых сотрудников.
Когда мы приглашаем к себе гостей, то нам совершенно не все равно, кто придет и кого приведет с собой. С точки зрения индивидуальных человеческих взаимоотношений ни один человек не равен другому. Когда мы говорим о всеобщем равенстве людей, то понимаем под этим

в действительности равенство прав перед законом, равноценность личностей, но не равенство индивидуальностей.

Рассмотрим множество животных, изображенных на рис. 2.1. Мы разбили их на следующие шесть групп: (1) сухопутные млекопитающие, (2) обитающие в воде, (3) насекомые, (4) птицы, (5) мифические существа и (6) пресмыкающиеся. Будем считать по определению животных, входящих в одну группу, одинаковыми. Можно вообразить ситуацию, когда одинаковые в этом смысле животные взаимозаменимы. Например, когда учителю биологии надо показать ученикам представителей разных типов.

Если мы внимательно проанализируем, что общего в употреблении слова «одинаковость» во всех приведенных примерах (а также в примерах, которые читатель сумеет теперь составить сам), то мы увидим следующее. Во-первых, *одинаковость* всегда понимается как бинарное отношение на некотором множестве объектов. Во-вторых, содержание этого отношения зависит от ситуации, в которой мы рассматриваем эти объекты, или от наблюдателя, который с выбранной им точки зрения судит об одинаковости объектов. В-третьих, слово «одинаковость» попадает в один синонимический ряд со словом «взаимозаменимость» (объектов в данной ситуации).

Действительно, одинаковость белых пешек или других одноименных и одноцветных фигур состоит в том, что любая из них может заменить другую. Каким бы шрифтом мы ни печатали слово в словаре, оно останется таким же словом. Кажется очень естественным предположить, что в данной ситуации взаимозаменимы те и только те объекты, которые обладают одним и тем же набором формальных признаков, существенных в данной ситуации. В следующем параграфе мы убедимся, что это предположение справедливо и ему можно придать точный смысл, если само понятие одинаковости, или взаимозаменимости, сформулировано точно.

Пусть теперь M — некоторое множество объектов, в котором некоторые объекты взаимозаменимы. Обозначим через M_x множество всех объектов, взаимозаменимых с объектом x. Очевидно, что $x \in M_x$ и объединение всех M_x (при всевозможных x из M) совпадает со всем множеством M:

$$M = \bigcup_{x \in M} M_x. \tag{2.1}$$

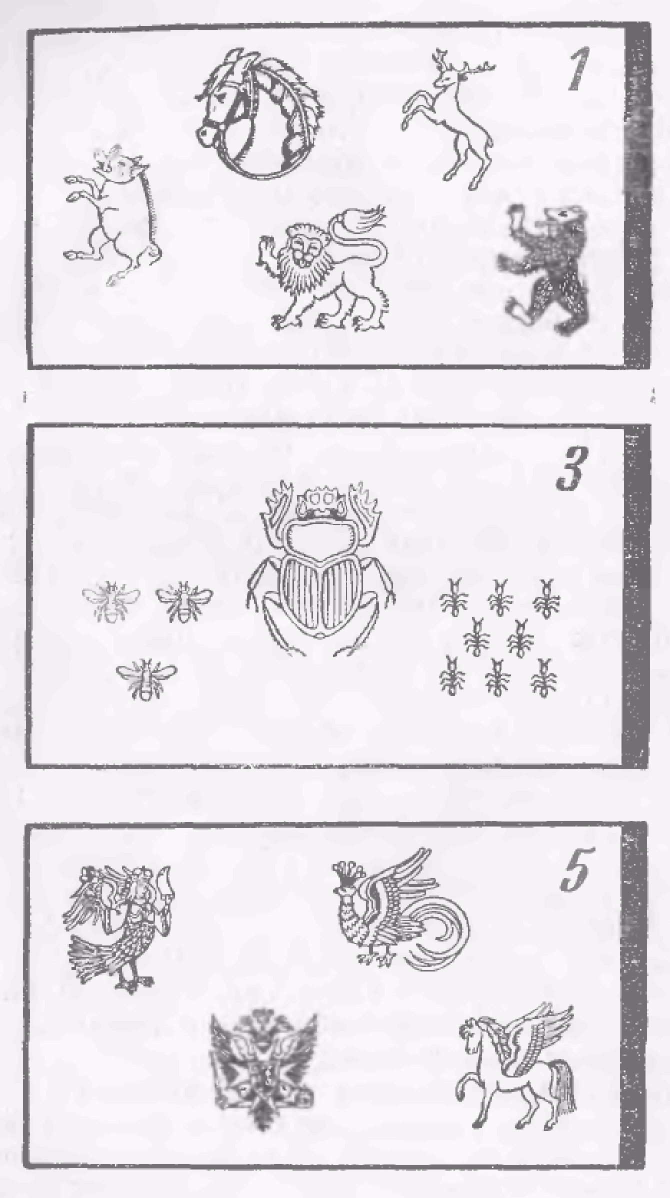


Рис. 2.1. Классы эквивалентности.

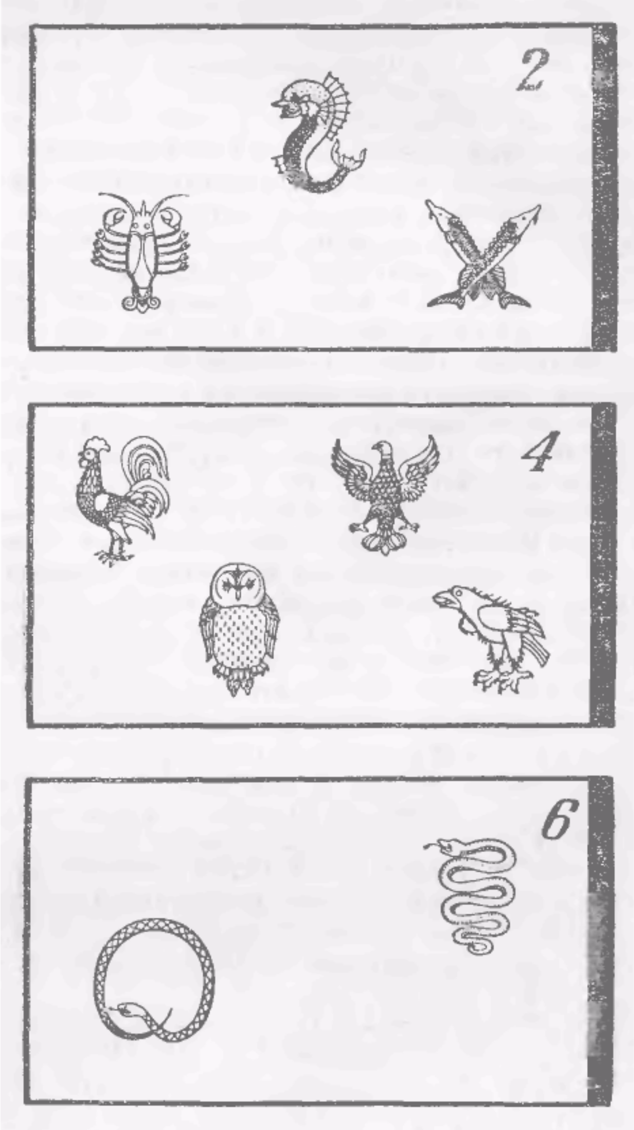


Рис. 2.1 (продолжение).   Классы эквивалентности.

Предположим, что $M_x \cap M_y \neq \varnothing$. Это значит, что существует некоторый элемент *z* такой, что он одновременно принадлежит M_x и M_y. Значит, *x* взаимозаменим с *z* и *z* взаимозаменим с *y*. Следовательно, *x* взаимозаменим с *y,* а значит и с любым элементом из M_y. Таким образом, $M_x \supseteq M_y$. Симметричным рассуждением можно показать, что $M_y \supseteq M_x$. Таким образом, встречающиеся в объединении (2.1) множества M_x либо целиком совпадают, либо не пересекаются.

Проведенное выше рассуждение наводит нас на мысль, как можно строго определить отношение одинаковости, или взаимозаменимости. В связи с этим полезно обратить внимание на способ употребления слов в математике. До сих пор мы имели дело со словами «одинаковость», «взаимозаменимость» (в данной ситуации). Эти слова никак не определялись, а использовались так, как мы привыкли их употреблять в обыденной речи. Но теперь, когда мы хотим дать точное определение (экспликацию), мы выберем и новое название. А именно, мы сейчас определим **отношение эквивалентности**, которое является **экспликацией понятия одинаковости**. Все же предыдущие соображения следует рассматривать как мотивировку именно такой экспликации.

Определение 2.1. Систему (нам совершенно несущественно, конечна эта система или бесконечна) непустых подмножеств $\{M_1, M_2, \dots\}$ множества *M* мы будем называть *разбиением* этого множества, если

1) $M = M_1 \cup M_2 \cup \dots$

и

2) $M_i \cap M_j = \varnothing$ при $i \neq j$.

Сами множества M_1, M_2, ... называются при этом *классами* данного разбиения.

Определение 2.2. Отношение *A* на множестве *M* называется *эквивалентностью* (или *отношением эквивалентности),* если существует разбиение $\{M_1, M_2, \dots\}$ множества *M* такое, что соотношение *xAy* выполняется тогда и только тогда, когда *x* и *y* принадлежат некоторому общему классу M_i данного разбиения.

Пусть $\{M_1, M_2, \dots\}$ - разбиение множества *M.*

Определим, исходя нз этого разбиения, отношение *A* на *M*: *xAy,* если *x* и *y* принадлежат некоторому общему классу M_i данного разбиения. Очевидно, отношение *A* является эквивалентностью. Назовем *A* отношением эквивалентности, *соответствующим* исходному разбиению.

Например, на рис. 2.1 изображено разбиение некоторого множества животных на шесть подмножеств. Соответствующее отношение эквивалентности — это определенное выше отношение одинаковости.

Еще пример: разбиение состоит из подмножеств множества *М,* содержащих ровно по одному элементу. Соответствующее отношение эквивалентности есть отношение равенства *Е.* Наконец, если разбиение множества *М* состоит из одного подмножества, совпадающего с самим *М,* то соответствующее отношение эквивалентности есть полное отношение: любые два элемента являются эквивалентными.

Читатель легко убедится, что пустое отношение (на непустом множестве!) не яиляется эквивалентностью.

Мы подошли к эквивалентности через понятие взаимозаменимости. Но что значит, что два объекта x и y взаимозаменимы в данной ситуации? Это всегда можно понимать так, что каждый из них содержит всю информацию о другом объекте, небезразличную в данной ситуации. Это утверждение не столь уж глубоко; оно означает только то, что взаимозаменимость объектов есть совпадение признаков, существенных в данной ситуации.

Например, пусть мы считаем одинаковыми автомобили, выпущенные в одной и той же серии одним и тем же заводом. Тогда, разобрав один экземпляр «Волги», мы в принципе можем составить комплект рабочих чертежей, который годится для выпуска однотипных «Волг». Однако, изучив один экземпляр «Волги», мы не можем угадать окраску кузова или характер вмятин на бампере у других односерийных экземпляров.

Когда мы выбираем из комплекта одну шахматную фигуру, то мы знаем, куда ее можно поставить в начальной позиции и как ходят все взаимозаменяемые с ней, т. е. одноименные и одноцветные, фигуры. В примере с животными на рис. 2.1, если мы выберем крылатого коня — пегаса, то уже тем самым знаем, что все эквивалентные ему животные возникли из мифов. А это и есть вся информация, существенная в данной классификации.

В данном случае все очень примитивно — объект заключает в себя полную информацию о каждом из эквивалентных ему объектов и не несет никакой информации о всех остальных объектах. Но для других типов отношений эта идея оценки информации, содержащейся в данном объекте относительно другого объекта, может быть развита несколько глубже.

Пусть теперь задано разбиение $\{M_1, M_2, \ldots\}$ множества *М.* Выберем в каждом множестве M_i некоторый содержащийся в нем элемент x_i. Этот элемент мы будем называть *эталоном* для всякого элемента *у,* входящего в то же множество M_i. Мы будем — по определению —

полагать выполненным соотношение x_iAy. Так определенное отношение A назовем отношением *«быть эталоном»*.

Легко видеть, что эквивалентность $\langle A\rangle$, соответствующая исходному разбиению, может быть определена так: $y\langle A\rangle z$, если y и z имеют общий эталон: x_iAy и x_iAz.

Ясно, что любое отношение эквивалентности может быть таким образом определено с помощью отношения «быть эталоном» и, наоборот, любое отношение «быть эталоном» определяет некоторую эквивалентность.

Пусть A — отношение эквивалентности, а Эт_A — такое отношение «быть эталоном», что xAy выполнено в том и только том случае, когда x и y имеют обший эталон z.

Иначе говоря, xAy равносильно существованию такого z, что $z\,\text{Эт}_A\,x$ и $z\,\text{Эт}_A\,y$. Поскольку $z\,\text{Эт}_A\,x = x(\text{Эт}_A)^{-1}z$, это означает, что

$$A = (\text{Эт}_A)^{-1}\,\text{Эт}_A.$$

Иначе говоря, эквивалентность можно алгебраически выразить через более простое отношение «быть эталоном». То, что отношение «быть эталоном» устроено более просто, видно из следующих соображений. Отношение Эт_A на множестве из n элементов можно задать графом, имеющим ровно n—m стрелок, где m — число классов эквивалентности: каждый элемент соединяется со своим единственным эталоном (эталоны можно и не соединять сами с собой). Граф, изображающий отношение эквивалентности, состоит из m полных подграфов, содержащих по n_i вершин $(n_1 + n_2 + \ldots + n_m = n)$. Таким образом, общее число ребер в этом графе равно

$$\sum_{i=1}^{m} \frac{n_i(n_i-1)}{2}.$$

Пример. Рассмотрим в качестве M множество всех целых неотрицательных чисел и возьмем его разбиение на множество M_0 четных чисел и множество M_1 нечетных чисел. Соответствующее отношение эквивалентности на множестве целых чисел обозначается так:

$$n \equiv m \pmod 2$$

и читается: *n сравнимо с m по модулю* 2. В качестве эталонов здесь естественно выбрать 0 — для четных чисел и 1 — для нечетных чисел. Аналогично, разбивая то же множество M на k подмножеств $M_0, M_1, \ldots, M_{k-1}$, где M_j состоит из всех чисел, дающих при делении на k в остатке j, мы придем к отношению эквивалентности:

$$n = m \pmod k,$$

которое выполняется, если n и m имеют одинаковый остаток при делении на k. В качестве эталона в каждом M_j естественно выбрать соответствующий остаток j.

2.1.2. Формальные свойства эквивалентности

Мы определили выше отношения эквивалентности с помощью разбиений, т. е. фактически задали их некоторой конструкцией. Можно было бы и по-другому определить эквивалентности: можно сформулировать свойства (аксиомы), которые выделяют отношения эквивалентности среди прочих бинарных отношений. Вместо определения 2.2 мы можем ввести следующее

Определение 2.3. Отношение A на множестве M называется *эквивалентностью* (или *отношением эквивалентности)*, если оно рефлексивно, симметрично и транзитивно.

Мы сейчас нарушили правила хорошего тона, принятые обычно в математике, тем, что дали два независимых определения одного и того же понятия. Сделали мы это для того, чтобы показать и **сравнить два разных способа введения математических понятий**: **<u>конструктивный и аксиоматический</u>.** Но теперь нам следует убедиться, что кроме правил хорошего тона ничто не нарушено, т. е. что оба определения эквивалентности равносильны. Соответствующим оправданием послужит

Теорема 2.1. *Если отношение A на множестве M рефлексивно, симметрично и транзитивно, то существует разбиение* $\{M_1, M_2, \ldots\}$ *множества M такое, что соотношение xAy выполнено в тех и только тех случаях, когда x и y принадлежат общему классу разбиения.*

Обратно: *если задано разбиение* $\{M_1, M_2, \ldots\}$ *множества M и бинарное отношение A определено как* «принадлежать общему классу разбиения», *то A рефлексивно, симметрично и транзитивно.*

Доказательство первой части. Рассмотрим рефлексивное, симметричное и транзитивное отношение A на M. Пусть для любого $x \in M$ множество M_x состоит из всех таких элементов z, для которых xAz.

Лемма. Для любых x и y либо $M_x = M_y$, либо

$$M_x \cap M_y = \varnothing.$$

Доказательство леммы. Пусть пересечение $M_x \cap M_y$ не пусто. Покажем, что $M_x = M_y$. Пусть $z \in M_x \cap M_y$, тогда выполнено xAz и yAz по самому определению множеств M_x и M_y. По симметричности имеем zAy, а по транзитивности из xAz и zAy следует xAy. Возьмем

теперь произвольный элемент $w \in M_y$. По определению *yAw*. Но из *xAy* и *yAw* следует *xAw*, т. е. $w \in M_x$. Итак, $M_y \subseteq M_x$.
Возьмем произвольный элемент $v \in M_x$; для него выполнено *xAv*. По симметричности отношения *A* имеем *yAx*. Но из *yAx* и *xAv* следует *yAv*. Значит, $v \in M_y$. Тем самым, мы показали, что $M_x \subseteq M_y$. В итоге можно заключить, что $M_x = M_y$. Лемма доказана.
Из леммы и рефлексивности отношения *A* следует, что множества вида M_x образуют разбиение множества *M*. (Это разбиение естественно назвать разбиением, *соответствующим* исходному отношению.) Пусть теперь выполнено соотношение *xAy*. Это значит, что $y \in M_x$. Но и $x \in M_x$ в силу *xAx*. Следовательно, оба элемента *x* и *y* входят в M_x. Итак, если *xAy,* то *x* и *y* входят в общин класс разбиения. Наоборот, пусть $u \in M_x$ и $v \in M_x$. Покажем, что *uAv* выполнено. Действительно, имеем *xAu* и *xAv*. Отсюда по симметричности *uAx*. По транзитивности из *uAx* и *xAv* следует *uAv*. Первая часть теоремы доказана.
Доказательство второй части. Пусть дано разбиение $\{M_1, M_2, \ldots\}$ множества *M.* Так как объединение всех классов разбиения совпадает с *M,* то всякий $x \in M$ входит в некоторый класс M_i. Отсюда следует *xAx,* т. е. отношение *A* рефлексивно. Если *x* и *y* входят в класс M_i, то *y* и *x* входят в тот же класс. Это означает, что из *xAy* вытекает *yAx,* т. е. отношение *A* симметрично. Пусть теперь выполнено *xAy* и *yAz*. Это означает, что *x* и *y* входят в класс M_i а *y* и *z* — в класс M_j. Поскольку M_i и M_j имеют общий элемент *y,* M_i и M_j совпадают. Значит, *x* и *z* входят в M_i, т. е. выполнено *xAz*. Итак, отношение *A* транзитивно, чем и завершается доказательство теоремы 2.1.
Заметим, что мы нигде не пользовались предположением о конечности ни множества *M,* ни его разбиения.
Из доказанной теоремы легко получается
Теорема 2.2. *Если M — конечное множество и A — отношение эквивалентности на нем, то существуют такие n и m, что каждому элементу* $x \in M$ *можно сопоставить кортеж* (упорядоченный набор) *из n + m двоичных признаков* (нулей или единиц):

$$x \to \langle \xi_1, \xi_2, \ldots, \xi_n; \xi_{n+1}, \ldots, \xi_{n+m} \rangle,$$

$$y \to \langle \eta_1, \eta_2, \ldots, \eta_n; \eta_{n+1}, \ldots, \eta_{n+m} \rangle$$

и т. д.,

так что 1) *разным элементам соответствуют разные кортежи признаков и* 2) *для того, чтобы было xAy, необходимо и достаточно,* *чтобы первые n признаков этих элементов совпадали:* $\xi_1 = \eta_1,\ \xi_2 = \eta_2,\ \ldots,\ \xi_n = \eta_n$.

Доказательство. Возьмем разбиение $\{M_1, M_2, \ldots\}$ множества *M,* соответствующее отношению *A.* В силу конечности множества *M* это разбиение конечно и каждый класс конечен. Перенумеруем элементы каждого класса. Тогда каждому элементу *x* можно сопоставить пару целых чисел: $x \to \langle p, q \rangle$, где *p*— номер класса M_p, в который попал *x,* а *q* — номер элемента *x* в своем классе. Ясно, что если $x \to \langle p_1, q_1 \rangle$, $y \to \langle p_2, q_2 \rangle$ и $x \neq y$, то $\langle p_1, q_1 \rangle \neq \langle p_2, q_2 \rangle$. Действительно, либо элементы *x* и *y* попали в разные классы — тогда у них различные первые номера: $p_1 \neq p_2$; либо они различаются номером в классе— тогда $q_1 \neq q_2$. Представим теперь числа *p* и *q* в двоичной системе счисления. Пусть *n* — наибольшее число разрядов у чисел *p,* а *m* — наибольшее число разрядов у чисел *q*. Если некоторое *p* имеет меньше, чем *n* разрядов, то дополним его слева нулями. Так же поступим и со вторыми номерами. Тем самым каждому элементу будет сопоставлен кортеж из $n + m$ двоичных признаков.

Для завершения доказательства достаточно заметить, что эквивалентность элементов *x* н *y* означает попадание в общий класс, т. е. совпадение первых номеров (первых *n* признаков).

Эта теорема оправдывает сделанное ранее утверждение, что любая эквивалентность (правда, на конечном множестве) может быть задана как совпадение некоторого набора общих признаков.

Итак, оба наши определения эквивалентности равносильны. Но теперь возникает вопрос, не являются ли некоторые аксиомы эквивалентности излишними. Например, быть может, из рефлексивности и симметричности уже следует транзитивность отношения? Далее мы будем изучать рефлексивные и симметричные отношения и увидим, что для них транзитивность вовсе не обязательна. Мы будем также заниматься рефлексивными и транзитивными отношениями и покажем, что они не обязаны быть симметричными. Наконец, попробуем доказать следующее

Утверждение. *Если отношение A симметрично и транзитивно, то оно рефлексивно.*

Будем рассуждать так. Возьмем произвольный элемент *x* и такое *y,* что выполнено соотношение *xAy.* Тогда, в силу симметричности, верно и соотношение *yAx.* Напишем эти соотношения рядом: *xAy* и *y Ax.* В силу транзитивности отсюда следует *xAx,* т. е. *A* рефлексивно. Предоставим читателю поразмыслить над тем, действительно ли мы доказали сформулированное утверждение.

П р и м е р. Пусть M — система каких-то множеств. Ранее мы определили, какие множества называются равномощными. Тем самым на M задано бинарное отношение «быть равномощными». Равномощность пары множеств V и W мы будем символически обозначать через $V \sim W$. По определению $V \sim W$ означает, что существует биективное отображение $\varphi\colon V \to W$. Ясно, что $V \sim V$, поскольку единичное отображение $\varepsilon_V\colon V \to V$ является биективным. Если существует биективное отображение $\varphi\colon V \to W$, то обратное отображение $\varphi^{-1}\colon W \to V$ также биективно, т. е. из $V \sim W$ вытекает $W \sim V$. Наконец, пусть выполнены соотношения $V \sim W$ и $W \sim V$. Тогда существуют биективные отображения $\varphi\colon V \to W$ и $\psi\colon W \to U$. Легко видеть, что их произведение $\varphi\psi = \theta$ есть биективное отображение $\theta\colon V \to U$ и, таким образом, $V \sim U$. Итак, мы доказали, что «равномощность» есть рефлексивное, симметричное и транзитивное отношение на системе M. Тем самым множества из произвольной системы M можно разбить на классы равномощных между собой. Например, если наша система множеств M состоит из всех подмножеств числовой прямой (т. е. множества действительных чисел), то она разбивается на классы из пустого множества, одноэлементных множеств, двухэлементных и т. д. Среди бесконечных множеств имеется по крайней мере два класса — счетные множества и множества, равномощные всей прямой (множества мощности континуума). Вопрос о существовании других классов бесконечных множеств составляет предмет так называемой проблемы континуума. Мы не беремся здесь обсуждать, в чем состоит недавно полученный Коэном замечательный результат, в некотором смысле решающий эту проблему.

Вернемся к обсуждению отношения *A: «x* является эталоном для *y*». Мы уже дали конструктивное определение этого отношения. Из него легко можно получить следующие свойства отношения A (быть эталоном):

1) для всякого y существует эталон x: xAy.
2) Если xAy, то xAx, т. е. любой эталон есть эталон для самого себя.
3) Эталон единствен, т. е. из xAy и zAy следует $x = z$.

Эти три свойства можно объявить аксиомами отношения «быть эталоном». Покажем, что из них следует определение эталона с помощью разбиения. Для этого сначала по отношению A построим новое отношение $\langle A \rangle$, определяемое правилом: $x\langle A \rangle y$, если x и y имеют общий эталон. Иначе говоря, если существует такое z, что zAx и zAy. Покажем, что $\langle A \rangle$ есть отношение эквивалентности. Действительно, по свойству 1) у каждого x есть эталон и, стало быть,

$x\langle A\rangle x$. Значит, $\langle A\rangle$ рефлексивно. Симметричность отношения $\langle A\rangle$ очевидна. Если $x\langle A\rangle y$ и $y\langle A\rangle z$, то это значит, что *x* и *y* имеют общий эталон, а *y* не может иметь эталона, отличного от эталона для *z*. Значит, $x\langle A\rangle z$.

Итак, доказано, что $\langle A\rangle$ есть отношение эквивалентности. Но тогда по теореме 2.1 существует разбиение $\{M_1, M_2, \ldots\}$ множества *M* на классы эквивалентных друг другу элементов — так называемые *классы эквивалентности.*

Очевидно, каждый класс эквивалентности M_i состоит из всех элементов, имеющих общий эталон x_i. По свойству 2) x_iAx_i и, значит $x_i \in M_i$. Таким образом, отношение *A,* определенное аксиоматически свойствами 1)—3), всегда может быть задано разбиением с выбранными представителями (эталонами) в каждом классе.

Пусть $\varphi\colon M \to S$— сюръективное отображение множества *M* на некоторое множество *S*. Рассмотрим на множестве *M* отношение «иметь общий образ» и обозначим это отношение A_φ. Иначе говоря, $xA_\varphi y$, если $\varphi(x) = \varphi(y)$. Обозначим через M_ξ множество всех элементов $x \in M$, имеющих данный образ $\xi \in S$, т. е. таких, что $\varphi(x) = \xi$. Ясно, что

$$\bigcup_{\xi\in S} M_\xi = M,$$

так как любой элемент из *M* имеет образ. Далее, при разных ξ и η, $M_\xi \cap M_\eta = \varnothing$, так как иначе элемент, попавший в пересечение $M_\xi \cap M_\eta$, имел бы два разных образа: ξ и η. Поскольку φ сюръективно, $M_\xi \neq \varnothing$ для любого $\xi \in S$. Итак, множества M_ξ образуют разбиение множества *M,* а отношение A_φ есть эквивалентность, соответствующая этому разбиению. Последнее следует из того, что $xA_\varphi y$ тогда и только тогда, когда *x* и *y* принадлежат общему множеству M_ξ.

Множество классов эквивалентности по отношению *A* принято обозначать M/A (читается: *фактор-множество множества M по отношению A).* Наши рассуждения показывают, что для всякого сюръективного отображения $\varphi\colon M \to S$ существует отношение эквивалентности *A* на множестве *M* такое, что *M/A* и S могут быть поставлены во взаимно-однозначное соответствие.

Наоборот, если имеется произвольное отношение эквивалентности *A* на *M,* то по нему можно построить отображение $\varphi\colon M \to S$, где $S = M/A$ и $\varphi(x)$ есть класс эквивалентности, содержащий *x.* Легко проверить, что φ сюръективно и построенное по этому отображению отношение эквивалентности A_φ есть исходное отношение *A.*

Рассмотрим частный случай, когда $\varphi\colon M \to S$ и $S \subseteq M$. Пусть, далее, отображение φ обладает тем свойством, что, при $x \in S,\ \varphi(x) = x$ или, как говорят в таких случаях, подмножество *S* *неподвижно* при отображении φ. Отсюда видно, что φ сюръективно. Действительно, всякий $x \in S$ есть образ по крайней мере самого $x\colon x = \varphi(x)$. Итак, каждому $y \in M$ однозначно сопоставлен некоторый элемент $x \in S$. При этом, если *x* сопоставлен какому-то элементу, то самому *x* сопоставлен он же.

Сравнивая с соответствующими свойствами, определяющими отношение «быть эталоном», мы видим, что отображение $\varphi\colon M \to S$ множества *M* на неподвижное подмножество *S* задает на *M* отношение *A* «быть эталоном» так, что *xAy* в том и только том случае, когда $\varphi(y) = x$.

Посмотрим теперь, что получится, если отказаться от условия, что φ определено на всем *M.* Рассмотрим функцию $\varphi\colon M \to S$, которая некоторым элементам *x* из *M* сопоставляет единственный образ *q(x)* из S. По отображению φ можно опять-таки построить отношение A_φ по правилу: $xA_\varphi y$, если $\varphi(x) = \varphi(y)$. Легко проверить, что A_φ будет симметрично и транзитивно. Выберем подмножество $M_0 \subseteq M$, состоящее из тех элементов, на которых определено отображение φ. Тогда если либо *x,* либо *y* не принадлежат M_o, то $xA_\varphi y$ заведомо не выполняется. Значит, если *x* не входит в M_0, то $xA_\varphi x$ также не выполнено. Следовательно, отношение A_φ теперь уже не обязано быть рефлексивным.

Читатель, дойдя до этого места, наверняка уже нашел ошибку в «доказательстве» того, что рефлексивность отношения следует из симметричности и транзитивности. Она состояла в том, что мы незаконно предположили для произвольного $x \in M$ существование такого *y,* что *xAy.* Для вышеопределенного отношения A_φ видно, что как раз при *x,* не входящих в M_0 (в область определения отображения φ), $xA_\varphi y$ не выполнено ни для какого *y.* Отсюда сразу видно, как построить содержательный пример симметричного и транзитивного, но не рефлексивного отношения. Пусть *M* — множество людей, а отношение *A* означает «быть уроженцем одного города». Легко видеть, что *A* симметрично и транзитивно, но если *x* родился не в городе, а в деревне, или, вообще, во время путешествия по морю, то *xAx* не выполиено. В этом примере S — множество городов, а отображение $\varphi\colon M \to S$ сопоставляет каждому человеку город, где он был рожден.

Из сказанного видно также, что условие рефлексивности можно в определении эквивалентности заменить более слабым. Достаточно потребовать, чтобы для каждого *x* существовал такой элемент *y,* что выполнено либо *xAy,* либо *yAx.* Тогда из этого свойства, а также симметричности и транзитивности можно получить рефлексивность отношения *А*.

Граф, изображающий отношение эквивалентности, выглядит следующим образом. Пусть *М* — множество его вершин. Тогда $M=\bigcup_i M_i$, где M_i— классы эквивалентности. Ясно, что в каждом подмножестве M_i все вершины соединены друг с другом. Но никакая из них не соединена с вершинами, не входящими в M_i. Итак, граф, изображающий отношение эквивалентности, состоит из отдельных, не связанных друг с другом полных подграфов.

На рис. 2.2 представлен граф отношения эквивалентности на множестве $M = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$ с классами эквивалентности $M_1= \{1,2\}$, $M_2= \{3\}$, $M_3 = \{4,5,6,7\}$, $M_4= \{8, 9, 10\}$.

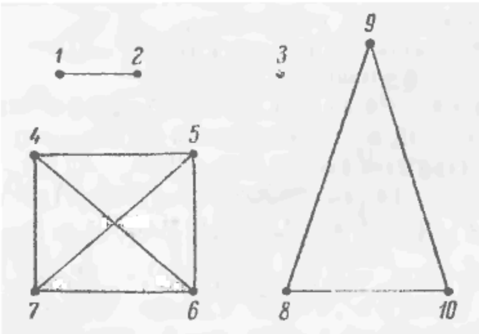


Рис. 2.2. Граф эквивалентности.

В соответствии с тем, что говорилось ранее, на графах отношений эквивалентности можно не изображать петель и стрелок. Так мы и поступили.

Пусть в нашем распоряжении имеются теперь два множества: M_1 и M_2, на каждом из которых задано отношение эквивалентности (соответственно, A_1 и A_2). Спрашивается: каким способом можно из них соорудить одно множество с отношением эквивалентности на нем? Вспомним, что, строго говоря, отношение — это пара $\langle A, M\rangle$, где *М* — множество элементов, вступающих в отношение, а *А* — множество пар, для которых выполняется данное отношение.

Один из наиболее простых типов композиции отношений дает следующее

Определение 2.4. *Прямой суммой* отношений $\langle A_1, M_1\rangle$ и $\langle A_2, M_2\rangle$ называется отношение $\langle A_1 \cup A_2, \quad M_1 \cup M_2\rangle$. Прямую: сумму отношений $\langle A_1, M_1\rangle$, $\langle A_2, M_2\rangle$ мы будем обозначать через $\langle A_1, M_1\rangle \oplus \langle A_2, M_2\rangle$. Итак,

$$\langle A_1, M_1\rangle \oplus \langle A_2, M_2\rangle = \langle A_1 \cup A_2, M_1 \cup M_2\rangle.$$

Таким образом, если $\langle A_1, M_1\rangle \oplus \langle A_2, M_2\rangle = \langle A, M\rangle$, то $M = M_1 \cup M_2$ и $A = A_1 \cup A_2$. Следовательно, соотношение *xAy* выполнено в следующих случаях:

1) $x \in M_1$, $y \in M_1$ и xA_1y;

2) $x \in M_2$, $y \in M_2$ и xA_2y.

На рис. 2.3 приведены два отношения: $\langle A_1, M_1\rangle$ и $\langle A_2, M_2\rangle$ — и их прямая сумма.

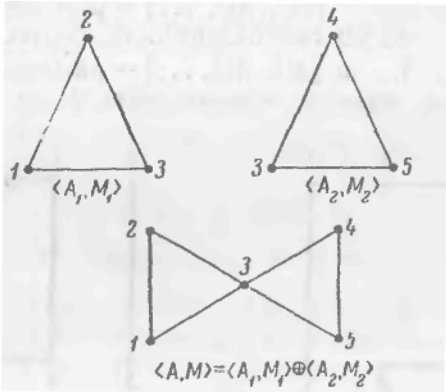


Рис. 2.3. Прямая сумма.

Из этого рисунка видно, что даже когда A_1 и A_2 — эквивалентности, прямая сумма A не обязана быть эквивалентностью. Однако, имеет место

Теорема 2.3. *Если* $M_1 \cap M_2 = \varnothing$, *а отношения* A_2 *и* A_2 — *эквивалентности, то их прямая сумма*

$$\langle A, M\rangle = \langle A_1, M_1\rangle \oplus \langle A_2, M_2\rangle$$

также является эквивалентностью.

Доказательство. Рефлексивность проверяется просто: если $x \in M_i$, то выполнено xA_ix и, следовательно, xAx. Симметричность также очевидна: если выполнено xAy, то либо x и y входят в M_1 и xA_1y, а значит, и yA_1x, т. е. yAx, либо x и y входят в M_2 и xA_2y, поэтому yA_2x и yAx. Докажем транзитивность отношения A. Пусть выполнены соотношения xAy и yAz. Рассмотрим случай, когда $x \in M_1$, $y \in M_1$ и xA_1y. Так как $M_1 \cap M_2 = \varnothing$, то y не входит в M_2. Но тогда соотношение yAz может выполняться только при $z \in M_1$ и yA_1z.

Однако, из xA_1y и yA_1z вытекает xA_1z и xAz. Случай, когда x и y принадлежат M_2, исследуется аналогично. Теорема доказана.

Замечание. Из этого доказательства видно, что условие непустоты пересечения работало только при проверке транзитивности. Значит, справедлива

Теорема 2.4. *Если отношения* $\langle A_1, M_1\rangle$ *и* $\langle A_2, M_2\rangle$ *рефлексивны и симметричны* (в частности, являются эквивалентностями), *то их прямая сумма* $\langle A_1, M_1\rangle \oplus \langle A_2, M_2\rangle$ *также рефлексивна и симметрична.*

Полное исследование условий, при которых прямая сумма эквивалентностей является эквивалентностью, можно провести с помощью теоремы 2.6.

Замечание. Если $\langle A, M\rangle = \langle A_1, M_1\rangle \oplus \langle A_2, M_2\rangle$, то каждое из отношений $\langle A_1, M_1\rangle$ и $\langle A_2, M_2\rangle$ есть сужение отношения $\langle A, M\rangle$ на свою область задания.

2.1.3. Операции над эквивалентностями

Посмотрим, какие операции над отношениями эквивалентности и при каких условиях дают в результате эквивалентность.

Первый результат такого типа был нами уже получен ранее. Мы установили там, что транзитивное замыкание транзитивного отношения совпадает с ним самим. Значит, транзитивное замыкание $\hat{A}$ отношения эквивалентности A является отношением эквивалентности.

В том же параграфе мы установили, что для симметричного отношения A обратное совпадает с ним самим: $A^{-1} = A$. Значит, отношение, обратное к эквивалентности, является эквивалентностью.

Из лемм 1.1, 1.3 и 1.7 вытекает, что если A и B - эквивалентности, то их пересечение $A\cap B$ также является отношением эквивалентности.

Пусть теперь $\{M_1^A, M_2^A, \ldots\}$— разбиение множества M на классы эквивалентности, соответствующее отношению A, а $\{M_1^B, M_2^B, \ldots\}$ — разбиение множества M на классы эквивалентности по B. Пусть $x \in M_i^A$ и одновременно $x \in M_j^B$. Элементы, с которыми x находится в отношении $A\cap B$, заполняют множество $M_i^A \cap M_j^B$. Таким образом, классы эквивалентности по $A\cap B$ суть пересечения классов эквивалентности по A и по B. Легко видеть, что система пересечений $M_i^A \cap M_j^B$ есть разбиение множества M, соответствующее отношению $A\cap B$.

Сложнее обстоит дело с объединением отношений эквивалентности. Вообще говоря, объединение эквивалентностей уже не обязано быть эквивалентностью. Это видно из примера на рис. 2.4.

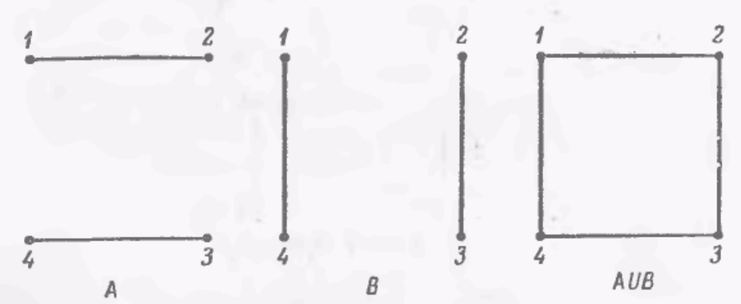


Рис. 2.4. Объединение эквивалснтностей.

Действительно, отношение *A* дает разбиение на два класса {1,2} и {3,4}, отношению *B* соответствует разбиение {{1,4}, {2,3}}, а отношение $A \cup B$ дает неполный связный граф. Теперь попробуем разобраться, когда объединение эквивалентностей дает в результате эквивалентность. Отметим сначала следующее тривиальное обстоятельство. Пусть $A \subseteq B,$ тогда из свойств теоретико-множественных операций следует

$$A \cup B = B,$$

т. е. $A \cup B$ есть эквивалентность. Точно так же, если $B \subseteq A,$ то $A \cup B$ является эквивалентностью.

Рассмотрим более общий случай, когда множество *M* можно разбить на два непересекающихся подмножества M_1 и M_2 (из которых одно может быть пустым) так, что

$$\begin{cases} \langle A,\ M\rangle = \langle A_1,\ M_1\rangle \oplus \langle A_2,\ M_2\rangle, \\ \langle B,\ M\rangle = \langle B_1,\ M_1\rangle \oplus \langle B_2,\ M_2\rangle, \end{cases} \tag{2.2}$$

и при этом

$$A_1 \subseteq B_1 \quad \text{и} \quad B_2 \subseteq A_2. \tag{2.3}$$

В этом случае отношения *A* и *B* мы назовем *когерентными.*

Легко видеть, что если $A \subseteq B$ или $B \subseteq A,$ то отношения *A* и *B* когерентны (надо положить $M_1=M,$ $M_2 = \varnothing$). Таким образом, сравнимость относительно «порядка», задаваемого включением, есть частный случай когерентности.

Из (2.3) следует, что для когерентных отношений эквивалентности *A* и *B:*

$$\langle A_1 \cup B_1,\ M_1\rangle = \langle B_1,\ M_1\rangle$$

и

$$\langle A_2 \cup B_2,\ M_2\rangle = \langle A_2,\ M_2\rangle.$$

Используя определение прямой суммы и (2.2), получаем

$$\langle A \cup B,\ M\rangle = \langle B_1,\ M_1\rangle \oplus \langle A_2,\ M_2\rangle.$$

Здесь $\langle B_1, M_1\rangle$ и $\langle A_2, M_2\rangle$ — эквивалентности (как сужения эквивалентностей $\langle B, M\rangle$ и $\langle A, M\rangle$), а M_1 и M_2 не пересекаются. По теореме 2.3 отсюда следует, что $A \cup B$ есть отношение эквивалентности. Оказывается, когерентность отношений *А, В* является не только достаточным, но и необходимым условием для того, чтобы объединение $A \cup B$ эквивалентностей *А* и *В* было эквивалентностью.

Теорема 2.5. *Для того чтобы объединение* $A \cup B$ *эквивалентностсй А и В само было отношением эквивалентности, необходимо и достаточно, чтобы А и В были когерентными.*

Нам понадобятся некоторые простые свойства разбиений на классы эквивалентности, которые мы сформулируем в виде самостоятельных лемм. Мы будем далее использовать некоторые словесные сокращения. Если *А* — эквивалентность и *xAy,* то мы будем говорить, что *х* и *у* *А-эквивалентны.* Разбиение, соответствующее эквивалентности *А,* мы будем называть *А-разбиением;* классы этого разбиения — *А-классами* и т. п.

Лемма 2.1. *Для того чтобы* $A \subseteq B$, *необходимо и достаточно, чтобы каждый А-класс содержался в некотором В-классе.*

Действительно, если $A \subseteq B$, то из *xAy* следует *xBy*. Значит, множество всех *у, А*-эквивалентных элементу *х,* содержится во множестве всех *у, В*-эквивалентых этому *х.* Обратный вывод столь же очевиден.

Л е м м а 2.2. *Для того чтобы* $B \subseteq A$, *необходимо и достаточно, чтобы каждый А-класс* M_i^A *целиком содержал любой В-класс* M_j^B, *имеющий с* M_i^A *непустое пересечение.*

Для доказательства необходимости выберем произвольный элемент $x \in M_j^B \cap M_i^A$. По предыдущей лемме M_j^B целиком содержится в некотором классе M_k^A. Но если бы M_k^A был бы отличен от M_i^A, то элемент *х* лежал бы сразу в двух классах *А*-разбнення, что невозможно. Значит, $M_j^B \subseteq M_i^A$. Для доказательства достаточности нужно только вспомнить, что из $M_j^B \cap M_i^A \neq \varnothing$ по условию вытекает $M_j^B \subseteq M_i^A$, и применить лемму 2.1.

Л е м м а 2.3. *Для того чтобы эквивалентности А и В были когерентными, необходимо и достаточно, чтобы всякий А-класс* M_i^A

либо содержался в некотором В-классе M_j^B, *либо целиком содержал любой В-класс* M_j^B, *имеющий с* M_i^A *непустое пересечение.*

(Очевидна следующая перефразировка этой леммы: *отношения эквивалентности А и В когерентны тогда и только тогда, когда любая пара классов эквивалентности* M_i^A *и* M_j^B *либо не пересекается, либо один* из *этих классов содержит другой.*)

Доказательство. Если *А* и *В* когерентны, то $M = M_1 \cup M_2,\ M_1 \cap M_2 = \varnothing$ и на М, имеем $A \subseteq B$, а на M_2 $A \supseteq B$. Тогда по лемме 2.1 для каждого класса M^A_i, содержащегося в M_2, существует такой класс $M_j^B \subseteq M_1$, что $M_i^A \subseteq M_j^B$. По лемме 2.2 каждый класс M^A_i, содержащийся в M_2, целиком содержит любой класс M^B_j, имеющий с M^A_i непустое пересечение. Поскольку M_1 и M_2 не пересекаются, из (2.2) вытекает, что всякий класс эквивалентности M^A_i содержится либо в M_1, либо в M_2; значит, наше рассуждение охватывает все классы.

Поведем доказательство в обратную сторону. Пусть каждый класс M^A_i обладает сформулированным в лемме 2.3 свойством. Обозначим через M_1 объединение всех тех классов M^A_i, для которых существует такой M^B_j, что $M_i^A \subseteq M_j^B$, а через M_2— объединение остальных классов M^A_i. Ясно, что $M_1 \cap M_2 = \varnothing$, $M_1 \cup M_2 = M$ и

$$\langle A,\ M\rangle = \langle A_1,\ M_1\rangle \oplus \langle A_2,\ M_2\rangle,$$
$$\langle B,\ M\rangle = \langle B_1,\ M_1\rangle \oplus \langle B_2,\ M_2\rangle,$$

где A_i и B_i — сужения отношений *А* и *В* на M_i. Наконец, очевидно, что $A_1 \subseteq B_1$ и $A_2 \supseteq B_2$, т. е. *А* и *В* когерентны.

Теперь мы подготовили все необходимое для доказательства теоремы 2.5. Будем вести доказательство от противного, т. е. предположим, что *А* и *В* не когерентны. Тогда по лемме 2.3 существует класс M^A_i и класс M^B_j такие, что $M_i^A \cap M_j^B \neq \varnothing$, но ни один из них не содержит другой. Значит, существует $x \in M_i^A \setminus M_j^B$, существует $y \in M_i^A \cap M_j^B$ и существует $z \in M_j^B \setminus M_i^A$ (рис. 2.5).

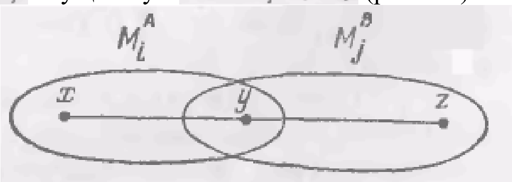


Рис. 2.5.

Имеем следующие соотношения: *xAy* и *yBz,* следовательно, $x A \cup B y$ и $y A \cup B z$. По транзитивности должно было бы быть также $x A \cup B z$. Однако, соотношения: *xAz* и *xBz* — оба не выполнены, так как *x* не лежит с *z* ни в общем *A*-классе, ни в общем *B*-классе. Значит, соотношение $x A \cup B z$ не выполнено. Полученное противоречие доказывает теорему.

Замечание. Понятие когерентности имеет смысл для любых отношений *A* и *B.* Но для эквивалентностей когерентность отношений *A* и *B* легко формулируется в терминах классов эквивалентности (лемма 2.3).

Лемма 2.4. *Если A и B рефлексивны, то*

$$A \cup B \subseteq AB. \qquad (2.4)$$

Доказательство. Если *xAy,* то, в силу *yBy,* выполнено и соотношение *xABy,* т. е. $A \subseteq AB$. Аналогично получается $B \subseteq AB$. Из этих двух включений следует (2.4).

Теорема 2.6. *Для того чтобы объединение* $A \cup B$ *эквивалентностей A и B само было отношением эквивалентности, необходимо и достаточно, чтобы*

$$AB = A \cup B. \qquad (2.5)$$

Доказательство. Пусть $A \cup B$ — эквивалентность. По лемме 2.4 $A \cup B \subseteq AB$. Для доказательства (2.5) остается доказать

$$AB \subseteq A \cup B. \qquad (2.6)$$

Пусть *xABy.* Тогда для некоторого *z* имеем *xAz* и *z By.* Следовательно, $x(A \cup B)z$ и $z(A \cup B)y$. Значит, $x(A \cup B)y$ и (2.6) доказано. Пусть теперь выполнено (2.5). По лемме 1.3 отношение $A \cup B$ симметрично. По (2.5) тогда симметрично и отношение *AB.* По лемме 1.4 *AB = BA.* По теореме 2.7 (см. ниже) получаем, что отношение *AB* — эквивалентность. Из (2.5) вытекает, что и $A \cup B$—эквивалентность. Теорема доказана.

Условие, при котором произведение *AB* двух отношений эквивалентности *A* и *B* само является эквивалентностью, было получено чешским математиком Шиком в 1954 г. Прежде всего отметим, что когда мы ранее приводили пример не коммутирующих отношений *A* и *B,* то *A* и *B* там были отношениями эквивалентности, но их произведение *AB* уже не было таковым (и даже не было симметричным). Связь с перестановочностью произведения здесь отнюдь не случайна, как показывает принадлежащая Шику

Теорема 2.7. *Для того чтобы произведение AB отношений эквивалентности A и B было эквивалентностью, необходимо и достаточно, чтобы A и B коммутировали.*

Доказательство. Пусть сначала

$$AB = BA. \qquad (2.7)$$

По лемме 1.1 *AB* рефлексивно. По лемме 1.4 *AB* симметрично. Транзитивность произведения доказывается так:

$$(AB)(AB) = A(BA)B = A(AB)B = (AA)(BB) = AB$$

— здесь мы использовали ассоциативный закон для произведения отношений, условие (2.7), а также транзитивность и рефлексивность отношений *A* и *B*. Итак,

$$(AB)(AB) = AB,$$

но это и означает транзитивность отношения *AB,* поскольку *AB* рефлексивно. Пусть теперь произведение *AB* есть эквивалентность. Тогда по лемме 1.4 *AB* = *BA*.

Ранее мы ввели еще операции $A \hat{\cup} B$ и $A \hat{\circ} B.$ Легко проверить, что если *A* и *B* — эквивалентности, то $A \hat{\cup} B$ и $A \hat{\circ} B$ также будут эквивалентностями.

Проверим это для первой операции. (Как мы увидим дальше, для второй из них надобности в проверке не будет.) Рефлексивность отношения $A \hat{\cup} B,$ вытекает из леммы 1.1. Симметричность вытекает из леммы 1.3 и следствия из леммы 1.4. Транзитивность следует из того, что любое отношение вида $\hat{C}$ транзитивно (теорема 1.4 и (1.20)).

Итак, операция $A \hat{\cup} B,$ будучи выполнена над отношениями эквивалентности, не выводит за этот класс отношений.

Оказывается, эта операция (ее иногда называют *объединением эквивалентностей,* имея в виду, что обычное объединение эквивалентностей может не быть эквивалентностью) ассоциативна, т. е. является «хорошей» алгебраической операцией.

Теорема 2.8. *Для любых транзитивных отношений A, B и C справедлив ассоциативный закон:*

$$(A \hat{\cup} B) \hat{\cup} C = A \hat{\cup} (B \hat{\cup} C). \qquad (2.8)$$

Докажем сначала две леммы.

Л е м м а 2.5. *Для любых отношений P, Q*

$$P \leqslant P \hat{\cup} Q, \qquad (2.9)$$

$$Q \leqslant P \hat{\cup} Q. \qquad (2.10)$$

(2.9) вытекает из $P \subseteq P \cup Q$ и (1.1). (2.10) доказывается аналогично.

Лемма 2.6. *Для любых транзитивных отношений P, Q, R из* $P \subseteq R$ *и* $Q \subseteq R$ *вытекает* $P \hat{\cup} Q \subseteq R.$

Из $P \subseteq R$ и $Q \subseteq R$ вытекает $P \cup Q \subseteq R.$ Из (1.17) и теоремы 1.3 получаем $P \hat{\cup} Q \subseteq R.$

Доказательство теоремы 2.8 Из леммы 2.5

$$B \subseteq B \hat{\cup} C \tag{2.11}$$

$$B \hat{\cup} C \subseteq A \hat{\cup} (B \hat{\cup} C). \tag{2.12}$$

Из (2.11) и (2.12)

$$B \subseteq A \hat{\cup} (B \hat{\cup} C). \tag{2.13}$$

Из леммы 2.5

$$A \subseteq A \hat{\cup} (B \hat{\cup} C). \tag{2.14}$$

Из (2.13), (2.14), леммы 2.6 и того, что любое отношение вида $\hat{C}$ транзитивно,

$$A \hat{\cup} B \subseteq A \hat{\cup} (B \hat{\cup} C). \tag{2.15}$$

Подобно тому, как мы доказали (2.13). доказывается

$$C \subseteq A \hat{\cup} (B \hat{\cup} C). \tag{2.16}$$

Подобно тому, как мы из (2.13) н (2.14) вывели (2.15), из (2.15) и (2.16) выводится

$$(A \hat{\cup} B) \hat{\cup} C \subseteq A \hat{\cup} (B \hat{\cup} C). \tag{2.17}$$

Из (2.17) и аналогично доказываемого «обратного» включения вытекает (2.8). Теорема доказана.

Нетрудно убедиться, что для любой эквивалентности *А*

$$A \hat{\cup} E = A, \tag{2.18}$$

где *Е*— диагональное отношение. Это следует из того, что $E \subseteq A$ (в силу рефлексивности отношения *А*); значит,

$$A \cup E = A \text{ и } A \hat{\cup} E = \hat{A} = A.$$

Покажем теперь, что операция $A \hat{\circ} B$ не дает ничего нового:

Теорема 2.9. *Если А и В— эквивалентности, то*

$$A \hat{\cup} B = A \hat{\circ} B \tag{2.19}$$

Доказательство. Заметим сначала, что, учитывая лемму 2.4,

$$A \cup B \subseteq AB \subseteq AB \cup BA = A \circ B.$$

Применяя транзитивное замыкание к обеим частям, ввиду свойства монотонности транзитивного замыкания (1.17) имеем

$$A \hat{\cup} B \subseteq A \hat{\circ} B. \tag{2.20}$$

Далее, применяя распределительный закон (1.13), получим

$$(A \cup B)^2 = A^2 \cup AB \cup BA \cup B^2 = A \cup AB \cup BA \cup B =$$

$$= AB \cup BA = A \circ B. \tag{2.21}$$

Мы использовали здесь тот факт, что для рефлексивного *А* выполнено включение $B \subseteq BA$, а следовательно, $BA \cup B = BA$.Запишем теперь выражение (1.2) для транзитивного замыкания, используя (2.21):

$$A \hat{\cup} B = (A \cup B) \cup (A \cup B)^2 \cup (A \cup B)^3 \cup (A \cup B)^4 \cup \ldots =$$
$$= (A \cup B) \cup (A \circ B) \cup (A \cup B)^3 \cup (A \circ B)^2 \cup \ldots$$

Отсюда ясно, что

$$A \hat{\cup} B \supseteq (A \circ B) \cup (A \circ B)^2 \cup \ldots,$$

т. е.

$$A \hat{\cup} B \supseteq A \hat{\circ} B. \quad (2.22)$$

Сравнивая включения (2.20) и (2.22), получим искомое соотношение (2.19).

Отсюда вытекает следующий результат, также принадлежащий Шику:

Теорема 2.10. *Если А и В — эквивалентности и АВ = ВА, то*

$$AB = A \hat{\cup} B. \quad (2.23)$$

В самом деле, по теореме 2.7 произведение *АВ* является эквивалентностью, а стало быть отношение $AB = AB \cup BA = A \circ B$ совпадает со своим транзитивным замыканием: $AB = A \hat{\circ} B.$ Но тогда из теоремы 2.9

$$AB = A \hat{\cup} B.$$

На этом мы закончим исследование свойств операций над эквивалентностями.

Полученные результаты об операциях над отношениями допускают алгебраическую интерпретацию. Множество $\mathfrak{M}$ всех отношений на *М* обладает структурой моноида относительно операции произведения отношений (если на некотором множестве $\mathfrak{M}$ определена ассоциативная операция и существует элемент *Е,* ведущий себя, как единица при этой операции, то говорят, что на множестве $\mathfrak{M}$ задана *структура моноида,*). Пусть множество $\mathfrak{M}_э \subseteq \mathfrak{M}$ состоит из всех отношений эквивалентности. Любое подмножество множества $\mathfrak{M}_э,$ замкнутое относительно операции произведения *АВ,* является коммутативным моноидом (теорема 2.7). Так как для множества *М,* содержащего не менее трех элементов, существуют некоммутирующие эквивалентности, то само $\mathfrak{M}_э$ не образует моноида относительно произведения отношений. Однако $\mathfrak{M}_э$ обладает структурой коммутативного моноида относительно операции $A \hat{\cup} B$ (теорема 2.8) или, что то же самое $A \hat{\circ} B$, (теорема 2.9). На подмоноидах из $\mathfrak{M}_э$ моноида $\mathfrak{M}$ (относительно *АВ)* операция $A \hat{\cup} B$ совпадает с операцией произведения *АВ* (теорема 2.10).

2.1.4. Отношения эквивалентности на числовой прямой

Пусть задано отношение *A* на множестве *M*. В случае, когда *M* — числовая прямая, отношение *A* отождествляется с некоторым подмножеством числовой плоскости, т. е. прямого произведения *M*×*M*. В этом параграфе будут рассмотрены геометрические свойства множества *A* на плоскости в случае, когда отношение *A* есть эквивалентность.

Согласно определению 2.3 отношение *A* называется *эквивалентностью,* если оно рефлексивно, симметрично и транзитивно. Каждое из этих свойств порождает некоторое геометрическое свойство множества *A*. Координаты точки на плоскости будем обозначать *{x,y)*.

1. Рефлексивность. Из того, что *xAx* для всех *x,* следует, что множество *A* содержит главную диагональ *(свойство R).*

2 Симметричность. Симметричность означает, что если $\langle x, y\rangle \in A$, то и $\langle y, x\rangle \in A$, т. е. что множество *A* симметрично относительно главной диагонали *(свойство S).*

3. Транзитивность. Транзитивность означает, что если $\langle x, y\rangle \in A$ и $\langle y, z\rangle \in A$, то и $\langle x, z\rangle \in A$.

Точка $\langle x, z\rangle$ является четвертой вершиной прямоугольника, три вершины которого находятся в точках $\langle x, y\rangle$, $\langle y, z\rangle$ и $\langle y, y\rangle$. Заметим, что вершина $\langle y, y\rangle$ лежит на биссектрисе координатного угла — главной диагонали координатной плоскости. Поэтому геометрически свойство транзитивности можно сформулировать следующим образом:

Множество A на плоскости определяет транзитивное отношение тогда и только тогда, когда для любого прямоугольника, одна вершина которо одна вершина которого α *лежит на главной диагонали, а две соседние с* α *вершины принадлежат A, вершина* α*', противоположная* α*, также принадлежит A* (*свойство* T_1; см. рис. 2.6).

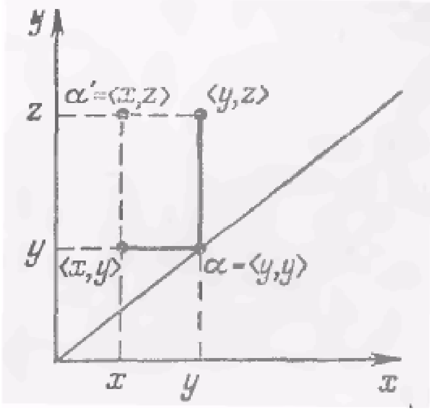


Рис. 2.6. Геометрический смысл транзитивности.

Замечание. Если отношение A является симметричным, то геометрическая формулировка транзитивности несколько упрощается. А именно:

Множество A на плоскости, симметричное относительно главной диагонали, определяет транзитивное отношение тогда и только тогда, когда для любого прямоугольника, одна вершина которого лежит на главной диагонали, а две другие принадлежат A, четвертая вершина также принадлежит A (свойство T_2).

Разница с предыдущим утверждением состоит в том, что вершины, принадлежащие A, не обязаны быть соседними с вершиной, лежащей на диагонали. Покажем, что для симметричного A свойство T_1 влечет T_2. Пусть, например, вершина, лежащая на диагонали, имеет координаты $\langle y, y\rangle$ и $\langle x, z\rangle \in A$ и $\langle y, z\rangle \in A$; покажем, что $\langle x, y\rangle \in A$. В самом деле, в силу симметрии, вместе с $\langle y, z\rangle \in A$ имеем $\langle z, y\rangle \in A$. Если в качестве вершины на диагонали взять теперь $\langle z, z\rangle$, а в качестве соседних с ней вершин, принадлежащих A, $\langle x, z\rangle$ и $\langle z, y\rangle$, то, в силу свойства T_1, получаем $\langle x, y\rangle \in A$.

Заметим, что класс эквивалентности, содержащий точку x_0, есть проекция пересечения множества A и прямой $x = x_0$ на ось ординат.

Сейчас мы приведем некоторые примеры множеств на плоскости, определяющих отношение эквивалентности.

Пример 1. Множество A — вся плоскость. Выполнение свойств R, S, T очевидно. Все точки исходной прямой M отождествляются, т. е. входят в один класс эквивалентности.

Замечание. Для любого ε > 0, если множество *A*, определяющее отношение эквивалентности, содержит полосу $|x-y|<\varepsilon$, то оно совпадает со всей плоскостью. В самом деле, из рис. 2.7 видно, что вместе с любой точкой $\langle y, y\rangle$ множество *A* содержит все внутренние точки квадрата с вершинами

$$\langle y-\varepsilon, y\rangle,\ \langle y, y\rangle,\ \langle y, y+\varepsilon\rangle,\ \langle y-\varepsilon,\ y+\varepsilon\rangle,$$

т. е. полосу $|x-y|<2\varepsilon$.

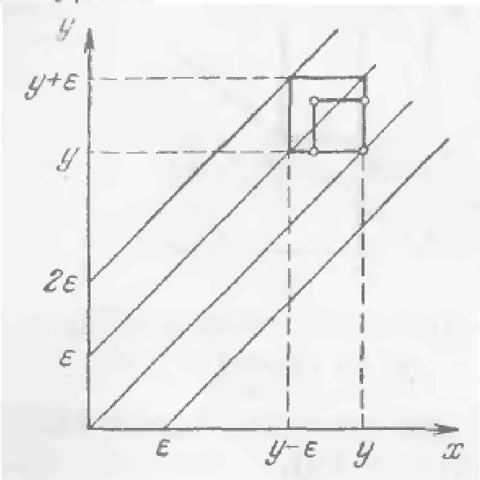


Рис 2.7.

Ясно, что таким образом свойство «принадлежать *A*» распространяется на все точки плоскости.

Пример 2 (периодичность). Возьмем некоторое число. Пусть множество *A* состоит из прямых *x* — *y* = — *kc*, где *k* — произвольное целое число. Выполнение свойств *R* и S очевидно, и если $x-y=k_1c,\ y-z=k_2c,$ то $x-z=(k_1+k_2)c.$

Пример 3. «Все константы равны единице, кроме нуля». (Такое утверждение высказал И. М. Гельфанд.) В этом примере множество *A* есть вся плоскость с выброшенными осями координат и добавленным началом координат. Иначе говоря, $\langle x, y\rangle\in A$ всегда, кроме случая *x*=0, *y*≠0 и ему симметричного. Если точки $\langle x, y\rangle,\ \langle y, z\rangle$ принадлежат *A*, то либо *x* = 0, и тогда *y* = 0, z = 0, либо *x* ≠0, и тогда *y* ≠ 0 и z ≠0. В обоих случаях $\langle x, z\rangle\in A$ (рис. 2.8).

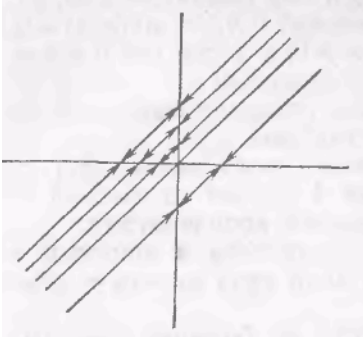


Рис. 2.8.

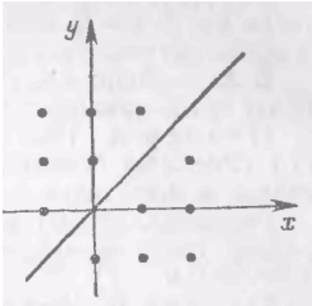


Рис. 2.9.

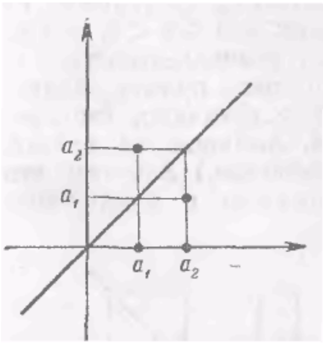


Рис. 2.10.

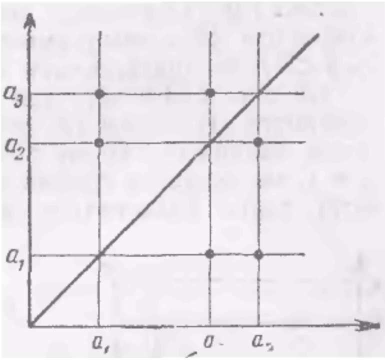


Рис 2.11.

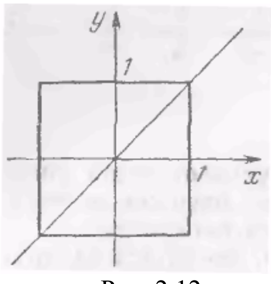


Рис. 2.12.

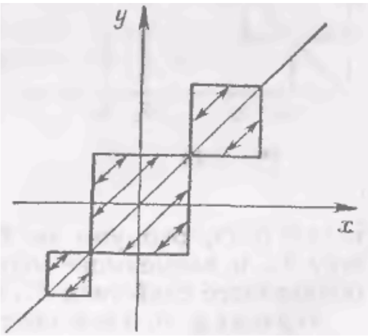


Рис. 2.13.

В дальнейших примерах проверка свойств *R, S, T* будет предоставляться читателям.

Пример 4. (Все целые числа равны друг другу.) (Рис. 2.9.) Множество *A* состоит из главной диагонали и всех точек с целыми координатами.

Очевидно, можно рассматривать и конечные варианты такой эквивалентности типа $a_1 = a_2 = \ldots = a_n$ (рис. 2.10 и 2.11).

Пример 5. (Все числа, не большие единицы по модулю, равны друг другу.) Множество *A* состоит из диагонали и замкнутого единичного квадрата (рис. 2.12). Очевидно, множество, состоящее из открытого (или полузамкнутого: $-1 \leqslant x < 1,\ -1 \leqslant y < 1$) квадрата, также дает эквивалентность.

На рис. 2.13 изображен еще один пример эквивалентности. (Стрелки на рисунке означают, что границы квадратов, кроме точек, лежащих на прямой $y = x$, не входят в график отношения.) Заметим, что если взять аналогичное множество с замкнутыми квадратами, оно уже не будет удовлетворять свойству T_1 и наименьшее множество, содержащее его и обладающее свойством T_1 есть вся плоскость.

Пример 6. (Все числа от a_1 до a_2 равны друг другу и все числа от a_3 до a_4 равны друг другу.) (Рис. 2.14.)

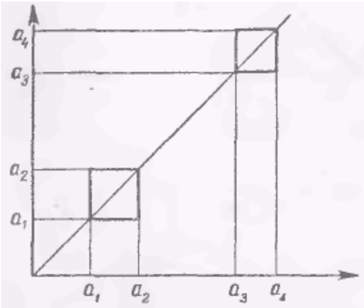


Рис. 2.14.

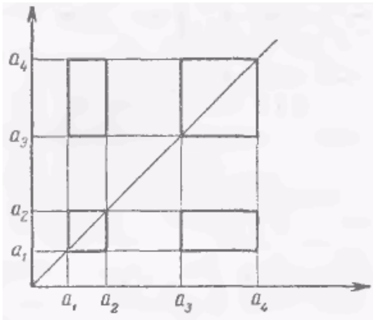


Рис. 2.15.

Пример 7. На рис. 2.15 изображено отношение: «Все числа от a_1 до a_2 и от a_3 до a_4 равны друг другу».

Пример 8. (Ковры Серпинского.) В заключение приведем два примера с множествами *A*, аналогичными «ковру Серпинского». На рис. 2.16 изображено множество *A* для следующего отношения эквивалентности: берется так называемое канторово совершенное множество и отождествляются точки каждого из интервалов, выбрасываемых из отрезка [0, 1] на n-м шаге (на рисунке $n = 3$).

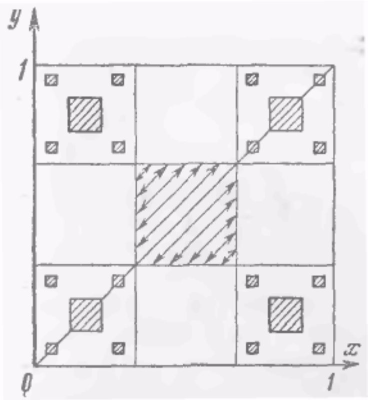


Рис. 2.16.

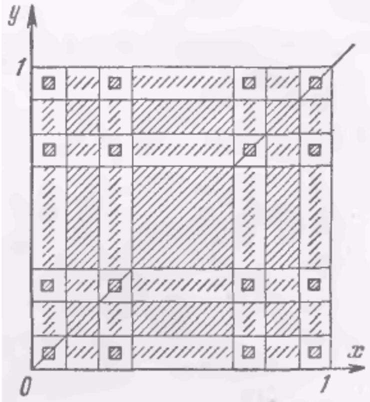


Рис. 2.17.

Если отождествляются все точки дополнения к канторову совершенному множеству, множество *A* имеет вид, изображенный на рис. 2.17.

2.1.5. Разбиение и отношения эквивалентности

Во многих вычислительных задачах берутся большие множества и разбиваются таким образом, чтобы все интересующие нас ситуации можно было исследовать на нескольких правильно выбранных примерах. Например, один с путей получения качественной оценки характеристик языка программирования — это посмотреть конкретные программы, написанные на этом языке. Однако каждый язык, включая язык высокого уровня, порождает бесконечно много программ, и, следовательно, мы должны выбирать программы так, чтобы они правильно отражали достоинства и недостатки языка. Чтобы быть более конкретными, будем в дальнейшем предполагать, что язык имеет три основные управляющие структуры и четыре метода доступа и более у него нет никаких особых свойств. Мы могли бы в качестве примера взять семь программ, каждая из которых включает только одну характеристику языка (хотя, вообще говоря, каждая программа может использовать более чем одну характеристику языка). Исследование этих программ тогда могло бы покрыть большую часть свойств языка. Математически это можно определить следующим образом

Определение. Пусть A — непустое множество и $\{A_i\}$ — совокупность подмножеств (n = 1, ..., n, $n \in N$) таких, что

$$\bigcup_{i=1}^{n} A_i = A.$$

Совокупность этих подмножеств называется *покрытием A*.

Пример 1.

$\{A, B\}$ — покрытие $A \cup B$,

$\{A, A \cup B, B, C)$ — покрытие $A \cup B \cup C$.

Используя понятие покрытия, можно обеспечить, что бы ни одно из свойств не было пропущено, так как каждый элемент включен по крайней мере в одно из подмножеств покрытия. Однако в общем случае могут встречаться случаи дублирования. Если в дальнейшем потребовать, чтобы элементы покрытия попарно не пересекались, то дублирования не будет. Отсюда возникает понятие разбиения.

Определение. *Разбиением* непустого множества A называется совокупность подмножеств $\mathcal{P}(A)$ таких, что объединение всех элементов $\mathcal{P}(A)$ совпадает с A и все элементы $\mathcal{P}(A)$ взаимно не пересекаются, т.е. A разбито таким образом, что каждый элемент A содержится только в одном подмножестве разбиения

Пример 2.

$\{A, A'\}$ — разбиение $\mathcal{E}$,

$\{A \cap B, A \cap B', A' \cap B, A' \cap B'\}$ - разбиение $\mathcal{E}$,

$\{A \backslash B, A \cap B, B \backslash A\}$ - разбиение $A \cup B$.

Разбиение определяется однозначно, и части разбиения индуцируют особый род отношения, называемого отношением *эквивалентности*. Эти отношения ведут себя подобно отношению «=» между числами или множествами. Выделяя основные свойства равенства, мы приходим к следующему определению.

Определение. Бинарным отношением на множестве называют *отношением эквивалентности,* если оно рефлексивно, симметрично и транзитивно.

Отношение эквивалентности - это обобщение понятия равенства. Эквивалентные элементы не различимы для теории в каком-то фиксированном смысле.

Отношение эквивалентности представляет собой экспликацию (перевод интуитивных представлений в ранг строгих математических понятий) таких обыденных слов как «одинаковость», «неотличимость», «взаимозаменяемость».

Примерами отношений эквивалентности есть отношения параллельности на множестве прямых некоторой плоскости; подобия на множестве треугольников; принадлежности к одной функциональной группе объектов или к одному классу типоразмеров и т.д.

Примеры отношений эквивалентности

- Самое наглядное и всем знакомое отношение эквивалентности — разделение контингента учащихся конкретной школы на классы.
- Равенство («$=$»), тривиальное отношение эквивалентности на любом множестве, в частности, вещественных чисел.
- Сравнение по модулю, («a ≡ b (mod n)»).
- В Евклидовой геометрии
    - Отношение конгруэнтности («$\cong$»).
    - Отношение подобия («$\sim$»).
    - Отношение параллельности прямых («$\parallel$»).
- Эквивалентность функций в математическом анализе:

Говорят, что функция $f(x)$ эквивалентна функции $g(x)$ при $x \to x_0$, если она допускает представление вида $f(x) = \alpha(x)g(x)$, где $\alpha(x) \to 1$ при $x \to x_0$. В этом случае пишут $f(x) \sim g(x)$, напоминая при необходимости, что речь идет о сравнении функций при $x \to x_0$. Если $g(x) \neq 0$ при $x \neq x_0$, эквивалентность функций $f(x)$ и $g(x)$ при $x \to x_0$, очевидно, равносильна соотношению

$$\lim_{x \to x0} \frac{f(x)}{g(x)} = 1.$$

- Отношение равномощности множеств.

Более сложный пример, но совершенно жизненно важный:

Когда врач выписывает вам лекарство, он, фактически в рецепте указывает класс эквивалентных лекарств, он не может указать совершенно конкретный экземпляр упаковки таблеток или ампул. То есть всевозможные лекарства разбиты на классы отношением эквивалентности. Если бы не этот факт, современная медицина просто не была бы возможна.

Таким образом, всевозможные рецепты салатов и коктейлей, ГОСТы и классификаторы также определяют жизненно важные отношения эквивалентности. Отношения эквивалентности заполняют всю нашу жизнь, а не являются абстрактной забавой математиков.

Термин *„отношение эквивалентности"* будем применять при выполненные следующих условий (свойств):

1) каждый элемент еквивалентен самому себе;

2) выражение, что два элемента являются эквивалентными, не нуждается в уточнении того, какой из элементов рассматривается первым, а какой вторым;

3) два элемента, которые эквивалентны третьему, эквивалентны между собой.

Эквивалентность, как уже было сказано, удовлетворяет условиям рефлексивности, симметричности, транзитивности и, обычно, обозначается знаком ~. При этом $x \sim y$ означает, что упорядоченная пара (x, y) принадлежит множеству $A \subset M \times M$, которое является отношением эквивалентности в множестве M.

Свойства эквивалентности записываются следующим:

1) $x \sim x$ (рефлексивность);

2) если $x \sim y$, то $y \sim x$ (симметричность);

3) из $x \sim y$ и $y \sim z$ следует $x \sim z$ (транзитивность).

Таким образом, отношение A называется отношением эквивалентности, если оно рефлексивно, симметрично, транзитивно.

Пример.

а) На множестве всех треугольников отношение, которое определяется как {(*x,y): x* и *y* имеют одинаковую площадь}, является отношением эквивалентности.

б) Имеет место следующее отношение, которое определено на множестве всех программ: {(*a,b): a* и *b* вычисляют одну и ту же функцию на определенной машине}. Это отношение является отношением эквивалентности.

Классы эквивалентности. Важнейшее значение эквивалентности заключается в том, что это отношение определяет признак, который допускает *разбиение* множества *M* на непересекающиеся подмножества, называемые *классами эквивалентности.* Наоборот, всякое разбиение множества *M* на непересекающиеся подмножества определяет между элементами этого множества некоторое отношение эквивалентности.

Например, отношение «проживать в одном доме» в множестве жителей города является эквивалентностью и разбивает это множество на непересекающиеся подмножества людей, являющихся соседями по дому.

Дадим формальное определение класса эквивалентности.

Назовем *классом эквивалентности* $K(x_a)$, *элемента* x_a множество всех элементов x_i каждый из которых находится с этим элементом в отношении эквивалентности (множество эквивалентных элементов).

$$K(x_a)=\{x_i/x_i \sim x_a\}.$$

Согласно свойству рефлексивности отношение ~, $x_a \in K(x_a)$. Из свойства транзитивности отношения эквивалентности

$$(x_a \sim x_b)\ \&\ (x_b \sim x_c) \to x_a \sim x_c$$

следует, что $K(x_a) \supset K(x_b)$, а из свойства симметричности —

$$x_a \sim x_b\ -\ K(x_a)=K(x_b).$$

Два различных класса эквивалентности $K(x_n)$ и $K(x_m)$, (x_n не эквивалентен x_m), не пересекаются: $K(x_n) \square K(x_m)=\varnothing$, так как в противном случае они бы совпали. Действительно, пусть найдется элемент x_k, который приналежить этим классам: $x_k \in K(x_n), K(x_m)$, но тогда, согласно приведенным выше свойствам,

$$K(x_n)=K(x_z)=K(x_m).$$

Иначе говоря, если найдется элемент x_z, который принадлежит двум классам эквивалентности $K(x_n)$ и $K(x_m)$, то $K(x_n)=K(x_m)$, что можно записать, используя вместо слова «найдется» обозначение $\exists$ (перевернутая буква E — первая буква английского слова Exist — существует), как

$$(\exists\ x_z \in K(x_n), K(x_m)) \rightarrow K(x_n)=K(x_m).$$

Представление множества X в виде попарно непересекающихся подмножеств $\{X_i\}$ будем называть *разбиением* этого множества:

$$\bigcup_i X_i = X,\ X_{i_a} \cap X_{i_b} = \varnothing,\ i_a \neq i_b.$$

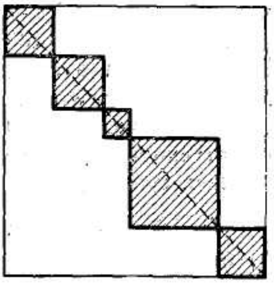


Рис. 2.18.

Таким образом, классы эквивалентности образуют разбиение множества. Тестом распознавания отношения эквивалентности, заданного матрицей смежности, может быть приведение с помощью перестановки строк (столбцов) матрицы смежности к виду, который изображен на рис. 2.18, где около главной диагонали расположены подматрицы, состоящие из единиц (они заштрихованы), другие элементы матрицы равны нулю. Каждая заштрихованная подматрица соответствует классу эквивалентности.

В тех случаях, когда рассматривается только одно отношение эквивалентности, мы можем также использовать обозначение ”=” (эквивалентно), поэтому при определении класа эквивалентности $[x]$ для $x \in X$, запишем:

$$[x] = (y\colon xAy),$$

т.е., $[x]$ есть множество всех элементов X, которые A-эквивалентны x.

Следующий пример иллюстрирует вышесказанное.

Пример. Пусть s фиксированный элемент N; определим отношение A_s на Z:

$A_s = \{(x,y): x\text{-}y=ns$, где $n \in Z\}$.

Рассмотрим случай $s=10$. Тогда

[1]={11, 21, -9, 10976631,...},

[1066]={66, 226, -24,...}

и.т.д.

В действительности есть только десять различных классов эквивалентности. Целые 0,1,2,3,4,5,6,7,8,9 принадлежат различным классам. Поэтому мы можем использовать их в качестве представителей этих классов.

Разбиение множеств

Пусть *A*— произвольное множество. Семейство $(B_i)_{i\in I}$ непустых и попарно не пересекающихся множеств называют **разбиением множества** *A*, если объединение множеств семейства $(B_i)_{i\in I}$ равно *A*, то есть

$$\bigcup_{i\in I} B_i = A\,.$$

Сами множества B_i называют элементами (или членами) разбиения $(B_i)_{i\in I}$.

Например, множества и $\left[\frac{2}{3};1\right]$ образуют разбиение отрезка [0 ;1]. Тривиальными разбиениями *A* являются, по определению, разбиение {*A*}, состоящее только из самого *A*, и разбиение, состоящее из всех одноэлементных подмножеств множества *A*.

Пусть ρ — эквивалентность на множестве *A* и $x \in A$. Множество всех элементов *A*, эквивалентных *x*, т.е. множество $\{y\colon\ y\,\rho\,x\}$, называют классом эквивалентности по отношению ρ и обозначают $[x]_\rho$. Отметим, что в силу рефлексивности для любого элемента $x \in A$ класс эквивалентности не пуст, так как $x \in [x]_\rho$.

Теорема 1. *Для любого отношения эквивалентности на множестве A множество классов эквивалентности образует* *разбиение множества А. Обратно, любое разбиение множества А задает на нем отношение эквивалентности, для которого классы эквивалентности совпадают с элементами разбиения.*

Покажем, что **отношение эквивалентности** ρ на множестве *А* определяет некоторое разбиение этого множества. Убедимся вначале, то любые два класса эквивалентности по отношению ρ либо не пересекаются, либо совпадают.

Пусть два класса эквивалентности $[x]_\rho$ и $[y]_\rho$ имеют общий элемент $z \in [x]_\rho \cap [y]_\rho$. Тогда *zρx* и *zρy*. В силу симметричности отношения ρ имеем *xρz*, и тогда *xρz* и *zρy*. В силу транзитивности отношения ρ получим *xρy*. Пусть $h \in [x]_\rho$, тогда *hρx*. Так как *xρy*, то *hρy* и, следовательно, $h \in [y]_\rho$.

Обратно, если $h \in [y]_\rho$, то в силу симметричности ρ получим *hρy*, *yρx* и в силу транзитивности — *hρx*, то есть $h \in [x]_\rho$. Таким образом, $[x]_\rho = [y]_\rho$.

Итак, любые два не совпадающих класса эквивалентности не пересекаются. Так как для любого $x \in A$ справедливо $x \in [x]_\rho$ (поскольку *xρx*), т.е. каждый элемент множества *А* принадлежит некоторому классу эквивалентности по отношению ρ, то множество всех классов эквивалентности по отношению ρ образует разбиение исходного множества *А*. Таким образом, любое отношение эквивалентности однозначно определяет некоторое разбиение.

Теперь пусть $(B_i)_{i \in I}$ — некоторое разбиение множества *А*. Рассмотрим отношение ρ, такое, что *xρy* имеет место тогда и только тогда, когда *x* и *y* принадлежат одному и тому же элементу B_i данного разбиения:

$$x\rho y \Leftrightarrow (\exists i \in I)(x \in B_i) \wedge (y \in B_i)$$

Очевидно, что введенное отношение рефлексивно и симметрично. Если для любых *x, y* и z имеет место *xρy* и *yρz*, то *x, y* и z в силу определения отношения ρ принадлежат одному и тому же элементу B_i разбиения. Следовательно, *xρz* и отношение ρ транзитивно. Таким образом, ρ — эквивалентность на *A*.

Фактор-множество

Теорема 1 позволяет отождествлять отношения эквивалентности и разбиения: любая эквивалентность определяет единственное разбиение и наоборот.

Множество всех классов эквивалентности по данному отношению эквивалентности ρ на множестве *A* называют **фактор-множеством** множества *A* по отношению ρ и обозначают *A/ρ*.

Пример 1. а. На множестве целых чисел $\mathbb{Z}$ определим отношение равенства по модулю *k*, где $k \in \mathbb{N}$. Положим $x \equiv_{(\mathrm{mod}\,k)} y$, если и только если (*x*-*y*) делится на *A*.

Легко проверяется, что это отношение эквивалентности. Действительно, рефлексивность следует из того, что для любо $m \in \mathbb{Z} m - m = 0$ и делится на *k*; симметричность — из того, что если (*m*-*n*) делится на *k*, то и (*n*-*m*) делится на *k*. Для доказательства транзитивности заметим, что если (*m*-*n*) делится на *k* и (*n*-*p*) делится на *k*, то и их сумма (*m*-*n*)+(*n*-*p*)= (*m*-*p*) делится на *k*. Другими словами, для любых целых *m, n, p* из $m \equiv_{(\mathrm{mod}\,k)} n$ и $n \equiv_{(\mathrm{mod}\,k)} p$следует $m \equiv_{(\mathrm{mod}\,k)} p$, что доказывает транзитивность отношения $\equiv_{(\mathrm{mod}\,k)}$.

Равенство чисел *m* и *n* по модулю *k* означает, что при делении на *k* эти числа дают одинаковые остатки. Действительно, для каждого $x \in \mathbb{Z}$ имеем $x = m \cdot k + r$, где *r*— остаток от деления *x* на *k*. Следовательно, $x - r = m \cdot k$, то есть $x \equiv_{(\mathrm{mod}\,k)} r$. Таким образом, каждое число попадает в тот же класс эквивалентности по отношению $\equiv_{(\mathrm{mod}\,k)}$, что и остаток от деления его на *k*. Поскольку всего различных остатков может быть ровно $k\colon\ 0, 1, \ldots, k-1$, получаем ровно *k* попарно различных классов эквивалентности по данному отношению:

$$[0]=_{(\mathrm{mod}\,k)}, [1]=_{(\mathrm{mod}\,k)}, \ldots [k-1]=_{(\mathrm{mod}\,k)}$$

где класс $[r]=_{(\mathrm{mod}\,k)}$ состоит из всех целых чисел, дающих при делении на k остаток r.

Отметим, что мы установили **взаимно однозначное соответствие** между фактор-множеством $\mathbb{Z}/=_{(\mathrm{mod}\,k)}$ и множеством $\mathbb{Z}_k$, состоящим из чисел $0, 1, \ldots, k-1$.

Второе множество дает нам как бы "наглядный образ" построенного фактор-множества. Нельзя считать, что фактор-множество $\mathbb{Z}/=_{(\mathrm{mod}\,k)}$ равно множеству $\{0, 1, \ldots, k-1\}$. Нет, указанное фактор-множество состоит из k элементов, каждый из которых есть не число, а множество всех целых чисел, при делении на k дающих фиксированный остаток. Но каждому такому классу эквивалентности однозначно сопоставляется целое число от 0 до k-1, и, наоборот, каждому целому числу от 0 до k-1 соответствует единственный класс эквивалентности по отношению $=_{(\mathrm{mod}\,k)}$. Заметим, что в математике часто используется прием сопоставления фактор-множеству такого находящегося с ним во взаимно однозначном соответствии множества, которое легко представить и описать.

б. На множестве $\mathbb{R}$ действительных чисел зададим отношение $a =_{(\mathrm{mod}\,1)} b$, полагая, что числа a и b равны по модулю 1 тогда и только тогда, когда число a-b является целым. Из определения следует, что каждое число по модулю 1 равно своей дробной части.

Примечание. Под дробной частью $< a >$ числа a понимается число из полуинтервала $[0; 1)$, такое, что $a = n+ < a >$ для некоторого целого n. Поэтому дробной частью отрицательного числа $(-a)$, где $a > 0$, будет число $(1- < a >)$. Так, дробной частью $(-1{,}23)$ будет не 0,23, а 0,77=1-0,23.

Так как отношение $=_{(\mathrm{mod}\,1)}$ определено через равенство, то легко понять, что все свойства отношения эквивалентности для него выполняются. Каждый класс эквивалентности будет содержать числа с равными дробными частями. Это значит, что каждый класс эквивалентности по данному отношению однозначно определяет некоторое число из полуинтервала $[0;1)$ и, наоборот, каждому числу $\gamma \in [0;1)$ однозначно сопоставляется класс эквивалентности, состоящий из всех действительных чисел, дробная часть которых равна γ. Таким образом, фактор-множество $\mathbb{R}/=_{(\mathrm{mod}\,1)}$ и полуинтервал $[0;1)$ на числовой прямой находятся во взаимно однозначном соответствии. Этот полуинтервал можно рассматривать как представление определенного выше фактор-множества.

Примеры

Расписание занятий в школе — это типичный пример факторизации. В данном случае X — множество всех учащихся школы, Y — множество всех учебных предметов, разнесенных по дням недели с уточнением времени проведения занятий. Классами эквивалентности являются классы (группы учащихся). Отображение f — расписание занятий, отображаемое в дневниках учащихся. Отображение f — расписание занятий по классам, вывешиваемое в вестибюле школы. Здесь же имеется и отображение p — списки классов. Этот пример очень наглядно демонстрирует практические выгоды факторизации: невозможно представить себе расписание занятий, как таблицу, в которой отражены все ученики школы в персональном порядке. Факторизация позволила отобразить нужную учащимся информацию в удобном для применения компактном виде в ситуации, где формулы применить не удается.

Однако этим выгоды факторизации не ограничены. Факторизация позволила провести разделение труда между участниками деятельности: завуч составляет расписание, а учащиеся записывают его себе в дневники. Аналогичным образом, факторизация выписки рецептов позволила провести разделение труда между медиком, ставящим диагноз и выписывающим рецепт, и аптекарем, обеспечивающим эквивалентность выписанных лекарств. Апофеозом факторизации является конвейер, реализующий максимальное разделение труда за счет стандартизации деталей.

Но и этим выгоды факторизации не ограничены. Факторизация позволила обеспечить модульность современной техники, что дает ей небывалую гибкость функций. Вы можете сохранить старую сим-карту и купить к ней совершенно новый телефон, или в свой старый компьютер вставить новую видеопамять. Все это — гибкость, модульность, в основе которой лежит факторизация.

Связь между понятиями эквивалентности и отображения

Установим теперь связь между понятиями эквивалентности и отображения. Заметим, что для любого отношения эквивалентности ρ на множестве *A* можно определить отображение $f_\rho\colon A \to A/\rho$, положив $f_\rho(x) = [x]_\rho$, т.е. сопоставив каждому $x \in A$ содержащий его класс эквивалентности. Это отображение сюръективно, так как каждый элемент множества *A* принадлежит некоторому классу эквивалентности, т.е. для каждого $[x]_\rho \in A/\rho$ справедливо $[x]_\rho = f_\rho(x)$.

Отображение f_ρ определенное таким образом, называют **канонической сюръекцией** множества *A*.

Покажем, что любое отображение однозначно определяет некоторое отношение эквивалентности.

Теорема 2. *Пусть f:A→B — произвольное отображение. Отношение ρf на множестве A, для которого* $x\,\rho_f\,y$*, если и только если* $f(x) = f(y)$*, является отношением эквивалентности, причем существует биекция фактор-множества* A/ρ_f *на множество f(A).*

Доказательство. Рефлексивность, симметричность и транзитивность отношения *ρf* следуют непосредственно из его определения, т.е. *ρf* — эквивалентность.

Зададим отображение φ фактор-множества *A/ρf* в множество *f(A)* следующим образом: $\varphi([x]_{\rho_f}) = f(x)$. Из способа задания отношения ρ следует, что отображение определено корректно, т.е.

каждому классу эквивалентности поставлен в соответствие единственный элемент $y \in f(A)$.

Докажем, что φ — биекция, для чего убедимся в том, что это инъекция и сюръекция одновременно. Пусть классы эквивалентности $[x]_{\rho_f}$ и $[y]_{\rho_f}$ не совпадают. В силу теоремы 1 это означает, что они не пересекаются, т.е. *x* не эквивалентно *y*. Из определения отношения *ρf* следует, что $f(x) \neq f(y)$. Таким образом, φ — инъекция. Если элемент $u \in f(A)$, то найдется такой элемент $x \in A$, что $u = f(x) = \varphi([x]_{\rho_f})$, то есть φ — сюръекция фактор-множества *A*/*ρf* на множество *f(A)*. Итак, φ — биекция.

Следовательно, в силу доказанных теорем 1 и 2 **существует связь между <u>тремя понятиями</u> — отображением множества, отношением эквивалентности на множестве и разбиением множества**. Но неверно, что существует взаимно однозначное соответствие между отображениями и отношениями эквивалентности (заметим, что теорема 2 этого и не утверждает). Два разных отображения могут определять одно и то же разбиение отображаемого множества, тем самым задавая на нем одно и то же отношение эквивалентности. Так, например, любое биективное отображение *f:A→B* задает на *A* одно и то же разбиение — тривиальное разбиение на одноэлементные множества.

Система представителей. Все элементы, которые принадлежат некоторому классу M_i разбиения $\{M_1, M_2, \ldots\}$ множества *M,* связаны отношениям эквивалентности. Они взаимозаменяемы в том смысле, что каждый из этих элементов определяет данный класс, т.е. может служить его *представителем (эталоном).* Подмножество из *M,* которое содержит по одному и только по одному элементу из каждого класса некоторого разбиения, называют *системой представителей* соответствующего отношения эквивалентности. Множество всех классов разбиения множества *M,* определенного отношением эквивалентности *A,* образовывает фактор-множество *M/A.*

Например, отношение параллельности определяет разбиение множества прямых на плоскости на классы, каждый из которых образован множеством параллельных между собой прямых и характеризуется некоторым направлением (следует также считать, что прямая параллельна сама себе). Любая из параллельных прямых может служить представителем данного класса, а само направление есть класс эквивалентности. Множество всех направлений составляет фактор-множество множества всех прямых по отношению параллельности.

Классы вычетов по модулю *m*. Рассмотрим отношение сравнения по модулю m на множестве натуральных чисел, что записывается как x=y (mod m) и означает: x сравнимо с y по модулю m (m — целое положительное число, не равное нулю), если x–y делится на m. Целые числа, которые сравнимы по модулю m, связаны соотношением x=y+km (k — целое число) и образовывают подмножество целых чисел, которые имеют одинаковый остаток j при делении на m. Так как эти подмножества не пересекаются, они являются классами эквивалентности, а в качестве представителя каждого из них естественно выбрать остаток j = 0, 1, 2, ..., m—1. Таким образом, *отношение сравнения по модулю* m определяет разбиение множества натуральных чисел на m классов $M_0, M_1, M_2, ..., M_{m—1}$, где $M_j=\{j, j+m, j+2m, ...\}$ — счетное множество, которое называют *классом вычетов по модулю* m.

Например, при m=4 имеем

M_0 ={0, 4, 8, 12, ...}; M_1 = {1, 5, 9, 13, ...};

M_2 = {2, 6, 10, 14, ...}; M_3 = {3, 7, 11, 15, ...}.

Представителями классов эквивалентности являются числа 0, 1, 2 и 3, так как

0 = 4(mod4) = 8(mod4) = ...; 1 = 5(mod4) = 9(mod4) = ...;

2 = 6(mod4) = 10 (mod4) =... и 3 = 7(mod4) = 11 (mod4) =....

Таким образом, множество целых чисел разбивается отношением сравнения по модулю 4 на четыре класса эквивалентности. Внутри каждого класса эти числа неразличимы (4 ~ 0, 5 ~ 1, 6 ~ 2, 7 ~ 3 и т.д.).

При m = 1 разбение состоит из единственного класса, который совпадает с исходным множеством, т.е. имеем *полное отношение эквивалентности,* при котором любые два элемента эквивалентны (все целые числа делятся на единицу). Отношение x = y(mod 2) разбивает множество целых чисел на классы парных и непарных чисел.

Значение рассмотренного примера настолько большое, что многие авторы принимают для отношения эквивалентности A обозначение a = b (mod A).

Идентификация элементов. Произвольное отношение эквивалентности определяет на некотором множестве обобщенную форму равенства. Классы эквивалентности состоят из всех тех элементов, которые неразличимы с точки зрения данного отношения эквивалентности. Разбиение множества на классы означает *идентификацию* эквивалентных между собой элементов. При этом каждый класс определяется его представителем (эталоном) и отождествляется с некоторым общим свойством или совокупностью свойств (параметров) входящих у него элементов (направление по отношению параллельности, остаток относительно сравнения по модулю m и т. д.).

Предельным случаем отношения эквивалентности есть *тождественное равенство.* Единствнный элемент, равный какому-либо данному элементу, есть этот самый элемент. Следовательно, имеем самое полное разбиение, при которой классы эквивалентности содержат только по одному элементу исходного множества.

Классы номинальных значений. В условиях массового производства стандартной детали (или компонента) определяется *ряд номинальных значений* характеризующей ее величины (емкость конденсатора, диаметр и твердость шарикоподшипника, чистота поверхности детали, содержание примесей вещества и т.п.). Например, для постоянных резисторов с допустимым отклонением сопротивления +20% задается ряд: 1,0; 1,5; 2,2; 3,3; 4,7; 6,8. Величины номинальных сопротивлений должны отвечать числам, которые получаются умножением этих значений на 10^k, где k — целое положительное или отрицательное число (1,5 ком; 4,7 Ом; 0,68 Ом; 1,0 МОм и т.п.). Этими указаниями руководствуются при проектировании изделий с резисторами и при их сортировке в процессе производства.

Заданне ряда номинальных значений $x_1, x_2, \ldots, x_n$ с допустимым отклонением Δx можно рассматривать как определение отношения эквивалентности в множестве M значений параметров x некоторых объектов. Неравенства $(x_i — \Delta x) \leq x < (x_i + \Delta x)$, $i = 1, 2.....n$ определяют n классов эквивалентности в множестве возможных значений x. Семейство представителей образуется совокупностью всех номинальных значений $\{x_1, x_2, \ldots, x_n\}$. Объекты, параметры которых принимают эти значения, взаимозаменяемы (возможность замены любого объекта эквивалентным) и неразличимы (все эквивалентные объекты маркируются номинальным значением).

Матрица отношения эквивалентности. Элементы, которые принадлежат некоторому классу эквивалентности, попарно эквивалентны между собой, а их сечения совпадают. Следовательно, столбцы матрицы отношения эквивалентности для элементов одного класса одинаковы и содержат единицы во всех строках, которые соответствуют этим элементам. Так как классы эквивалентности не пересекаются, то в столбцах разных классов не будет единиц в одинаковых строках.

Расположим элементы множества так, чтобы в каждом классе эквивалентности принадлежащие ему элементы стояли рядом. Тогда единичные элементы матрицы отношения эквивалентности образуют непересекающиеся квадраты, диагонали которых располагаются по главной диагонали матрицы. Например, для разбиения на классы эквивалентности

$$M_1 = \{x_1, x_2, x_3\};\ M_2 = \{x_4\};\ M_3 = \{x_5, x_6, x_7, x_8\}$$

имеем:

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	1	1	1					
x_2	1	1	1					
x_3	1	1	1					
x_4				1				
x_5					1	1	1	1
x_6					1	1	1	1
x_7					1	1	1	1
x_8					1	1	1	1

Граф отношения эквивалентности. На рис. 2.19 изображен граф отношения эквивалентности, матрица которого приведена выше.

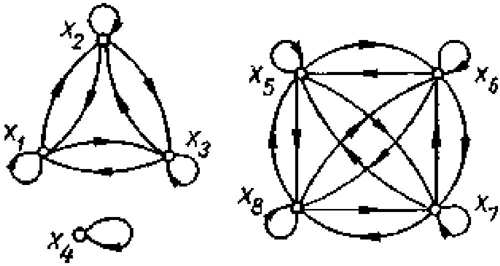


Рис. 2.19. Граф отношения эквивалентности на множестве X.

Каждому классу эквивалентности соответствует отдельная часть графа, которая представляет собой *полный направленный граф* на множестве ее вершин. Как видно, матрицы и графы отношения эквивалентности содержат избыточную информацию и полностью определяются заданием классов эквивалентности.

Разбиение и отображение. Разбиение множества на классы можно связать с отображением $f : X \to Y$, ставящим каждому элементу из X в соответствие один и только один элемент из Y (рис. 2.20). Собирая в один класс все те элементы из X, образы которых в Y совпадают, приходим к некоторому разбиению на непересекающиеся подмножества $\{X_1, X_2, ...\}$.

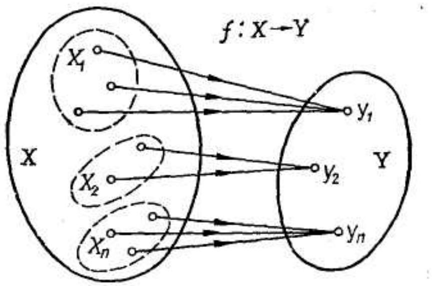


Рис. 2.20. Отображение $f : X \to Y$, порождающее отношение эквивалентности на X.

Каждое подмножество X_i характеризуется соответствующим ему образом $y_i \in Y$ и является классом эквивалентности. Обратно, если задана некоторая совокупность классов эквивалентности $\{X_1, X_2, ...\}$ множества X, то каждому элементу $x \in X$ можно поставить в соответствие тот класс X_i к которому принадлежит x. В результате получаем отображение множества X на множество классов $\{X_1, X_2, ...\}$.

Пусть, например, заданы отображения

$$\{(x_1 ,y_1), (x_2 ,y_3), (x_3 ,y_1), (x_4 ,y_2), (x_5 ,y_1), (x_6 ,y_3)\}.$$

Тогда классы эквивалентности, соответствующие образам y_1, y_2 и y_3, будут:

$$X_1=\{ x_1, x_3, x_5 \}; X_2=\{x_4\}; \; X_3 =\{x_2, x_6\}.$$

Отображение X на $\{X_1, X_2, X_3\}$ выразится множеством упорядоченных пар

$$\{(x_1, X_1), (x_2, X_3), (x_3, X_1), (x_4, X_2), (x_5, X_1), (x_6, X_3),\}.$$

Итак, любое отображение $f : X \to Y$ порождает отношение эквивалентности на множестве X, причем $x_i \sim x_j$, если и только если $f(x_i)= f(x_j)$. Образа y_i классов эквивалентности X_1, X_2, ... могут служить эталонами и образовывают в совокупности систему представителей.

Измерения. Измерительный прибор можно рассматривать как устройство, которое отображает множества возможных значений измеренных величин x в множество элементов *функциональной шкалы* прибора. При использовании цифровых измерительных приборов *результат измерения* выходит в виде некоторого n-разрядного числа $\alpha \in Y$, которое отвечает измеренной величине x, заключенной в интервале $(\alpha_i-0,5) \leq x < (\alpha_i+0,5)$. Множество возможных значений x разбивается на 10^n классов эквивалентности, каждый из которых характеризуется соответствующим ему образом α_i, из множества чисел $\{0,1,2,..., 10^n — 1\}$.

В аналоговых приборах функциональная шкала значений α_i обычно наносится в виде меток на отрезок дуги или прямой, а результат измерения α_i определяется положением подвижного указателя относительно шкалы (иногда таким указателем является сам объект измерения, например, при измерении длины с помощью линейки или рулетки). Множество классов эквивалентности определяется соотношениями $(\alpha_i-\Delta\alpha) \leq x < (\alpha_i+\Delta\alpha)$, где $\Delta\alpha$ равно половине расстояния между соседними метками шкалы (предполагается, что шкала равномерная).

При различных измерениях, а также для градуировки приборов используют также *натуральные шкалы.* Например, шкала твердости минералов задается системой неравенства

$$x<\beta_0;\ \beta_0\leq x<\beta_1;\ \ldots;x<\beta_9\leq x<\beta_{10};\ \beta_{10}\leq x,$$

соответствующих 0, 1, 2, ..., 9, 10 баллам, причем β_i — твердости некоторых минералов (тальк, гипс, известковый шпат, ..., корунд, алмаз). Каждое неровненство определяет класс эквивалентности для твердости, а представителями этих классов является твердость β_i выраженная в баллах (β_i= 0, 1, ..., 10). Аналогично строятся натуральные шкалы температур, двенадцатибальная шкала скорости ветра, шкала чувствительности фотопленок и т.п.

Следует отметить, что изложенное выше является идеализированной моделью, не учитывающей различных погрешностей, которые неизбежно сопровождают процесс измерения в реальных условиях.

3. Отношение порядка

3.1. Основные положения

Упорядоченность. *Отношение порядка* имеет свойство рефлексивности, транзитивности и антисимметричности. Его принято обозначать символом $\leq$. Запись $x \leq y$ означает, что пара (x, y) принадлежит множеству $A \subset M \times M$, являющемуся отношением порядка в множестве M, причем x *предшествует* y (или y следует за x). В принятых обозначениях свойства отношения порядка запишутся следующим образом:

1) $x \leq x$ (рефлексивность);

2) если $x \leq y$ и $y \leq z$, то $x \leq z$ (транзитивность);

3) из $x \leq y$ и $y \leq x$ следует $x = y$ (антисимметричность).

Множество, на котором определено отношение порядка, называют *упорядоченным,* и говорят, что *порядок введен* этим отношением. Множество *совершенно (линейно, просто), упорядочено,* если для любых двух его элементов имеет место, по крайней мере, $x \leq y$ или $y \leq x$ (его называют также *цепью).* Например, множество натуральных или действительных чисел с естественным отношением порядка $\leq$, множество значений длин волн на шкале радиоприемника и т г.

В общем случае может оказаться, что для некоторых пар (x, y) ни одно из соотношений $x \leq y$ и $y \leq x$ не имеет места (такие элементы называют *несравнимыми).* Тогда говорят, что множество *частично упорядочено.* Типичными примерами частичного порядка являются включение, отношение «быть делителем» и т.п. Так, отношение включения на множестве подмножеств некоторого универсуума рефлексивно ($X \subset X$), транзитивно (если $X \subset Y$ и $B \subset Z$, то $X \subset Z$) и антисимметрично (из $X \subset Y$ и $Y \subset X$ следует $X=Y$), но среди всевозможных подмножеств есть такие, что ни одно из соотношений $X \subset Y$ и $Y \subset X$ для них не имеет места. Аналогично не все пары элементов из множества целых чисел находятся в отношении «быть делителем».

Рассмотрим пример частично упорядоченного множества, графическое изображение которого приведено на рис. 3.1 (в качестве отношения $\leq$ рассмотрено отношение включения множеств $\subset$). Часто частично упорядоченное множество изображают в виде графов $H=<V, \leq>$, в которых изъятые все петли и все транзитивно замыкающие дуги. Граф $H=<V, \leq>$, задающий частично упорядоченное множество с изъятыми петлями и транзитивно замыкающими дугами, называется *диаграммой Хассе* H. Диаграмма Хассе H, задающая частично упорядоченное множество, которое показано на рис. 3.1, изображена на рис. 3.2,*а*.

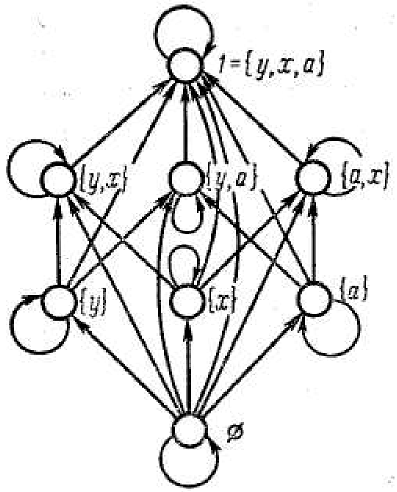


Рис. 3.1.

Диаграмма Хассе известна с конца XIX в., и в течение многих лет ее применяли в генеалогии для определения родства. Понятие непосредственного старшего легко задается в частично упорядоченном множестве следующим определением: x_i *покрывает* x_j — это означает, что $x_j < x_i$, и не найдется такого элемента x_p, что $x_j < x_p < x_i$.

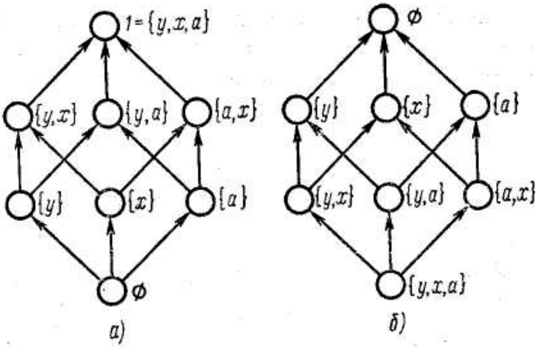


Рис. 3.2.

Рассмотрим подмножество X' упорядоченного множества X. Если найдется элемент $x_\alpha \in X$ такой, что $x_i \leq x_\alpha$ для любого элемента x_i подмножества X', то элемент x_α называется *мажорантой* подмножества X'. Аналогично, если обнаружится элемент $x_\beta \in X'$ такой, что $x_\beta \leq x_i$ для любого элемента x_i подмножества X', то элемент x_β называется *минорантой* подмножества X'. На диаграмме Хассе Н (рис. 3.2, *а*) элемент $\{x\}$ является минорантой подмножества $\{\{y, x\}, \{a, x\}, \{y, x, a\}\}$, элемент $\{a, x\}$ — мажорантой подмножества $\{\{x\}, \{a\}, \varnothing\}$.

Если мажоранта x_α подмножества X' принадлежит X', $x_\alpha \in X'$, то она называется *максимальным элементом* $x_{макс}$ этого подмножества. Аналогично, если миноранта x_β подмножества X' принадлежит X', $x_\beta \in X'$, то она называется *минимальным элементом* $x_{мін}$ подмножества X'. В диаграмме Хассе *Н* (рис. 3.2, *а*) минимальным и максимальным элементами подмножества $\{\{x\}, \{y, x\}, \{a, x\}, \{y, x, a\}\}$ являются соответственно $\{x\}$, $\{y, x, a\}$. Для пары элементов линейно упорядоченного множества всегда существуют максимальный (равный одному з них) и минимальный (равный другому) элементы. Часто пары элементов линейно упорядоченного множества называют *сравнимыми,* т.е. элементами x_i, x_j, для которых $x_i \leq x_j$ или $x_j \leq x_i$.

Если множество мажорант (минорант), в свою очередь, имеет максимальный (минимальный) элемент, то его называют *верхней (нижней) гранью подмножества* X'. Верхнюю (нижнюю) грань множества X' обозначают $\sup X'$ ($\inf X'$).

Верхняя (нижняя) грань подмножества X', принадлежащая X', называется *наибольшим (наименьшим) элементом подмножества* X'.

Теорема 1. *Упорядоченное множество X содержит не более одного наибольшего (наименьшего) элемента.*

Докажем теорему для случая наибольшего элемента. Действительно, если x_α, x_β — два наибольших элемента, то $x_\alpha \leq x_\beta$ и $x_\beta \leq x_\alpha$, откуда в силу антисимметричности $\leq$ $x_\alpha = x_\beta$. Доказательство для наименьшего элемента аналогично.

Наибольший элемент, если он существует, упорядоченного множества X называют *единичным* и обозначают 1. Наименьший элемент, если он существует, упорядоченного множества X называют *нулевым* и обозначают 0.

В упорядоченном семействе множества X пустое множество соответствует нулевому элементу, X — единичному элементу.

Элемент, покрывающий 0 в частично упорядоченном множестве X, т.е. минимальный элемент в подмножестве множества X, полученном исключением 0, называется *атомом* или *точкой.* При заданнии с помощью графа семейства множества X точке (атому) соответствует элемент универсума.

Под *изоморфизмом* между двумя упорядоченными множествами X и X^* будем понимать взаимно однозначное соответствие η между X и X^* такое, что из $x_i \leq x_j$ следует $\eta\ (x_i) \leq \eta\ (x_j)$ и из $\eta\ (x_i) \leq \eta\ (x_j)$ следует $x_i \leq x_j$.

Два упорядоченных множества X и X^* называются *изоморфными* тогда и только тогда, когда между ними существует изоморфизм.

Под отношением $\overline{R}$, обратным R, понимают такое отношение, при котором $(x_i, x_j) \in \overline{R}$ тогда и только тогда, когда $(x_j, x_i) \in R$.

Принцип двойственности. *Отношение, обратное отношению упорядоченности, также является отношением упорядоченности.*

Двойственным к частично упорядоченному множеству X называется частично упорядоченное множество X, которое определено на том же носителе с помощью обратного отношения. Часто принцип двойственности формулируют следующим образом: *если теорема справедлива для частично упорядоченных множеств, то справедлива и двойственная ей теорема.*

Очевидно, что подмножество X' упорядоченного множества X является упорядоченным множеством, и если это упорядоченное множество является линейным, то подмножество X' является *цепью* X' в X. Одной из важных числовых характеристик цепи X' является ее *длина* l, равная $|X'|$-1, где $|X'|$ — мощность носителя линейно упорядоченного подмножества X'. Каждая цепь длины l изоморфна цепи целых чисел 1, 2,...., l+1.

Высотой $d(x_i)$ элемента x_i упорядоченного множества X называется максимум длины $l_{\text{макс}}$ цепей $x_0<x_1<...<x_i$ в X, для которых x_i — наибольший элемент (x_0 — минимальный элемент множества X).

Длиной $d(X)$ упорядоченного множества X называется максимум длин цепей в X. Другими словами, длиной $d(X)$ упорядоченного множества называется максимум высот ($d_i(x_i)$ его элементов

$$d(X) = \max_i d_i(x_i), \quad x_i \in X.$$

Наименьшей верхней гранью называется верхняя грань, меньшая всякой другой верхней грани. *Наибольшая нижняя грань* определяется аналогично. Очевидно, что подмножество упорядоченного множества имеет не больше одной наименьшей верхней и одной наибольшей нижней грани.

Наконец, отметим, что использование естественного порядка на R определяет новые множества. Их называют *интервалами:*

$$[a, b] = \{x: x \in R, \ a \leq x \leq b\}$$

есть *замкнутый интервал (отрезок)* от a до b;

$$]a, b[= \{x: x \in R, \ a < x < b\}$$

есть *открытый интервал* от a до b.

В каждом случае числа a и b называются концевыми точками.

Замкнутый интервал включает в себя концевые точки, а открытый нет. Удобно также определить полуоткрытые интервалы:

$$[a, b[= \{x: x \in R, \ a \leq x < b\},$$
$$]a, b] = \{x: x \in R, \ a < x \leq b\}.$$

Для удобства будем использовать следующие обозначения:

$$]-\infty, a] = \{x: x \leq a\},$$
$$]-\infty, a[= \{x: x < a\},$$
$$[a, \infty[= \{x: x \leq a\},$$
$$]a, \infty[= \{x: x < a\},$$
$$]-\infty, \infty[= R.$$

Отношение строгого порядка. Отношения, которое наделено свойствами транзитивности и антирефлексивности (следствиями этих двух свойств есть также ассиметричность и антисимметричность), называют *отношением строгого порядка* и обозначают символом <. Свойство антирефлексивности означает, что элемент множества не может сравниваться сам с собой (как в случае строгого неравенства или строгого включения). Между отношениями строгого и нестрогого порядка имеют место соотношения: (≤) = (<) □ E и (<) = (≤) \ E, где E — тождественное отношение.

Отношение строгого порядка характерно для различного рода *иерархий* с подчинением одного объекта другому (или другим). Если для некоторой совокупности элементов из X справедливо соотношение $x_1 < x_2 < \ldots < x_n$, то в соответствии со свойством транзитивности $x_i < x_j$ $(i<j\leq n)$, т.е. отношение строгого порядка обусловливает как прямое, так и косвенное подчинение по старшинству. Говорят, что x_{i+j} покрывает x_i, если $x_i < x_j$, и не существует такого промежуточного элемента x, что $x_i < x < x_j$.

Последовательности. Элементы любого конечного множества M можно пронумеровать порядковыми числами 1, 2, 3, ..., n. Для счетного множества нумерацию следует понимать как взаимно-однозначное отображение множества натуральных чисел N на M, где каждому числу i ставит в соответствие некоторый элемент x_i из M. Упорядоченное таким отображением множество $\{x_1, x_2, x_3, \ldots\}$ называется *последовательностью* (конечной или бесконечной). Элемент x_i из M называют *членом последовательности* с индексом i.

Если отношение строгого порядка на конечном множестве совершенно, то на этом множестве всегда можно выбрать такую последовательность всех его элементов $\{x_1, x_2, \ldots, x_n\}$, что соотношение $x_i<x_j$ будет выполняться в том и только в том случае, когда $i < j \leq n$. Другими словами, любой совершенно строгий порядок на конечном множестве равносилен естественному порядку следования натуральных чисел. Если же порядок на конечном множестве не является совершенным, то элементы этого множества нельзя пронумеровать так, чтобы большим номерам соответствовали старшие элементы.

Нумерация элементов множества устанавливает совершенно строгий порядок на этом множестве. Например, на спортивных соревнованиях жеребьевкой каждому спортсмену ставится в соответствие номер; термины в предметном указателе располагаются в соответствии с порядком следования букв алфавита (здесь предполагается соответствие между последовательностью букв и отрезком натурального ряда чисел).

Весовые функции. Пусть на множестве M определено отображения $f : M \to R$ (R — множество действительных чисел), ставящее в соответствие каждому объекту x из M некоторое действительное число $f(x)$. Это число называют *весом,* а отображение f

— *весовой функцией.* Иногда понятие веса совпадает с буквальным смыслом этого слова (вес детали какого-нибудь механизма, атомный вес химического элемента, полезный груз автомашины в колонне и т.п.). Но весом может служить любая числовая характеристика объекта (сопротивление резистора, объем тела, площадь участка, число баллов спортсмена и т.п.).

Если отображение f взаимно-однозначное, то на множестве M можно определить совершенно строгий порядок условием: $x<y$, если $f(x)<f(y)$. Действительно, поскольку не существует объектов с равными весовыми функциями, то для любой пары (x, y) справедливо или $f(x)<f(y)$, или $f(y)<f(x)$, т.е. все элементы сравнимы, и отношение антирефлексивно. В то же время оно транзитивно, так как для элементов $x, y, z \in M$ из $f(x)<f(y)$ и $f(y)<f(z)$ следует $f(x)<f(z)$.

Примерами совершенно строгого упорядочения множества, на котором определено инъективное отображение (весовая функция) являются: периодическая система Менделеева, расположение спортсменов по совокупности полученных баллов при условии, что нет одинаковых результатов и т.п.

Квазипорядок. Если отображение $f: M \to R$ не инъективно, т.е. два разных объекта x и y из M могут иметь равные веса $f(x)=f(y)$, то отношение между ними не является антисимметричным и, следовательно, не удовлетворяет определению порядка. В то же время, как показано выше, с отображением f можно связать разбиение множества M на классы эквивалентности $\{M_1, M_2, ..., M_j, ...\}$. Каждый из них объединяет различные элементы из M с равными весами, причем этот вес служит представителем соответствующего класса.

Теперь можно говорить об упорядочении совокупности классов эквивалентности $\{M_1, M_2, ... \}$ некоторого множества M по их представителям $\alpha_1, \alpha_2, ...$ Так как система представителей $\{\alpha_1, \alpha_2, ...\}$ не содержит одинаковых элементов (в противном случае соответствующие им классы объединились бы в общий класс эквивалентности), то на этой системе как на множестве можно определить строгий порядок. Такое упорядочение отождествляет элементы множества M, принадлежажещее к одному и тому же классу эквивалентности, и определяет на этом множестве *квазипорядок (предпорядок).* Говорят также, что строгий порядок на множестве классов эквивалентности $\{M_1, M_2, ...\}$ множества M *индуцируется* квазипорядком.

Квазипорядок удовлетворяет условиям рефлексивности и транзитивности. Он является обобщением эквивалентности (в определение не входит свойство симметричности) и нестрогого порядка (не обязательно свойство антисимметричности). Отношение, являющееся одновременно эквивалентностью и нестрогим порядком, есть *тождественное равенство.* Можно также показать, что если A — квазипорядок, то $A \cap A^{-1}$ — эквивалентность. Совершенный квазипорядок индуцирует и совершенно строгий порядок на множестве классов эквивалентности.

Рассмотренный раньше ряд номинальных значений можно рассматривать как строго упорядоченное множество, а упорядоченное множество объектов, приведенных в соответствие этому ряду, вводится квазипорядком. Аналогично можно говорить, что квазипорядок на множестве возможных значений измеряемых величин индуцирует строгий порядок на функциональной шкале (множества реперных точек) измерительного прибора.

Области уровня. Классы эквивалентности множества M с квазипорядком, представляющие собой такие множества, где весовая функция f принимает фиксированные значения, обычно называются *областями уровня.* Например, в множестве конденсаторов, которые характеризуются номинальными значениями емкости, области уровня — это подмножества конденсаторов с одинаковыми номинальными емкостями. Другим примером служит квазипорядок A на множестве комплексных чисел $z=a+bi$ такой, что z_iAz_k, если $a_i \leq a_k$. При этом различные комплексные числа с одинаковыми действительными частями объединяются в классы эквивалентности, множество которых может быть упорядочено по их представителям.

Пусть M — множество точек на топографической карте и $h(a)$ высота точки a над уровнем моря. Отношение $a \leq b$, если $h(a) \leq h(b)$, определяет квазипорядок на множестве M. Областями уровня служат *горизонтали (изолинии)* — геометрические места (различных) точек, высота которых над уровнем моря одинакова. Обычно такие горизонтали строят для некоторых фиксированных уровней h_1, h_2, ..., h_n и по ним судят о характере изображаемого рельефа (рис. 3.3).

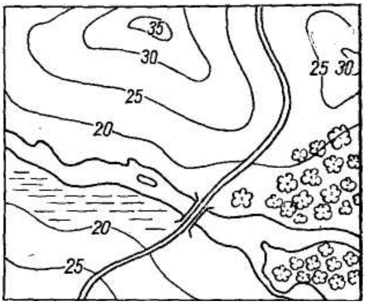


Рис. 3.3. Линии уровня (горизонтали) на топографической карте.

Аналогичный прием используется для представления на плоскости функций двух переменных $\varphi(x, y)=z$. Полагая $z = h_1, h_2,..,h_n,$ строят кривые (изолинии), соответствующие $\varphi_i(x, y) = h_i$ $(i = 1, 2, ..., n)$.

На практике в такой форме обычно представляют рассчитанные или полученные экспериментально значения различных величин, которые характеризуют точки плоскости (или земной поверхности), температуры (изотермы), потенциала (эквипотенциальные линии) и т.п.

Комплексный показатель качества. Сравнение различных изделий (или любых объектов) по некоторой числовой характеристике сводится, как об этом уже говорилось, к упорядочению множества соответствующих им весов, которые можно рассматривать как некоторый показатель качества. Сложное изделие характеризуется несколькими показателями качества $x_1, x_2, ..., x_n$ (стоимость, надежность, габаритные размеры, масса и т.п.).

Для оценки различных типов изделий одинакового назначения используется *комплексный показатель качества,* который выражается некоторым числом $x=f(x_1, x_2, ..., x_n)$. Простейший способ определения этого числа основан на соотношении $x = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n$, где α_i — *коэффициент весомости* показателя x_i. Обычно под $x_1, x_2, ..., x_n$ понимают *относительные показатели* в сравнении с соответствующими показателями некоторого изделия, принятого в качестве базисного. Коэффициенты весомости α_i являются численными выражениями значимости показателей и их определение находится в компетенции специалистов конкретной отрасли промышленности.

Определив комплексные показатели качества некоторой совокупности изделий одинакового назначения и упорядочив множество этих показателей, можно судить о качестве изделий и сравнивать их между собой. Изделия с одинаковыми показателями качества являются в этом отношении эквивалентными. Необходимо, однако, отметить, что порядок или квазипорядок на множестве изделий зависит от того, как определены коэффициенты весомости α_i, что составляет основные трудности при оценке качества.

Матрицы отношений порядка. Отношению порядка соответствует матрица, у которой главная диагональ заполнена единицами (рефлексивность). Для каждой пары единичных элементов, один из которых расположен в i-м столбце и j-й строке, а второй — в j-м столбце и k-й строке, обязательно существует единичный элемент в i-м столбце и k-й строке (транзитивность). Кроме того, ни один единичный элемент не имеет симметричного относительно главной диагонали (антисимметричность). Например, матрица отношения «быть делителем» на множестве $\{1, 2, 3, 4, 6, 7, 12, 14, 21, 28, 42, 84\}$ имеет вид:

	1	2	3	4	6	7	12	14	21	28	42	84
1	1											
2	1	1										
3	1		1									
4	1	1		1								
6	1	1	1		1							
7	1					1						
12	1	1	1	1	1		1					
14	1	1				1		1				
21	1		1			1			1			
28	1	1		1		1		1		1		
42	1	1	1		1	1		1	1		1	
84	1	1	1	1	1	1	1	1	1	1	1	1

Матрица отношений строгого порядка отличается тем, что все элементы главной диагонали нулевые (антирефлексивнасть), а квазипорядка — допустимостью симметричных единичных элементов.

3.2. Эквивалентность и порядок. Изоморфизмы

3.2.1. Отношения эквивалентности и порядка

Напомним, что бинарным отношением на множестве *X* называется подмножество $R \subset X \times X$; вместо $\langle x_1, x_2 \rangle \in R$ часто пишут x_1Rx_2.

Бинарное отношение *R* на множестве *X* называется отношением эквивалентности, если выполнены следующие свойства:

- (рефлексивность) *xRx* для всех $x \in X$;
- (симметричность) $xRy \Rightarrow yRx$ для всех $x, y \in X$;
- (транзитивность) xRy и $yRz \Rightarrow xRz$ для любых элементов $x, y, z \in X$.

Имеет место следующее часто используемое утверждение:

Теорема 1. (а) Если множество *X* разбито в объединение непересекающихся подмножеств, то отношение " лежать в одном подмножестве" является отношением эквивалентности.

(б) Всякое отношение эквивалентности получается описанным способом из некоторого разбиения.

Доказательство. Первое утверждение совсем очевидно; мы приведем доказательство второго, чтобы было видно, где используются все пункты определения эквивалентности. Итак, пусть *R* - отношение эквивалентности. Для каждого элемента $x \in X$ рассмотрим его класс эквивалентности - множество всех $y \in X$, для которых верно *xRy*.

Докажем, что для двух различных x_1, x_2 такие множества либо не пересекаются, либо совпадают. Пусть они пересекаются, то есть имеют общий элемент z. Тогда x_1Rz и x_2Rz, откуда zRx_2 (симметричность) и x_1Rx_2 (транзитивность), а также x_2Rx_1 (симметричность). Поэтому для любого z из x_1Rz следует x_2Rz (транзитивность) и наоборот.

Осталось заметить, что в силу рефлексивности каждый элемент x принадлежит задаваемому им классу, то есть действительно все множество X разбито на непересекающиеся классы.

Теорема Рамсея. Множество всех k-элементных подмножеств бесконечного множества A разбито на l классов (k, l - натуральные числа). Найдется бесконечное множество $B \subset A$, все k- элементные подмножества которого принадлежат одному классу. (При k=1 это очевидно: если бесконечное множество разбито на конечное число классов, то один из классов бесконечен. При k=2 и l=2 утверждение можно сформулировать так: из бесконечного множества людей можно выбрать либо бесконечно много попарно знакомых, либо бесконечно много попарно незнакомых. Конечный вариант этого утверждения - о том, что среди любых шести людей есть либо три попарно знакомых, либо три попарно незнакомых, - известная задача для школьников.)

Множество классов эквивалентности называют фактор - множеством множества X по отношению эквивалентности R. (Если отношение согласовано с дополнительными структурами на X, получаются фактор - группы, фактор - кольца и т.д)

Отношения эквивалентности нам не раз еще встретятся, но сейчас наша основная тема - отношения порядка.

Бинарное отношение $\leq$ на множестве X называется отношением частичного порядка, если выполнены такие свойства:

- (рефлексивность) $x \leq x$ для всех $x \in X$;
- (антисимметричность) $x \leqslant y$ и $y \leqslant x \Rightarrow x = y$ для всех $x, y \in X$;
- (транзитивность) $x \leqslant y$ и $y \leqslant z \Rightarrow x \leqslant z$ для всех

$x, y, z \in X$.

(Следуя традиции, мы используем символ ≤ (а не букву) как знак отношения порядка.) Множество с заданным на нем отношением частичного порядка называют частично упорядоченным.

Говорят, что два элемента *x, y* частично упорядоченного множества сравнимы, если $x \leq y$ или $y \leq x$. Заметим, что определение частичного порядка не требует, чтобы любые два элемента множества были сравнимы. Добавив это требование, мы получим определение линейного порядка (линейно упорядоченного множества).

Приведем несколько примеров частичных порядков:

- Числовые множества с обычным отношением порядка (здесь порядок будет линейным).
- На множестве $\mathbb{R} \times \mathbb{R}$ всех пар действительных чисел можно ввести частичный порядок, считая, что $\langle x_1, x_2 \rangle \leqslant \langle y_1, y_2 \rangle$, если x_1Rx_2 и y_1Ry_2. Этот порядок уже не будет линейным: пары <0, 1> и <1, 0> не сравнимы.
- На множестве функций с действительными аргументами и значениями можно ввести частичный порядок, считая, что $f \leq g$, если $f(x) \leq g(x)$ при всех $x \in \mathbb{R}$. Этот порядок не будет линейным.
- На множестве целых положительных чисел можно определить порядок, считая, что $x \leq y$, если *x* делит *y*. Этот порядок тоже не будет линейным.
- Отношение " любой простой делитель числа *x* является также и делителем числа *y*" не будет отношением порядка на множестве целых положительных чисел (оно рефлексивно и транзитивно, но не антисимметрично).
- Пусть *U* - произвольное множество. Тогда на множестве *P(U)* всех подмножеств множества *U* отношение включения $\subset$ будет частичным порядком.
- На буквах русского алфавита традиция определяет некоторый порядок (а≤б≤в≤…≤я). Этот порядок линеен - про любые две буквы можно сказать, какая из них раньше (при необходимости заглянув в словарь).

- На словах русского алфавита определен лексикографический порядок (как в словаре). Формально определить его можно так: если слово x является началом слова y, то $x \leq y$ (например, кант≤кантор). Если ни одно из слов не является началом другого, посмотрим на первую по порядку букву, в которой слова отличаются: то слово, где эта буква меньше в алфавитном порядке, и будет меньше. Этот порядок также линеен (иначе что бы делали составители словарей?).
- Отношение равенства ($(x \leqslant y) \Leftrightarrow (x = y)$) также является отношением частичного порядка, для которого никакие два различных элемента не сравнимы.
- Приведем теперь бытовой пример. Пусть есть множество X картонных коробок. Введем на нем порядок, считая, что $x \leq y$, если коробка x целиком помещается внутрь коробки y (или если x и y - одна и та же коробка). В зависимости от набора коробок этот порядок может быть или не быть линейным.

Пусть x, y - элементы частично упорядоченного множества X. Говорят, что $x<y$, если $x \leq y$ и $x \neq y$. Для этого отношения выполнены такие свойства:

$$x \nless x;$$

$$(x < y) \text{ и } (y < z) \ \Rightarrow \ x < z.$$

(Первое очевидно, проверим второе: если $x<y$ и $y<z$, то есть $x \leq y$, $x \neq y$, $y \leq z$, $y \neq z$, то $x \leq z$ по транзитивности; если бы оказалось, что $x=z$, то мы бы имели $x \leq y \leq x$ и потому $x=y$ по антисимметричности, что противоречит предположению.)

Терминологическое замечание: мы читаем знак ≤ как " меньше или равно", а знак < - как " меньше", неявно предполагая, что $x \leq y$ тогда и только тогда, когда $x \leq y$, $x<y$ или $x=y$. К счастью, это действительно так. Еще одно замечание: выражение $x>y$ (" x больше y") означает, что $y<x$, а выражение $x \geq y$ (" x больше или равно y ") означает, что $y \leq x$.

В некоторых книжках отношение частичного порядка определяется как отношение <, удовлетворяющее двум указанным свойствам. В этом случае отношение $x \leqslant y \Leftrightarrow [(x < y) \text{ или } (x = y)]$ является отношением частичного порядка в смысле нашего определения.

Во избежание путаницы отношение < иногда называют отношением строгого порядка, а отношение ≤ - отношением нестрогого порядка. Одно и то же частично упорядоченное множество можно задавать по - разному: можно сначала определить отношение нестрогого порядка ≤ (рефлексивное, антисимметричное и транзитивное) и затем из него получить отношение строгого порядка <, а можно действовать и наоборот.

Опуская требование антисимметричности в определении частичного порядка, получаем определение предпорядка. Любой предпорядок устроен так: множество делится на непересекающиеся классы, при этом $x \leq y$ для любых двух элементов x, y из одного класса, а на фактор - множестве задан частичный порядок, который и определяет результат сравнения двух элементов из разных классов.

Рассмотрим несколько конструкций, позволяющих строить одни упорядоченные множества из других.

- Пусть Y - подмножество частично упорядоченного множества $(X, \leq)$. Тогда на множестве Y возникает естественный частичный порядок, индуцированный из X. Формально говоря,

$$(\leqslant_Y) = (\leqslant) \cap (Y \times Y).$$

  Если порядок на X был линейным, то и индуцированный порядок на Y, очевидно, будет линейным.

- Пусть X и Y - два непересекающихся частично упорядоченных множества. Тогда на их объединении можно определить частичный порядок так: внутри каждого множества элементы сравниваются как раньше, а любой элемент множества X по определению меньше любого элемента Y. Это множество естественно обозначить $X+Y$. (Порядок будет линейным, если он был таковым на каждом из множеств.)

  Это же обозначение применяют и для пересекающихся (и даже совпадающих множеств). Например, говоря об упорядоченном множестве $\mathbb{N} + \mathbb{N}$, мы берем две непересекающиеся копии натурального ряда $\{0,1,2,...\}$ и $\{\bar{0},\bar{1},\bar{2},...\}$ и рассматриваем множество $\{0,1,2,...,\bar{0},\bar{1},\bar{2},...\}$, причем $k \leqslant \bar{l}$ при всех *k* и *l*, а внутри каждой копии порядок обычный.

- Пусть $(X, \leq_X)$ и $(Y, \leq_Y)$ - два упорядоченных множества. Можно определить порядок на произведении $X \times Y$ несколькими способами. Можно считать, что $\langle x_1, y_1 \rangle \leqslant \langle x_2, y_2 \rangle$, если $x_1 \leq_X x_2$ и $y_1 \leq_Y y_2$ (покоординатное сравнение). Этот порядок, однако, не будет линейным, даже если исходные порядки и были линейными: если первая координата больше у одной пары, а вторая у другой, как их сравнить? Чтобы получить линейный порядок, договоримся, какая координата будет " главной" и будем сначала сравнивать по ней, а потом (в случае равенства) - по другой. Если главной считать *X* - координату, то $\langle x_1, y_1 \rangle \leqslant \langle x_2, y_2 \rangle$, если $x_1 <_X x_2$ или если $x_1 = x_2$, а $y_1 \leq_Y y_2$. Однако по техническим причинам удобно считать главной вторую координату. Говоря о произведении двух линейно упорядоченных множеств как о линейно упорядоченном множестве, мы в дальнейшем подразумеваем именно такой порядок (сначала сравниваем по второй координате).

В частично упорядоченном множестве $\mathbb{N} \times \mathbb{N}$ (порядок покоординатный) нет бесконечного подмножества, любые два элемента которого были бы несравнимы..

Всякий частичный порядок на конечном множестве можно продолжить до линейного (" продолжить" означает, что если $x \leq y$ в исходном порядке, то и в новом это останется так).

Дано бесконечное частично упорядоченное множество *X*. В нем всегда найдется либо бесконечное подмножество попарно несравнимых элементов, либо бесконечное подмножество, на котором индуцированный порядок линеен.

(Конечный вариант предыдущего утверждения.) Даны целые положительные числа m и n. Во всяком частично упорядоченном множестве мощности $mn+1$ можно указать либо $m+1$ попарно несравнимых элементов, либо $n+1$ попарно сравнимых.

Элемент частично упорядоченного множества называют наибольшим, если он больше любого другого элемента, и максимальным, если не существует большего элемента. Если множество не является линейно упорядоченным, то это не одно и то же: наибольший элемент автоматически является максимальным, но не наоборот. (Одно дело коробка, в которую помещается любая другая, другое - коробка, которая никуда больше не помещается.)

Аналогичным образом определяются наименьшие и минимальные элементы.

Легко понять, что наибольший элемент в данном частично упорядоченном множестве может быть только один, в то время как максимальных элементов может быть много.

3.2.2. Изоморфизмы

Два частично упорядоченных множества называются изоморфными, если между ними существует изоморфизм, то есть взаимно однозначное соответствие, сохраняющее порядок. (Естественно, что в этом случае они равномощны как множества.) Можно сказать так: биекция $f{:}A\to B$ называется изоморфизмом частично упорядоченных множеств A и B, если

$$a_1 \leqslant a_2 \;\Leftrightarrow\; f(a_1) \leqslant f(a_2)$$

для любых элементов $a_1, a_2 \in A$ (слева знак $\leq$ обозначает порядок в множестве A, справа - в множестве B).

Очевидно, что отношение изоморфности рефлексивно (каждое множество изоморфно самому себе), симметрично (если X изоморфно Y, то и наоборот) и транзитивно (два множества, изоморфные третьему, изоморфны между собой). Таким образом, все частично упорядоченные множества разбиваются на классы изоморфных, которые называют порядковыми типами. (Правда, как и с мощностями, тут необходима осторожность - изоморфных множеств слишком много, и потому говорить о порядковых типах как множествах нельзя.)

Теорема 2. Конечные линейно упорядоченные множества из одинакового числа элементов изоморфны.

Доказательство. Конечное линейно упорядоченное множество всегда имеет наименьший элемент (возьмем любой элемент; если он не наименьший, возьмем меньший, если и он не наименьший, еще меньший - и так далее; получим убывающую последовательность $x>y>z>\ldots$, которая рано или поздно должна оборваться). Присвоим наименьшему элементу номер 1. Из оставшихся снова выберем наименьший элемент и присвоим ему номер 2 и так далее. Легко понять, что порядок между элементами соответствует порядку между номерами, то есть что наше множество изоморфно множеству $\{1, 2, \ldots, n\}$.

Будем рассматривать финитные последовательности натуральных чисел, то есть последовательности, у которых все члены, кроме конечного числа, равны 0. На множестве таких последовательностей введем покомпонентный порядок: $(a_0, a_1, \ldots) \leqslant (b_0, b_1, \ldots)$, если $a_i \leq b_i$ при всех i. Это множество изоморфно множеству всех положительных целых чисел с отношением " быть делителем" в качестве порядка.

Взаимно однозначное отображение частично упорядоченного множества A в себя, являющееся изоморфизмом, называют автоморфизмом частично упорядоченного множества A. Тождественное отображение всегда является автоморфизмом, но для некоторых множеств существуют и другие автоморфизмы. Например, отображение прибавления единицы ($x \mapsto x+1$) является автоморфизмом частично упорядоченного множества $\mathbb{Z}$ целых чисел (с естественным порядком). Для множества натуральных чисел та же формула не дает автоморфизма (нет взаимной однозначности).

Не существует автоморфизма упорядоченного множества $\mathbb{N}$ натуральных чисел, отличного от тождественного.

Множество целых положительных чисел, частично упорядоченное отношением " *x* делит *y*", имеет континуум различных автоморфизмов.

Вот несколько примеров равномощных, но не изоморфных линейно упорядоченных множеств (в силу теоремы 12 они должны быть бесконечными).

- Отрезок [0, 1] (с обычным отношением порядка) не изоморфен множеству $\mathbb{R}$, так как у первого есть наибольший элемент, а у второго нет. (При изоморфизме наибольший элемент, естественно, должен соответствовать наибольшему.)
- Множество $\mathbb{Z}$ (целые числа с обычным порядком) не изоморфно множеству $\mathbb{Q}$ (рациональные числа). В самом деле, пусть $\alpha\colon \mathbb{Z} \to \mathbb{Q}$ является изоморфизмом. Возьмем два соседних целых числа, скажем, 2 и 3. При изоморфизме α им должны соответствовать какие - то два рациональных числа α(2) и α(3), причем α(2) < α(3), так как 2<3. Но тогда рациональным числам между α(2) и α(3) должны соответствовать целые числа между 2 и 3, которых нет.
- Более сложный пример - множества $\mathbb{Z}$ и $\mathbb{Z} + \mathbb{Z}$. Возьмем в $\mathbb{Z} + \mathbb{Z}$ две копии нуля (из той и другой компоненты); мы обозначали их 0 и $\bar{0}$. При этом $0 < \bar{0}$. При изоморфизме им должны соответствовать два целых числа *a* и *b*, для которых *a*<*b*. Тогда всем элементам между 0 и $\bar{0}$(их бесконечно много: 1, 2, 3, $\ldots$, $-\bar{3}$, $-\bar{2}$, $-\bar{1}$) должны соответствовать числа между *a* и *b* - но их лишь конечное число.

  Этот пример принципиально отличается от предыдущих тем, что здесь разницу между свойствами множеств нельзя записать формулой. Как говорят, упорядоченные множества $\mathbb{Z}$ и $\mathbb{Z} + \mathbb{Z}$ " элементарно эквивалентны".

Отображение $x \mapsto \sqrt{2}x$ осуществляет изоморфизм между интервалами (0, 1) и (0, √2). Но уже не так просто построить изоморфизм между множествами рациональных точек этих интервалов (то есть между $\mathbb{Q} \cap (0,1)$ и $\mathbb{Q} \cap (0,\sqrt{2})$), поскольку умножение на √2 переводит рациональные числа в иррациональные. Тем не менее изоморфизм построить можно. Для этого надо взять возрастающие последовательности рациональных чисел $0 < x_1 < x_2 < \ldots$ и $0 < y_1 < y_2 < \ldots$, сходящиеся соответственно к 1 и √2 и построить кусочно - линейную функцию *f*, которая переводит x_i в y_i и линейна на каждом из отрезков $[x_i, x_{i+1}]$ (рис.1). Легко понять, что она будет искомым изоморфизмом.

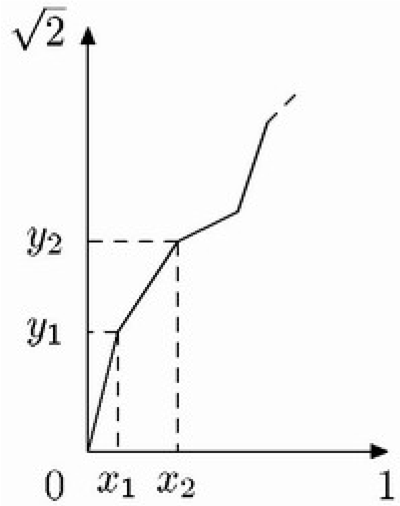


Рис. 1. Ломаная осуществляет изоморфизм

Множество рациональных чисел интервала (0, 1) и множество $\mathbb{Q}$ изоморфны. (Указание: здесь тоже можно построить ломаную; впрочем, у этой задачи есть и другое решение, которое начинается с того, что функция $x \mapsto 1/x$ переводит рациональные числа в рациональные.)

Более сложная конструкция требуется в следующей задаче (видимо, ничего проще, чем сослаться на общую теорему 3, тут не придумаешь).

Множество двоично - рациональных чисел интервала (0, 1) изоморфно множеству $\mathbb{Q}$. (Число считается двоично - рациональным, если оно имеет вид $m/2^n$, где m- целое число, а n - натуральное.)

Два элемента x, y линейно упорядоченного множества называют соседними, если $x<y$ и не существует элемента между ними, то есть такого z, что $x<z<y$. Линейно упорядоченное множество называют плотным, если в нем нет соседних элементов (то есть между любыми двумя есть третий).

Теорема 3. Любые два счетных плотных линейно упорядоченных множества без наибольшего и наименьшего элементов изоморфны.

Доказательство. Пусть X и Y - данные нам множества. Требуемый изоморфизм между ними строится по шагам. После n шагов у нас есть два n- элементных подмножества $X_n \subset X$ и $Y_n \subset Y$, элементы которых мы будем называть " охваченными", и взаимно однозначное соответствие между ними, сохраняющее порядок. На очередном шаге мы берем какой-то неохваченный элемент одного из множеств (скажем, множества X) и сравниваем его со всеми охваченными элементами X. Он может оказаться либо меньше всех, либо больше, либо попасть между какими-то двумя. В каждом из случаев мы можем найти неохваченный элемент в Y, находящийся в том же положении (больше всех, между первым и вторым охваченным сверху, между вторым и третьим охваченным сверху и т.п.). При этом мы пользуемся тем, что в Y нет наименьшего элемента, нет наибольшего и нет соседних элементов, - в зависимости от того, какой из трех случаев имеет место. После этого мы добавляем выбранные элементы к X_n и Y_n, считая их соответствующими друг другу.

Чтобы в пределе получить изоморфизм между множествами X и Y, мы должны позаботиться о том, чтобы все элементы обоих множеств были рано или поздно охвачены. Это можно сделать так: поскольку каждое из множеств счетно, пронумеруем его элементы и будем выбирать неохваченный элемент с наименьшим номером (на нечетных шагах - из X, на четных - из Y). Это соображение завершает доказательство.

Теорема 4. Всякое счетное линейно упорядоченное множество изоморфно некоторому подмножеству множества $\mathbb{Q}$.

Доказательство. Заметим сразу же, что вместо множества $\mathbb{Q}$ можно было взять любое плотное счетное всюду плотное множество без первого и последнего элементов, так как они все изоморфны.

Доказательство этого утверждения происходит так же, как и в теореме 3 - с той разницей, что новые необработанные элементы берутся только с одной стороны (из данного нам множества), а пары к ним подбираются в множестве рациональных чисел.

3.3. Фундированные и вполне упорядоченные множества

3.3.1.Фундированные множества

Принцип математической индукции в одной из возможных форм звучит так:

Пусть *A*(*n*) - некоторое свойство натурального числа *n*. Пусть нам удалось доказать *A*(*n*) в предположении, что *A*(*m*) верно для всех *m*, меньших *n*. Тогда свойство *A*(*n*) верно для всех натуральных чисел *n*. (Заметим, что по условию доказательство *A*(0) возможно без всяких предположений, поскольку меньших чисел нет.)

Для каких частично упорядоченных множеств верен аналогичный принцип? Ответ дается следующей простой теоремой:

Теорема 5. Следующие три свойства частично упорядоченного множества *X* равносильны:

(а) любое непустое подмножество *X* имеет минимальный элемент;

(б) не существует бесконечной строго убывающей последовательности $x_0 > x_1 > x_2 > \ldots$ элементов множества *X*;

(в) для множества *X* верен принцип индукции в следующей форме: если (при каждом $x \in X$) из истинности *A*(*y*) для всех *y*<*x* следует истинность *A*(*x*), то свойство *A*(*x*) верно при всех *x*. Формально это записывают так:

$$\forall x\,(\forall y\,((y < x) \Rightarrow A(y)) \Rightarrow A(x)) \Rightarrow \forall x\, A(x).$$

Доказательство. Сначала докажем эквивалентность первых двух свойств. Если $x_0 > x_1 > x_2 > \ldots$ - бесконечная убывающая последовательность, то, очевидно, множество ее значений не имеет минимального элемента (для каждого элемента следующий еще меньше). Поэтому из (а) следует (б). Напротив, если *B* - непустое множество, не имеющее минимального элемента, то бесконечную убывающую последовательность можно построить так. Возьмем произвольный элемент $b_0 \in B$. По предположению он не является минимальным, так что можно найти $b_1 \in B$, для которого $b_0 > b_1$. По тем же причинам можно найти $b_2 \in B$, для которого $b_1 > b_2$ и т.д. Получается бесконечная убывающая последовательность. Теперь выведем принцип индукции из существования минимального элемента в любом подмножестве. Пусть *A*(*x*) - произвольное свойство элементов множества *X*, верное не для всех элементов *x*. Рассмотрим непустое множество *B* тех элементов, для которых свойство *A* неверно. Пусть *x* - минимальный элемент множества *B*. По условию меньших элементов в множестве *B* нет, поэтому для всех *y*<*x* свойство *A*(*y*) выполнено. Но тогда по предположению должно быть выполнено и *A*(*x*) - противоречие.

Осталось доказать существование минимального элемента в любом непустом подмножестве, исходя из принципа индукции. Пусть *B* - подмножество без минимальных элементов. Докажем по индукции, что *B* пусто; другими словами, в качестве *A*(*x*) возьмем свойство $x \notin B$. В самом деле, если *A*(*y*) верно для всех *y*<*x*, то никакой элемент, меньший *x*, не лежит в *B*. Если бы *x* лежал в *B*, то он был бы там минимальным, а таких нет.

Множества, обладающие свойствами (а)-(в), называются фундированными. Какие есть примеры фундированных множеств? Прежде всего, наш исходный пример - множество натуральных чисел.

Другой пример - множество $\mathbb{N} \times \mathbb{N}$ пар натуральных чисел (меньше та пара, у которой второй член меньше; в случае равенства сравниваем первые). В самом деле, проверим условие (б). Нам будет удобно сформулировать его так: всякая последовательность $u_0 \geqslant u_1 \geqslant u_2 \geqslant \ldots$ элементов множества рано или поздно стабилизируется (все члены, начиная с некоторого, равны); очевидно, что это эквивалентная формулировка.

Пусть дана произвольная последовательность пар

$$\langle x_0, y_0 \rangle \geqslant \langle x_1, y_1 \rangle \geqslant \langle x_2, y_2 \rangle \geqslant \ldots$$

По определению порядка (сначала сравниваются вторые члены) $y_0 \geqslant y_1 \geqslant y_2 \geqslant \ldots$ и потому последовательность натуральных чисел y_i с какого - то места не меняется. После этого уже x_i должны убывать - и тоже стабилизируются. Что и требовалось.

То же самое рассуждение пригодно и в более общей ситуации.

Теорема 6. Пусть *A* и *B* - два фундированных частично упорядоченных множества. Тогда их произведение *A*×*B*, в котором

$$\langle a_1, b_1 \rangle \leqslant \langle a_2, b_2 \rangle \Leftrightarrow [(b_1 < b_2) \text{ или } (b_1 = b_2 \text{ и } a_1 \leqslant a_2)],$$

является фундированным.

Доказательство. В последовательности $\langle a_0, b_0 \rangle \geqslant \langle a_1, b_1 \rangle \geqslant \ldots$ стабилизируются сначала вторые, а затем и первые члены. Отсюда вытекает аналогичное утверждение для $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$, для $\mathbb{N}^k$ или вообще для произведения конечного числа фундированных множеств.

Еще проще доказать, что сумма *A* +*B* двух фундированных множеств *A* и *B* фундирована: последовательность $x_0 \leqslant x_1 \leqslant x_2 \leqslant \ldots$ либо целиком содержится в *B* (и мы ссылаемся на фундированность *B*), либо содержит элемент из *A*. В последнем случае все следующие элементы также принадлежат *A*, и мы используем фундированность *A*.

Часто в программировании (или в олимпиадных задачах) нам нужно доказать, что некоторый процесс не может продолжаться бесконечно долго. Например, написав цикл, мы должны убедиться, что рано или поздно из него выйдем. Это можно сделать так: ввести какой - то натуральный параметр и убедиться, что на каждом шаге цикла этот параметр уменьшается. Тогда, если сейчас этот параметр равен N, то можно гарантировать, что не позже чем через N шагов цикл закончится.

Однако бывают ситуации, в которых число шагов заранее оценить нельзя, но тем не менее гарантировать завершение цикла можно, поскольку есть параметр, принимающий значения в фундированном множестве и убывающий на каждом шаге цикла.

Вот пример олимпиадной задачи, где по существу такое рассуждение и используется.

Задача. Бизнесмен заключил с чертом сделку: каждый день он дает черту одну монету, и в обмен получает любой набор монет по своему выбору, но все эти монеты меньшего достоинства (видов монет конечное число). Менять (или получать) деньги в другом месте бизнесмен не может. Когда монет больше не останется, бизнесмен проигрывает. Докажите, что рано или поздно черт выиграет, каков бы ни был начальный набор монет у бизнесмена.

Решение: пусть имеется k видов монет. Искомый параметр определим так: посчитаем, сколько монет каждого вида есть у бизнесмена (n_1 - число монет минимального достоинства, n_2 - число следующих, и так далее до n_k). Заметим, что в результате встречи с чертом набор $\langle n_1, \ldots, n_k \rangle$ уменьшается (в смысле введенного нами порядка, когда мы сравниваем сначала последние члены, затем предпоследние и т.д.). Поскольку множество $\mathbb{N}^k$ фундировано, этот процесс должен оборваться.

Имеется конечная последовательность нулей и единиц. За один шаг разрешается сделать такое действие: найти в ней группу 01 и заменить на 100…00 (при этом можно написать сколько угодно нулей). Такие шаги нельзя выполнять бесконечно много раз.

Рассмотрим множество невозрастающих последовательностей натуральных чисел, в которых все члены, начиная с некоторого, равны нулю. Введем в нем порядок так: сначала сравниваем первые члены, при равенстве первых вторые и т.д. Это (линейно) упорядоченное множество фундировано.

Рассмотрим множество всех многочленов от одной переменной x, коэффициенты которых - натуральные числа. Упорядочим его так: многочлен P больше многочлена Q, если $P(x)>Q(x)$ для всех достаточно больших x. Это определение задает линейный порядок и получающееся упорядоченное множество фундировано.

3.3.2. Вполне упорядоченные множества

Фундированные линейно упорядоченные множества называются вполне упорядоченными, а соответствующие порядки - полными. Для линейных порядков понятия наименьшего и минимального элемента совпадают, так что во вполне упорядоченном множестве всякое непустое подмножество имеет наименьший элемент.

Заметим, что частично упорядоченное множество, в котором всякое непустое подмножество имеет наименьший элемент, автоматически является линейно упорядоченным (в самом деле, всякое двухэлементное множество имеет наименьший элемент, поэтому любые два элемента сравнимы).

Примеры вполне упорядоченных множеств: $\mathbb{N}$, $\mathbb{N}+k$ (здесь k обозначает конечное линейно упорядоченное множество из k элементов), $\mathbb{N}+\mathbb{N}$, $\mathbb{N}\times\mathbb{N}$.

Наша цель - понять, как могут быть устроены вполне упорядоченные множества. Начнем с нескольких простых замечаний.

- Вполне упорядоченное множество имеет наименьший элемент. (Непосредственное следствие определения.)
- Для каждого элемента x вполне упорядоченного множества (кроме наибольшего) есть непосредственно следующий за ним элемент y (это значит, что $x>y$, но не существует z, для которого $y>z>x$). В самом деле, если множество всех

элементов, больших x, непусто, то в нем есть минимальный элемент y, который и будет искомым. Такой элемент логично обозначать $x+1$, следующий за ним – $x+2$ и т.д.

- Некоторые элементы вполне упорядоченного множества могут не иметь непосредственно предыдущего. Например, в множестве $\mathbb{N}+\mathbb{N}$ есть два элемента, не имеющих непосредственно предыдущего (наименьший элемент, а также наименьший элемент второй копии натурального ряда). Такие элементы называют предельными.
- Всякий элемент упорядоченного множества имеет вид $z+n$, где z - предельный, а n - натуральное число (обозначение $z+n$ понимается в описанном выше смысле). В самом деле, если z не предельный, возьмем предыдущий, если и он непредельный - то его предыдущий и т.д., пока не дойдем до предельного (бесконечно продолжаться это не может, так как множество вполне упорядочено). Очевидно, такое представление однозначно (у элемента может быть только один непосредственно предыдущий).
- Любое ограниченное сверху множество элементов вполне упорядоченного множества имеет точную верхнюю грань. (Как обычно, подмножество X частично упорядоченного множества A называется ограниченным сверху, если оно имеет верхнюю границу, т.е. элемент $a \in A$, для которого $x \le a$ при всех $x \in X$. Если среди всех верхних границ данного подмножества есть наименьшая, то она называется точной верхней гранью.)

В самом деле, множество всех верхних границ непусто и потому имеет наименьший элемент. (Заметим в скобках, что вопрос о точной нижней грани для вполне упорядоченного множества тривиален, так как всякое множество имеет наименьший элемент.)

Пусть A - произвольное вполне упорядоченное множество. Его наименьший элемент обозначим через 0. Следующий за ним элемент обозначим через 1, следующий за 1 - через 2 и т.д. Если множество конечно, процесс этот оборвется. Если бесконечно, посмотрим, исчерпали ли мы все элементы множества A. Если нет, возьмем минимальный элемент из оставшихся. Обозначим его w. Следующий за ним элемент (если он есть) обозначим w +1, затем w +2 и т.д. Если и на этом множество не исчерпается, то возьмем наименьший элемент из оставшихся, назовем его $w\cdot 2$, и повторим всю процедуру. Затем будут $w\cdot 3$, $w\cdot 4$ и т.д. Если и на этом множество не кончится, минимальный из оставшихся элементов назовем w^2. Затем пойдут ω^2+1, ω^2+2, $\ldots$, $\omega^2+\omega$, $\ldots$, $\omega^2+\omega\cdot 2$, $\ldots$, $\omega^2\cdot 2$, $\ldots$, $\omega^2\cdot 3$, $\ldots$, ω^3, $\ldots$ (мы не поясняем сейчас подробно обозначения).

Что, собственно говоря, доказывает это рассуждение? Попытаемся выделить некоторые утверждения. При этом полезно такое определение: если линейно упорядоченное множество A разбито на две (непересекающиеся) части B и C, причем любой элемент B меньше любого элемента C, то B называют начальным отрезком множества A. Другими словами, подмножество B линейно упорядоченного множества A является начальным отрезком, если любой элемент B меньше любого элемента $A\backslash B$. Еще одна переформулировка: $B\subset A$ является начальным отрезком, если из $a,b\in A$, $b\in B$ и $a\leq b$ следует $a\in B$. Заметим, что начальный отрезок может быть пустым или совпадать со всем множеством.

Отметим сразу же несколько простых свойств начальных отрезков:

- Начальный отрезок вполне упорядоченного множества (как, впрочем, и любое подмножество) является вполне упорядоченным множеством.
- Начальный отрезок начального отрезка есть начальный отрезок исходного множества.
- Объединение любого семейства начальных отрезков (в одном и том же упорядоченном множестве) есть начальный отрезок того же множества.
- Если x - произвольный элемент вполне упорядоченного множества A, то множества $[0, x)$ (все элементы множества A, меньшие x) и $[0, x]$ (элементы множества A, меньшие или равные x) являются начальными отрезками.
- Всякий начальный отрезок I вполне упорядоченного множества A, не совпадающий со всем множеством, имеет вид $[0, x)$ для некоторого $x\in A$. (В самом деле, если $I\neq A$, возьмем наименьший элемент x в множестве $A\backslash I$. Тогда все меньшие элементы принадлежат I, сам x не принадлежит I и

все бóльшие x элементы не принадлежат I, иначе получилось бы противоречие с определением начального отрезка.)

- Любые два начальных отрезка вполне упорядоченного множества сравнимы по включению, т.е. один есть подмножество другого. (Следует из предыдущего.)
- Начальные отрезки вполне упорядоченного множества A, упорядоченные по включению, образуют вполне упорядоченное множество. Это множество состоит из наибольшего элемента (все A) и остальной части, изоморфной множеству A. (В самом деле, начальные отрезки множества A, не совпадающие с A, имеют вид $[0, x)$, и соответствие $[0,x) \leftrightarrow x$ будет изоморфизмом.)

Возвратимся к нашему рассуждению с последовательным выделением различных элементов из вполне упорядоченного множества. Его первую часть можно считать доказательством такого утверждения: если вполне упорядоченное множество бесконечно, то оно имеет начальный отрезок, изоморфный w. (Говоря о множестве натуральных чисел вместе с порядком, обычно употребляют обозначение w, а не $\mathbb{N}$.)

Но на этом наше рассуждение не оканчивается. Его следующая часть может считаться доказательством такого факта: либо A изоморфно некоторому начальному отрезку множества w^2, либо оно имеет начальный отрезок, изоморфный w^2. (Здесь w^2 - вполне упорядоченное множество пар натуральных чисел: сравниваются сначала вторые компоненты пар, а при их равенстве - первые.)

Вообще верно такое утверждение: для любых двух вполне упорядоченных множеств одно изоморфно начальному отрезку другого, и доказательство состоит более или менее в повторении проведенного рассуждения. Но чтобы сделать это аккуратно, нужна некоторая подготовка.

4. Трансфинитная индукция

Термины " индукция" и " рекурсия" часто употребляются вперемежку. Например, определение факториала $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$ как функции $f(n)$, для которой $f(n) = n \cdot f(n-1)$ при $n>0$ и $f(0)=1$, называют и " индуктивным", и " рекурсивным". Мы будем стараться разграничивать эти слова так: если речь идет о доказательстве чего-то сначала для $n=0$, затем для $n=1, 2, \ldots$, причем каждое утверждение опирается на предыдущее, то это индукция. Если же мы определяем что-то сначала для $n=0$, потом для $n=1, 2, \ldots$, причем определение каждого нового значения использует ранее определенные, то это рекурсия.

Наша цель - научиться проводить индуктивные доказательства и давать рекурсивные определения не только для натуральных чисел, но и для других вполне упорядоченных множеств.

Доказательства по индукции мы уже обсуждали, говоря о фундированных множествах, и сейчас ограничимся только одним примером.

Теорема 7. Пусть A - вполне упорядоченное множество, а $f{:}A\to A$ - возрастающее отображение (то есть $f(x)<f(y)$ при $x<y$). Тогда $f(x) \geq x$ для всех $x \in A$.

Доказательство. Согласно принципу индукции (теорема 5) достаточно доказать неравенство $f(x)\geq x$, предполагая, что $f(y)\geq y$ при всех $y<x$. Пусть это не так и $f(x)<x$. Тогда по монотонности $f(f(x))<f(x)$. Но, с другой стороны, элемент $y=f(x)$ меньше x, и потому по предположению индукции $f(y)\geq y$, то есть $f(f(x))\geq f(x)$.

Если угодно, можно в явном виде воспользоваться существованием наименьшего элемента и изложить это же рассуждение так. Пусть утверждение теоремы неверно. Возьмем наименьшее x, для которого $f(x)<x$. Но тогда $f(f(x))<f(x)$ по монотонности и потому x не является наименьшим вопреки предположению.

Наконец, это рассуждение можно пересказать и так: если $x>f(x)$, то по монотонности

$$x > f(x) > f(f(x)) > f(f(f(x))) > \ldots,$$

но бесконечных убывающих последовательностей в фундированном множестве быть не может.

Теперь перейдем к рекурсии. В определении факториала $f(n)$ выражалось через $f(n\text{-}1)$. В общей ситуации значение $f(n)$ может использовать не только одно предыдущее значение функции, но и все значения на меньших аргументах. Например, можно определить функцию $f\colon \mathbb{N} \to \mathbb{N}$, сказав, что $f(n)$ на единицу больше суммы всех предыдущих значений, то есть $f(n) = f(0) + f(1) + \ldots + f(n-1) + 1$; это вполне законное рекурсивное определение (надо только пояснить, что пустая сумма считается равной нулю, так что $f(0)=1$).

Как обобщить эту схему на произвольные вполне упорядоченные множества вместо натурального ряда? Пусть A вполне упорядочено. Мы хотим дать рекурсивное определение некоторой функции $f{:}A{\to}B$ (где B - некоторое множество). Такое определение должно связывать значение $f(x)$ на некотором элементе $x \in A$ со значениями $f(y)$ при всех $y<x$. Другими словами, рекурсивное определение задает $f(x)$, предполагая известным ограничение функции f на начальный отрезок $[0, x)$. Вот точная формулировка:

Теорема 8. Пусть A - вполне упорядоченное множество. Пусть B - произвольное множество. Пусть имеется некоторое рекурсивное правило, то есть отображение F, которое ставит в соответствие элементу $x \in A$ и функции $g\colon [0, x) \to B$ некоторый элемент множества B. Тогда существует и единственна функция $f{:}A{\to}B$, для которой

$$f(x) = F(x, f|_{[0,x)})$$

при всех $x \in A$. (Здесь $f|_{[0,x)}$ обозначает ограничение функции f на начальный отрезок $[0, x)$ - мы отбрасываем все значения функции на элементах, больших или равных x.)

Доказательство. Неформально можно рассуждать так: значение f на минимальном элементе определено однозначно, так как предыдущих значений нет (сужение $f|_{[0,0)}$ пусто). Тогда и на следующем элементе значение функции f определено однозначно, поскольку на предыдущих (точнее, единственном предыдущем) функция f уже задана, и т.д.

Конечно, это надо аккуратно выразить формально. Вот как это делается. Докажем по индукции такое утверждение о произвольном элементе $a \in A$:

существует и единственно отображение *f* отрезка [0, *a*] в множество *B*, для которого рекурсивное определение (равенство, приведенное в условии) выполнено при всех $x \in [0, a]$.

Будем называть отображение *f*: [0, *a*] →*B*, обладающее указанным свойством, корректным. Таким образом, мы хотим доказать, что для каждого $a \in A$ есть единственное корректное отображение отрезка [0, *a*] в *B*.

Поскольку мы рассуждаем по индукции, можно предполагать, что для всех $c<a$ это утверждение выполнено, то есть существует и единственно корректное отображение $f_c : [0, c] \to B$. (Корректность f_c означает, что при всех $d \le c$ значение $f_c(d)$ совпадает с предписанным по рекурсивному правилу.)

Рассмотрим отображения f_{c_1} и f_{c_2} для двух различных c_1 и c_2. Пусть, например, $c_1<c_2$. Отображение f_{c_2} определено на большем отрезке [0, c_2]. Если ограничить f_{c_2} на меньший отрезок [0, c_1], то оно совпадет с f_{c_1}, поскольку ограничение корректного отображения на меньший отрезок корректно (это очевидно), а мы предполагали единственность на отрезке [0, c_1].

Таким образом, все отображения f_c согласованы друг с другом, то есть принимают одинаковое значение, если определены одновременно. Объединив их, мы получаем некоторое единое отображение *h*, определенное на [0, *a*). Применив к *a* и *h* рекурсивное правило, получим некоторое значение $b \in B$. Доопределим *h* в точке *a*, положив *h(a)=b*. Получится отображение *h*: [0, *a*] →*B*; легко понять, что оно корректно.

Чтобы завершить индуктивный переход, надо проверить, что на отрезке [0, *a*] корректное отображение единственно. В самом деле, его ограничения на отрезки [0, *c*] при $c<a$ должны совпадать с f_c, поэтому осталось проверить однозначность в точке *a* - что гарантируется рекурсивным определением (выражающим значение в точке *a* через предыдущие). На этом индуктивное доказательство заканчивается.

Осталось лишь заметить, что для разных *a* корректные отображения отрезков [0, *a*] согласованы друг с другом (сужение корректного отображения на меньший отрезок корректно, применяем единственность) и потому вместе задают некоторую функцию $f\colon A \to B$, удовлетворяющую рекурсивному определению.

Существование доказано; единственность тоже понятна, так как ограничение этой функции на любой отрезок [0, *a*] корректно и потому однозначно определено, как мы видели.

Прежде чем применить эту теорему и доказать, что из двух вполне упорядоченных множеств одно является отрезком другого, нам потребуется ее немного усовершенствовать. Нам надо предусмотреть ситуацию, когда рекурсивное правило не всюду определено. Пусть, например, мы определяем последовательность действительных чисел соотношением $x_n = \operatorname{tg} x_{n-1}$ и начальным условием $x_0 = a$. При некоторых значениях *a* может оказаться, что построение последовательности обрывается, поскольку тангенс не определен для соответствующего аргумента.

Множество всех таких " исключительных" *a* (когда последовательность конечна) счетно.

Аналогичная ситуация возможна и для общего случая.

Теорема 9. Пусть отображение *F*, о котором шла речь в теореме 8, является частичным (для некоторых *x* и функций g:[0, *a*)→*B* оно может быть не определено). Тогда существует функция *f*, которая

- либо определена на всем *A* и согласована с рекурсивным определением;

- либо определена на некотором начальном отрезке $[0, a)$ и на нем согласована с рекурсивным определением, причем для точки a и функции f рекурсивное правило неприменимо (отображение F не определено).

Доказательство. Это утверждение является обобщением, но одновременно и следствием предыдущей теоремы 8. В самом деле, добавим к множеству B специальный элемент $\perp$ ("неопределенность") и модифицируем рекурсивное правило: новое правило дает значение $\perp$, когда старое было не определено. (Если среди значений функции на предыдущих аргументах уже встречалось $\perp$, новое рекурсивное правило тоже дает $\perp$.)

Применив теорему 8 к модифицированному правилу, получим некоторую функцию f'. Если эта функция нигде не принимает значения $\perp$, то реализуется первая из двух возможностей, указанных в теореме (при $f=f'$). Если же функция f' принимает значение $\perp$ в какой-то точке, то она имеет то же значение $\perp$ и во всех больших точках. Заменив значение $\perp$ на неопределенность, мы получаем из функции f' функцию f. Область определения функции f есть некоторый начальный отрезок $[0, a)$ и реализуется вторая возможность, указанная в формулировке теоремы.

Теперь у нас все готово для доказательства теоремы о сравнении вполне упорядоченных множеств.

Теорема 10. Пусть A и B - два вполне упорядоченных множества. Тогда либо A изоморфно некоторому начальному отрезку множества B, либо B изоморфно некоторому начальному отрезку множества A.

Доказательство. Отметим прежде всего, что начальный отрезок может совпадать со всем множеством, так что случай изоморфных множеств A и B также покрывается этой теоремой.

Определим отображение f из A в B таким рекурсивным правилом: для любого $a \in A$ $f(a)$ есть наименьший элемент множества B, который не встречается среди $f(a')$ при $a'<a$.

Это правило не определено в том случае, когда значения $f(a')$ при $a'<a$ покрывают все B. Применяя теорему 9, мы получаем функцию f, согласованную с этим правилом. Теперь рассмотрим два случая:

- Функция f определена на всем A. Заметим, что рекурсивное определение гарантирует монотонность, поскольку $f(a)$ определяется как минимальный еще не использованный элемент; чем больше a, тем меньше остается неиспользованных элементов и потому минимальный элемент может только возрасти (из определения следует также, что одинаковых значений быть не может). Остается лишь проверить, что множество значений функции f, то есть $f(A)$, будет начальным отрезком. В самом деле, пусть $b<f(a)$ для некоторого $a \in A$; надо проверить, что b также является значением функции f. Действительно, согласно рекурсивному определению $f(a)$ является наименьшим неиспользованным значением, следовательно, b уже использовано, то есть встречается среди $f(a')$ при $a'<a$.
- Функция f определена лишь на некотором начальном отрезке $[0, a)$. В этом случае этот начальный отрезок изоморфен B, и функция f является искомым изоморфизмом. В самом деле, раз $f(a)$ не определено, то среди значений функции f встречаются все элементы множества B. С другой стороны, f сохраняет порядок в силу рекурсивного определения.

Таким образом, в обоих случаях утверждение теоремы верно.

Может ли быть так, что A изоморфно начальному отрезку B, а B изоморфно начальному отрезку A? Нет - за исключением тривиального случая, когда начальные отрезки представляют собой сами множества A и B. Это вытекает из такого утверждения:

Теорема 11. Никакое вполне упорядоченное множество не изоморфно своему начальному отрезку (не совпадающему со всем множеством).

Доказательство. Пусть вполне упорядоченное множество A изоморфно своему начальному отрезку, не совпадающему со всем множеством. Как мы видели, этот отрезок имеет вид $[0, a)$ для некоторого элемента $a \in A$. Пусть $f: A \to [0, a)$ - изоморфизм.

Тогда f строго возрастает, и по теореме 7 имеет место неравенство $f(a) \geqslant a$, что противоречит тому, что множество значений функции fесть $[0, a)$.

Если множество A изоморфно начальному отрезку множества B, а множество B изоморфно начальному отрезку множества A, то композиция этих изоморфизмов дает изоморфизм между множеством A и его начальным отрезком (начальный отрезок начального отрезка есть начальный отрезок). Этот начальный отрезок обязан совпадать со всем множеством A, так что это возможно лишь если A и B изоморфны.

Сказанное позволяет сравнивать вполне упорядоченные множества. Если A изоморфно начальному отрезку множества B, не совпадающему со всем B, то говорят, что порядковый тип множества A меньше порядкового типа множества B. Если множества A и B изоморфны, то говорят, что у них одинаковые порядковые типы. Наконец, если B изоморфно начальному отрезку множества A, то говорят, что порядковый тип множества A больше порядкового типа множества B. Как мы только что доказали, верно такое утверждение:

Теорема 12. Для любых вполне упорядоченных множеств A и B имеет место ровно один из указанных трех случаев.

Если временно забыть о проблемах оснований теории множеств и определить порядковый тип упорядоченного множества как класс изоморфных ему упорядоченных множеств, то можно сказать, что мы определили линейный порядок на порядковых типах вполне упорядоченных множеств (на ординалах, как говорят). Этот порядок будет полным. Мы переформулируем это утверждение так, чтобы избегать упоминания классов.

Теорема 13. Всякое непустое семейство вполне упорядоченных множеств имеет " наименьший элемент" - множество, изоморфное начальным отрезкам всех остальных множеств.

Доказательство. Возьмем какое-то множество X семейства. Если оно наименьшее, то все доказано. Если нет, рассмотрим все множества семейства, которые меньше его, то есть изоморфны его начальным отрезкам вида $[0, x)$. Среди всех таких элементов x выберем наименьший. Тогда соответствующее ему множество и будет наименьшим.

Следствием доказанных теорем является то, что любые два вполне упорядоченных множества сравнимы по мощности (одно равномощно подмножеству другого). Сейчас мы увидим, что всякое множество может быть вполне упорядочено (теорема Цермело), и, следовательно, любые два множества сравнимы по мощности.

4.1.Теорема Цермело и трансфинитная индукция

4.1.1.Теорема Цермело

Теорема 14. (Цермело) Всякое множество может быть вполне упорядочено.

Доказательство. Доказательство этой теоремы существенно использует аксиому выбора и вызывало большие нарекания своей неконструктивностью. На счетных множествах полный порядок указать легко (перенеся с $\mathbb{N}$). Но уже на множестве действительных чисел никакого конкретного полного порядка указать не удается, и доказав (с помощью аксиомы выбора) его существование, мы так и не можем себе этот порядок представить.

Объясним, в какой форме используется аксиома выбора. Пусть *А* - данное нам множество. Мы принимаем, что существует функция φ, определенная на всех подмножествах множества *А*, кроме самого *А*, которая указывает один из элементов вне этого подмножества:

$$X \subsetneq A \Rightarrow \varphi(X) \in A \setminus X.$$

После того, как такая функция фиксирована, можно построить полный порядок на *А*, и в этом построении уже нет никакой неоднозначности. Вот как это делается.

Наименьшим элементом множества *А* мы объявим элемент $a_0 = \varphi(\varnothing)$. За ним идет элемент $a_1 = \varphi(\{a_0\})$; по построению он отличается от a_0. Далее следует элемент $a_2=\varphi(\{a_0,a_1\})$. Если множество *A* бесконечно, то такой процесс можно продолжать и получить последовательность $\{a_0,a_1,\ldots\}$ элементов множества *A*. Если после этого остаются еще не использованные элементы множества *A*, рассмотрим элемент $a_\omega=\varphi(\{a_0,a_1,a_2,\ldots\})$ и так будем продолжать, пока все *A* не кончится; когда оно кончится, порядок выбора элементов и будет полным порядком на *A*.

Конечно, последняя фраза нуждается в уточнении - что значит " так будем продолжать"? Возникает желание применить теорему о трансфинитной рекурсии (у нас очень похожая ситуация: следующий элемент определяется рекурсивно, если известны все предыдущие). И это можно сделать, если у нас есть другое вполне упорядоченное множество *B*, и получить взаимно однозначное соответствие либо между *A* и частью *B*, либо между *B* и частью *A*. В первом случае все хорошо, но для этого надо иметь вполне упорядоченное множество *B* по крайней мере той же мощности, что и *A*, так что получается некий порочный круг.

Тем не менее из него можно выйти. Мы сделаем это так: рассмотрим все потенциальные кусочки будущего порядка и убедимся, что их можно склеить.

Пусть $(S, \leq_S)$ - некоторое подмножество множества *A* и заданный на нем порядок. Будем говорить, что $(S, \leq_S)$ является корректным фрагментом, если оно является вполне упорядоченным множеством, причем

$$s=\varphi([0, s))$$

для любого $s\in S$. Здесь $s[0, s)$ - начальный отрезок множества *S*, состоящий из всех элементов, меньших *s* с точки зрения заданного на *S* порядка.

Например, множество $\{\varphi(\varnothing)\}$ является корректным фрагментом (порядок здесь можно не указывать, так как элемент всего один). Множество $\{\varphi(\varnothing), (\{\varphi(\varnothing)\}\}$ (первый из выписанных элементов считается меньшим второго) также является корректным фрагментом. Это построение можно продолжать и дальше, но нам надо каким-то образом " перескочить" через бесконечное (и очень большое в смысле мощности) число шагов этой конструкции.

План такой: мы докажем, что любые два корректных фрагмента в определенном смысле согласованы, после чего рассмотреть объединение всех корректных фрагментов. Оно будет корректным и будет совпадать со всем множеством A (в противном случае его можно было бы расширить и получить корректный фрагмент, не вошедший в объединение).

Лемма 1. Пусть $(S, \leq_S)$ и $(T, \leq_T)$ - два корректных фрагмента. Тогда один из них является начальным отрезком другого, причем порядки согласованы (два общих элемента все равно как сравнивать - в смысле $\leq_S$ или в смысле $\leq_T$).

Заметим, что по теореме 10 один из фрагментов изоморфен начальному отрезку другого. Пусть S изоморфен начальному отрезку T и $h:S\to T$ - их изоморфизм. Лемма утверждает, что изоморфизм h является тождественным, то есть что $h(x)=x$ при всех $x \in S$. Докажем это индукцией по $x \in S$ (это законно, так как S вполне упорядочено по определению корректного фрагмента). Индуктивное предположение гарантирует, что $h(y)=y$ для всех $y<x$. Мы хотим доказать, что $h(x)=x$. Рассмотрим начальные отрезки $[0, x)_S$ и $[0, h(x))_T$ (с точки зрения порядков $\leq_S$ и $\leq_T$ соответственно). Они соответствуют друг другу при изоморфизме h, поэтому по предположению индукции совпадают как множества. Но по определению корректности $x=\varphi([0, x))$ и $h(x)= \varphi([0, h(x)))$, так что $x=h(x)$. Лемма 1 доказана.

Рассмотрим объединение всех корректных фрагментов (как множеств). На этом объединении естественно определен линейный порядок: для всяких двух элементов найдется фрагмент, которому они оба принадлежат (каждый принадлежит своему, возьмем больший из фрагментов), так что их можно сравнить. По лемме 1 порядок не зависит от того, какой фрагмент будет выбран для сравнения.

Лемма 2. Это объединение будет корректным фрагментом.

Чтобы доказать лемму 2, заметим, что на этом объединении определен линейный порядок. Он будет полным. Для разнообразия объясним это в терминах убывающих последовательностей. Пусть $x_0 \geq x_1 \geq \ldots$; возьмем корректный фрагмент F, которому принадлежит x_0. Из леммы 1 следует, что все x_i также принадлежат этому фрагменту (поскольку фрагмент F будет начальным отрезком в любом большем фрагменте), а F вполне упорядочен по определению, так что последовательность стабилизируется. Лемма 2 доказана.

Утверждение леммы 2 можно переформулировать таким образом: существует наибольший корректный фрагмент. Осталось доказать, что этот фрагмент (обозначим его S) включает в себя все множество A. Если $S \neq A$, возьмем элемент $a=\varphi(S)$, не принадлежащий S, и добавим его к S, считая, что он больше всех элементов S. Полученное упорядоченное множество S' (сумма S и одноэлементного множества) будет, очевидно, вполне упорядочено. Кроме того, условие корректности также выполнено (для a - по построению, для остальных элементов - поскольку оно было выполнено в S). Таким образом, мы построили больший корректный фрагмент, что противоречит максимальности S. Это рассуждение завершает доказательство теоремы Цермело.

Как мы уже говорили, из теоремы Цермело и теоремы 10 о сравнении вполне упорядоченных множеств немедленно вытекает такое утверждение:

Теорема 15. Из любых двух множеств одно равномощно подмножеству другого.

Понятие вполне упорядоченного множества ввел Кантор в работе 1883 года; в его итоговой работе 1895-1897 годов приводится доказательство того, что любые два вполне упорядоченных множества сравнимы (одно изоморфно начальному отрезку другого).

Утверждения о возможности полного упорядочения любого множества и о сравнении мощностей (теоремы 14 и 15) неоднократно встречаются в работах Кантора, но никакого внятного доказательства он не предложил, и оно было дано лишь в 1904 году немецким математиком Э.Цермело.

4.1.2. Трансфинитная индукция и базис Гамеля

Вполне упорядоченные множества и теорема Цермело позволяют продолжать индуктивные построения в трансфинитную область (если выражаться торжественно). Поясним это на примере из линейной алгебры.

Всякое линейно независимое множество векторов в конечномерном пространстве может быть дополнено до базиса. Как это доказывается? Пусть S - данное нам линейно независимое множество. Если оно не является базисом, то некоторый вектор x_0 через него не выражается. Добавим его к S, получим линейно независимое множество $S \cup \{x_0\}$. Если и оно не является базисом, то некоторый вектор x_1 через него не выражается, и т.д. Либо на каком-то шаге мы получим базис, либо процесс не оборвется и мы получим бесконечную последовательность линейно независимых векторов, что противоречит конечномерности.

Теперь с помощью трансфинитной индукции (точнее, рекурсии) мы избавимся от требования конечномерности.

Пусть дано произвольное векторное пространство. Говорят, что множество (возможно, бесконечное) векторов линейно независимо, если никакая нетривиальная линейная комбинация конечного числа векторов из этого множества не равна нулю. (Заметим в скобках, что говорить о бесконечных линейных комбинациях в принципе можно лишь если в пространстве определена сходимость, чего мы сейчас не предполагаем.) Линейно независимое множество векторов называется базисом Гамеля (или просто базисом) данного пространства, если любой вектор представим в виде конечной линейной комбинации элементов этого множества.

Как и в конечной ситуации, максимальное линейно независимое множество (которое становится линейно зависимым при добавлении любого нового элемента) является, очевидно, базисом.

Теорема 16. Всякое линейно независимое множество векторов может быть расширено до базиса Гамеля.

Доказательство. Пусть S - линейно независимое подмножество векторного пространства V. Рассмотрим вполне упорядоченное множество I достаточно большой мощности (большей, чем мощность пространства V). Определим функцию f из I в V с помощью трансфинитной рекурсии:

$f(i)$= элемент пространства V, не выражающийся линейно через элементы S и значения $f(i)$ при $j<i$.

Заметим, что это рекурсивное правило оставляет $f(i)$ неопределенным, если такого невыразимого элемента не существует. (Кроме того, можно отметить, что мы снова используем аксиому выбора. Более подробно следовало бы сказать так: по аксиоме выбора существует некоторая функция, которая по каждому подмножеству пространства V, через которое не все V выражается, указывает один из невыразимых элементов. Затем эта функция используется в рекурсивном определении. Впрочем, аксиома выбора и так уже использована для доказательства теоремы Цермело.)

Это определение гарантирует, что f является инъекцией; более того можно утверждать, что все значения f вместе с множеством S образуют линейно независимое множество. В самом деле, пусть линейная комбинация некоторых значений функции f и элементов множества S равна нулю. Можно считать, что все коэффициенты в этой комбинации отличны от нуля (отбросив нулевые слагаемые). Входящие в комбинацию значения функции f имеют вид $f(i)$ при различных i. Посмотрим на тот из них, который имеет наибольшее i; по построению он должен быть линейно независим от остальных - противоречие.

Поскольку мы предположили, что множество I имеет большую мощность, чем V, рекурсивное определение задает функцию не на всем I, а только на некотором начальном отрезке $[0, i)$, а в точке I рекурсивное правило не определено (теорема 9). Это означает, что все векторы пространства V выражаются через элементы множества S и значения функции f на промежутке $[0, i)$. Кроме того, как мы видели, все эти векторы независимы. Таким образом, искомый базис найден.

На самом деле можно обойтись без множества большей мощности, упорядочив само пространство V. При этом на каждом шаге рекурсии надо либо добавлять очередной элемент к будущему базису (если он не выражается через предыдущие), либо оставлять базис без изменений.

Базис Гамеля может быть использован для построения разных экзотических примеров. Вот некоторые из них:

Теорема 17. Существует (всюду определенная) функция $f\colon \mathbb{R} \to \mathbb{R}$, для которой $f(x+y) = f(x) + f(y)$ при всех *x* и *y*, но которая не есть умножение на константу.

Доказательство. Рассмотрим $\mathbb{R}$ как векторное пространство над полем $\mathbb{Q}$. В нем есть базис Гамеля. Пусть α - один из векторов базиса. Рассмотрим функцию *f*, которая с каждым числом *x* (рассматриваемым как вектор в пространстве $\mathbb{R}$ над полем $\mathbb{Q}$) сопоставляет его α - координату (коэффициент при α в единственном выражении *x* через векторы базиса). Эта функция линейна над $\mathbb{Q}$, поэтому $f(x+y) = f(x) + f(y)$ для всех $x, y \in \mathbb{R}$. Она отлична от нуля ($f(\alpha) = 1$) и принимает лишь рациональные значения, поэтому не может быть умножением на константу.

Всякая функция, обладающая указанными в теореме 17 свойствами, не ограничена ни на каком отрезке и, более того, ее график всюду плотен в $\mathbb{R}^2$.

Теорема 18. Аддитивные группы $\mathbb{R}$ и $\mathbb{R} \oplus \mathbb{R}$ изоморфны друг другу.

Доказательство. Рассмотрим $\mathbb{R}$ как векторное пространство над $\mathbb{Q}$ и выберем базис в этом пространстве. Очевидно, он бесконечен. Базис в $\mathbb{R} \oplus \mathbb{R}$ может быть составлен из двух частей, каждая из которых представляет собой базис в одном из экземпляров $\mathbb{R}$. Как мы увидим чуть позже, для любого бесконечного множества *B* удвоенная мощность *B* (мощность объединения двух непересекающихся множеств, равномощных *B*) равна мощности *B*. Наконец, осталось заметить, что пространства над одним и тем же полем с равномощными базисами изоморфны как векторные пространства и тем более как группы.

Любой базис в пространстве $\mathbb{R}$ над полем $\mathbb{Q}$ имеет мощность континума. (При доказательстве пригодятся результаты изложенные позже.)

Мы видели, что трансфинитная индукция позволяет доказать существование базиса в любом векторном пространстве. Продолжая эту линию, можно доказать, что любые два базиса векторного пространства равномощны. (Таким образом, понятие размерности как мощности базиса корректно определено и для бесконечномерных векторных пространств.) Мы вернемся к этому позже (теорема 26).

Отметим, что существование базиса Гамеля можно использовать и quot "в мирных целях", а не только для построения экзотических примеров. Известная " третья проблема Гильберта " состояла в доказательстве того, что многогранники равного объема могут не быть равносоставлены. (Это значит, что один из них нельзя разрезать на меньшие многогранники и сложить из них другой многогранник.) Для многоугольников на плоскости ситуация иная: если два многоугольника равновелики (имеют равную площадь), то они равносоставлены.

Теорема 19. Куб нельзя разрезать на части, из которых можно было бы составить правильный тетраэдр (независимо от объема последнего).

Доказательство. Введем понятие псевдообъема многогранника. Как и объем, псевдообъем будет аддитивен (если многогранник разбит на части, сумма их псевдообъемов равна псевдообъему исходного многогранника); псевдообъемы равных многогранников будут равны. Отсюда следует, что псевдообъемы равносоставленных многогранников будут равны. Мы подберем псевдообъем так, чтобы у куба он равнялся нулю, а у тетраэдра нет - и доказательство будет завершено.

Псевдообъем многогранника мы определим как сумму $\sum l_i\varphi(\alpha_i)$, где сумма берется по всем ребрам многогранника, l_i - длина i-го ребра, α_i - двугранный угол при этом ребре, а φ - некоторая функция. Такое определение автоматически гарантирует, что равные многогранники имеют равные псевдообъемы. Что нужно от функции φ, чтобы псевдообъем был аддитивен? Представим себе, что многогранник разрезается плоскостью на две части, и плоскость проходит через уже имеющееся ребро длины *l*. Тогда двугранный угол α при этом ребре разбивается на две части β и γ. Поэтому в выражении для псевдообъема вместо слагаемого *l*φ(α) появляются слагаемые $l\varphi(\beta)+l\varphi(\gamma)$, и φ(α) должно равняться $\varphi(\beta)+\varphi(\gamma)$. Кроме того, разрезающая плоскость может образовать новое ребро, пересекшись с какой-то гранью. Обозначим длину этого ребра за *l'*. Тогда в псевдообъеме появятся слагаемые $l'\varphi(\alpha)+l'\varphi(\pi-\alpha)$ (два образовавшихся двугранных угла дополнительны), которые в сумме должны равняться нулю.

Теперь ясно, какими свойствами должна обладать функция φ. Нужно, чтобы $\varphi(\beta+\gamma)=\varphi(\beta)+\varphi(\gamma)$ и чтобы φ(π)=0. Тогда псевдообъем будет и впрямь аддитивен. Аккуратная проверка требует точного определения понятия многогранника (что не так и просто), и мы ее проводить не будем. Наглядно аддитивность кажется очевидной, особенно если учесть, что все разрезы можно проводить плоскостями (при этом могут получиться более мелкие части, но это не страшно).

Итак, для завершения рассуждения достаточно построить функцию $\varphi\colon \mathbb{R}\to\mathbb{R}$, для которой

- $\varphi(\beta+\gamma)=\varphi(\beta)+\varphi(\gamma)$ для всех $\beta,\gamma\in\mathbb{R}$;
- $\varphi(\pi)=0$(это свойство вместе с предыдущим гарантирует аддитивность псевдообъема);
- $\varphi(\pi/2)=0$(псевдообъем куба равен нулю; это свойство, впрочем, легко следует из двух предыдущих);
- $\varphi(\theta)\neq 0$, где θ - двугранный угол при ребре правильного тетраэдра.

Существенно здесь то, что отношение θ/π иррационально. Проверим это. Высоты двух соседних граней, опущенные на общее ребро, образуют равнобедренный треугольник со сторонами √3, √3, 2; надо доказать, что углы этого треугольника несоизмеримы с π. Удобнее рассмотреть не θ, а другой угол треугольника (два других угла треугольника равны); обозначим его β. Это угол прямоугольного треугольника со сторонами 1, √2 и √3, так что $(\cos\beta + i\sin\beta) = (1+\sqrt{-2})/\sqrt{3}$. Если бы угол θ был соизмерим с π, то и β был бы соизмерим, поэтому некоторая степень этого комплексного числа равнялась бы единице. Можно проверить, однако, что это не так, поскольку кольцо чисел вида $m+n\sqrt{-2}$ ($m, n \in \mathbb{Z}$) евклидово и разложение на множители в нем однозначно.

Дальнейшее просто: рассмотрим числа π и θ. Они независимы как элементы векторного пространства $\mathbb{R}$ над $\mathbb{Q}$, дополним их до базиса и рассмотрим $\mathbb{Q}$-линейный функционал $\varphi\colon \mathbb{R} \to \mathbb{Q}$, равный коэффициенту при θ в разложении по этому базису. Очевидно, все требования при этом будут выполнены.

Некоторое усложнение этого рассуждения позволяет обойтись без базиса Гамеля: достаточно определять φ не на всех действительных числах, а только на линейных комбинациях углов, встречающихся при разрезании куба и тетраэдра на части.

4.1.3. Лемма Цорна и свойства операций

Лемма Цорна и ее применения

В современных учебниках редко встречается трансфинитная индукция как таковая: она заменяется ссылкой на так называемую лемму Цорна. Сейчас мы покажем, как это делается, на примере теоремы о существовании базиса в линейном пространстве.

Теорема 20 (лемма Цорна) Пусть *Z* - частично упорядоченное множество, в котором всякая цепь имеет верхнюю границу. Тогда в этом множестве есть максимальный элемент, и, более того, для любого элемента $a \in Z$ существует элемент $b \geq a$, являющийся максимальным в *Z*. (Цепь - это подмножество, любые два элемента которого сравнимы. Верхняя граница цепи - элемент, больший или равный любого элемента цепи.)

Доказательство. Прежде всего отметим, что Z лишь частично упорядочено, поэтому надо различать максимальные и наибольшие элементы. По этой же причине мы вынуждены употреблять грамматически некорректную конструкцию " больший или равный любого (любому?)", поскольку сказать " не меньше любого" (стандартный выход из положения) означало бы изменить смысл.

Доказательство повторяет рассуждения при построении базиса, но в более общей ситуации (теперь у нас не линейно независимые семейства, а произвольные элементы Z).

Пусть дан произвольный элемент a. Предположим, что не существует максимального элемента, большего или равного a. Это значит, что для любого $b \geq a$ найдется $c>b$. Тогда $c>a$ и потому найдется $d>c$ и т.д. Продолжая этот процесс достаточно долго, мы исчерпаем все элементы Z и придем к противоречию.

Проведем рассуждение аккуратно (пока что мы даже не использовали условие леммы, касающееся цепей). Возьмем вполне упорядоченное множество I достаточно большой мощности (большей, чем мощность Z). Построим строго возрастающую функцию $f:I \rightarrow Z$ по трансфинитной рекурсии. Ее значение на минимальном элементе I будет равно a. Предположим, что мы уже знаем все ее значения на всех элементах, меньших некоторого i. В силу монотонности эти значения попарно сравнимы. Поэтому существует их верхняя граница s, которая, в частности, больше или равна a. Возьмем какой-то элемент $t>s$ и положим $f(i)=t$; по построению монотонность сохранится. Тем самым I равномощно части Z, что противоречит его выбору.

В этом рассуждении, формально говоря, есть пробел: мы одновременно определяем функцию по трансфинитной рекурсии и доказываем ее монотонность с помощью трансфинитной индукции. Наше рекурсивное определение имеет смысл, лишь если уже построенная часть функции монотонна. Формально говоря, надо воспользоваться теоремой 9, считая, что следующее значение не определено, если уже построенный участок не монотонен, и получить функцию, определенную на всем I или на начальном отрезке. Если она определена на некотором начальном отрезке, то она монотонна на нем по построению, поэтому следующее значение тоже определено - противоречие.

Как и при построении базиса Гамеля, можно обойтись без множества большей мощности. Вполне упорядочим множество *Z* с помощью теоремы Цермело. Этот порядок никак не связан с исходным порядком на *Z*; мы будем обозначать его символом $\prec$. Построим с помощью трансфинитной рекурсии функцию $f: Z \to Z$ с такими свойствами: (1) $f(z) \geqslant a$ для любого $z \in Z$; (2) *f* монотонна в следующем смысле: если $x \prec y$, то $f(x) \leqslant f(y)$; (3) *f*(*z*) не может быть строго меньше *z* (в смысле исходного порядка ≤) ни при каком *z*.

Делается это так. Значение $f(z_0)$ для $\prec$ - наименьшего элемента z_0 мы положим равным либо *a*, либо z_0 (последнее - если $z_0>a$). Значение *f*(*z*) для остальных *z* есть либо верхняя граница значений *f*(*z′*) при $z' \prec z$ (по предположению индукции множество таких значений линейно упорядочено и потому имеет некоторую верхнюю границу α), либо само *z* (последнее - если *z*>α).

В силу монотонности множество значений функции *f* линейно упорядочено и имеет верхнюю границу. Эта граница (обозначим ее β) больше или равна *a* (которое есть *f*(z_0)) и является искомым максимальным элементом: если β<*z* для некоторого *z*, то $f(z) \leqslant \beta < z$, что противоречит свойству (3).

Теперь повторим доказательство теоремы о базисе, используя лемму Цорна. Пусть *V*- произвольное векторное пространство. Рассмотрим частично упорядоченное множество *Z*, состоящее из линейно независимых подмножеств пространства *V*. Порядок на *Z* задается отношением "быть подмножеством".

Проверим, что условия леммы выполнены. Пусть имеется некоторая цепь, то есть семейство линейно независимых множеств, причем любые два множества этого семейства сравнимы. Объединим все эти множества и покажем, что полученное множество будет линейно независимым (тем самым оно будет верхней границей элементов цепи).

В самом деле, нетривиальная линейная комбинация включает в себя какое-то конечное число векторов, каждый из своего множества. Этих множеств конечное число, и потому среди них есть наибольшее по включению (в конечном линейно упорядоченном множестве есть наибольший элемент). Это наибольшее множество содержит все векторы нетривиальной линейной комбинации, и линейно независимо по предположению, так что наша нетривиальная линейная комбинация отлична от нуля.

Таким образом, можно применить лемму Цорна и заключить, что любое линейно независимое множество векторов содержится в максимальном линейно независимом множестве векторов. К нему уже нельзя добавить ни одного вектора, не создав линейной зависимости, и оно является искомым базисом.

Аналогичным образом можно доказать существование ортогонального базиса в гильбертовом пространстве (там определение базиса другое: разрешаются бесконечные линейные комбинации, понимаемые как суммы рядов) или существование базиса трансцендентности (максимальная алгебраически независимая система элементов в расширении полей).

Мы приведем другой пример применения леммы Цорна, где фигурируют уже известные нам понятия.

Теорема 21. Всякий частичный порядок может быть продолжен до линейного.

Доказательство. Пусть $(X, \leq)$ - частично упорядоченное множество. Теорема утверждает, что существует отношение порядка $\leq'$ на X, продолжающее исходное (это значит, что $x \leqslant y \Rightarrow x \leqslant' y$) и являющееся отношением линейного порядка. (Кстати, отметим, что слово "линейного" в формулировке теоремы нельзя заменить на слово "полного" - например, если исходный порядок линейный, но не полный.)

Готовясь к применению леммы Цорна, рассмотрим частично упорядоченное множество Z, элементами которого будут частичные порядки на X (то есть подмножества множества $X \times X$, обладающие свойствами рефлексивности, транзитивности и антисимметричности), упорядоченные по включению: $\leq_1$ считается меньшим или равным $\leq_2$, если $\leq_2$ продолжает $\leq_1$ (из $x\leq_1 y$ следует $x\leq_2 y$).

Легко проверить, что условие леммы Цорна выполнено: если у нас есть семейство частичных порядков, линейно упорядоченное по включению, то объединение этих порядков является частичным порядком, и этот порядок будет верхней границей семейства. (Проверим, например, что объединение обладает свойством транзитивности. Пусть $x\leq_1 y$ в одном из порядков семейства ($\leq_1$), а $y\leq_2 z$ в другом; один из порядков (например, $\leq_1$) продолжает другой, тогда $x \leqslant_1 y \leqslant_1 z$ и потому $x\leq z$ в объединении. Рефлексивность и антисимметричность проверяются столь же просто.)

Следовательно, по лемме Цорна на множестве X существует максимальный частичный порядок, продолжающий исходный. Обозначим его как $\leq$ (путаницы с исходным порядком не возникнет, так как исходный нам больше не нужен). Нам надо показать, что он будет линейным. Пусть $x, y \in X$ - два несравнимых элемента. Расширим порядок до нового порядка $\leq'$, при котором $x\leq' y$. Этот новый порядок определяется так: $a\leq' b$, если (1) $a\leq b$ или (2) $a\leq x$ и $y\leq b$. Несложно проверить, что $\leq'$ будет частичным порядком. Рефлексивность очевидна. Транзитивность: если $a\leq' b$ и $b\leq' c$, то есть четыре возможности. Если в обоих случаях имеет место случай (1), то $a \leqslant b \leqslant c$ и все очевидно. Если $a\leq' b$ в силу (1), а $b\leq c$ в силу (2), то $a \leqslant b \leqslant x$ и $y\leq c$, так что $a\leq' c$ в силу (2). Аналогично рассматривается и симметричный случай. Наконец, двукратная ссылка на (2) невозможна, так как тогда $(a \leqslant x)$, $(y \leqslant b)$, $(b \leqslant x)$ и $(y \leqslant c)$, и получается, что $y \leqslant b \leqslant x$, а мы предполагали, что x и y не сравнимы. Антисимметричность доказывается аналогично. Таким образом, отношение $\leq'$ будет частичным порядком, строго содержащим $\leq$, что противоречит максимальности.

Любое бинарное отношение без циклов (цикл образуется, если xRx , или $xRyRx$, или $xRyRzRx$ и т.д) может быть продолжено до линейного порядка. (Для конечных множеств поиск такого продолжения обычно называют " топологической сортировкой ".)

Множество на плоскости называется выпуклым, если вместе с любыми двумя точками оно содержит соединяющий их отрезок. Любые два непересекающихся выпуклых множества можно разделить прямой (каждое множество лежит по одну сторону от прямой, возможно, пересекаясь с ней). (Указание. Используя лемму Цорна, можно расширить исходные непересекающиеся множества A и B до взаимно дополнительных выпуклых множеств A' и B'. Затем можно убедиться, что граница между A' и B'представляет собой прямую.)

Свойства операций над мощностями

Теперь мы можем доказать несколько утверждений о мощностях.

Теорема 22. Если A бесконечно, то множество $A \times \mathbb{N}$ равномощно A.

Доказательство. Вполне упорядочим множество A. Мы уже знаем, что всякий элемент множества A однозначно представляется в виде $z+n$, где z - предельный элемент (не имеющий непосредственно предыдущего), а n - натуральное число. Это означает, что A равномощно $B \times \mathbb{N}$, где B - множество предельных элементов. (Тут есть небольшая трудность - последняя группа элементов конечна, если в множестве есть наибольший элемент. Но мы уже знаем, что добавление конечного или счетного множества не меняет мощности, так что этим можно пренебречь.)

Теперь утверждение теоремы очевидно: $A \times \mathbb{N}$ равномощно $(B \times \mathbb{N}) \times \mathbb{N}$, то есть $B \times (\mathbb{N} \times \mathbb{N})$ и тем самым $B \times \mathbb{N}$ (произведение счетных множеств счетно), то есть A.

По теореме Кантора - Бернштейна отсюда следует, что промежуточные мощности (в частности, $|A|+|A|$, а также любое произведение A и конечного множества) совпадают с $|A|$. Еще одно следствие полезно выделить:

Теорема 23. Сумма двух бесконечных мощностей равна их максимуму.

Доказательство. Прежде всего напомним, что любые две мощности сравнимы (теорема 15). Пусть, скажем, $|A|\leq|B|$. Тогда $|B| \leqslant |A| + |B| \leqslant |B| + |B| \leqslant |B| \times \aleph_0 = |B|$ (последнее неравенство - утверждение предыдущей теоремы). Остается воспользоваться теоремой Кантора-Бернштейна и заключить, что $|B| = |A + B|$.

Теперь можно доказать более сильное утверждение.

Теорема 24. Если A бесконечно, то $A \times A$ равномощно A.

Доказательство. Заметим, что для счетного множества (как, впрочем, и для континуума - но это сейчас не важно) мы это уже знаем. Поэтому в A есть подмножество, равномощное своему квадрату.

Рассмотрим семейство всех таких подмножеств вместе с соответствующими биекциями. Элементами этого семейства будут пары $\langle B, f\rangle$, где B - подмножество A, а $f{:}B\to B\times B$ - взаимно однозначное соответствие. Введем на этом семействе частичный порядок: $\langle B_1, f_1\rangle \leqslant \langle B_2, f_2\rangle$, если $B_1 \subset B_2$ и ограничение отображения f_2 на B_1 совпадает с f_1 (рис. 1).

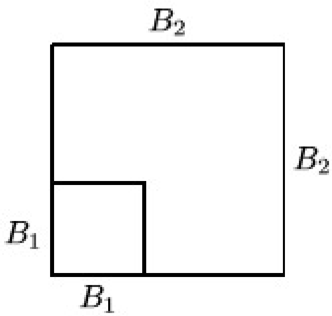


Рис. 1.

Отображение f_1 - взаимно однозначное соответствие между малым квадратом и его стороной; f_2 добавляет к нему взаимно однозначное соответствие между $B_2\backslash B_1$ и " уголком" $(B_2\times B_2\ )\backslash(B_1\times B_1)$. Теперь применим лемму Цорна. Для этого нужно убедиться, что любое линейно упорядоченное (в смысле описанного порядка) множество пар указанного вида имеет верхнюю границу. В самом деле, объединим все первые компоненты этих пар; пусть B - их объединение. Как обычно, согласованность отображений (гарантируемая определением порядка) позволяет соединить отображения в одно. Это отображение (назовем его f) отображает B в $B\times B$. Оно будет инъекцией: значения $f(b')$ и $f(b'')$ при различных b' и b'' различны (возьмем большее из множеств, которым принадлежат b' и b''; на нем f является инъекцией по предположению). С другой стороны, f является сюръекцией: для любой пары $\langle b', b''\rangle \in B\times B$ возьмем множества, из которых произошли b' и b'', выберем из них большее и вспомним, что мы имели взаимно однозначное соответствие между ним и его квадратом.

По лемме Цорна в нашем частично упорядоченном множестве существует максимальный элемент. Пусть этот элемент есть $\langle B, f\rangle$. Мы знаем, что f есть взаимно однозначное соответствие между B и $B\times B$ и потому $|B| = |B| \times |B|$. Теперь есть две возможности. Если B равномощно A, то $B\times B$ равномощно $A\times A$ и все доказано. Осталось рассмотреть случай, когда B не равномощно A, то есть имеет меньшую мощность (большей оно иметь не может, будучи подмножеством). Пусть C - оставшаяся часть A, то есть $A\backslash B$. Тогда $|A| = |B| + |C| = \max(|B|, |C|)$, следовательно, C равномощно A и больше B по мощности. Возьмем в C часть C', равномощную B, и положим $B'=B+C'$ (рис. 2).

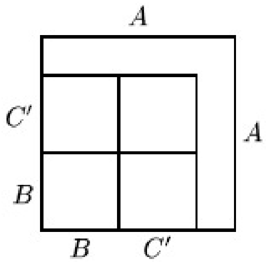


Рис. 2. Продолжение соответствия с B на $B'=B+C'$

Обе части множества B' равномощны B. Поэтому $B'\times B'$ разбивается на 4 части, каждая из которых равномощна $B\times B$, и, следовательно, равномощна B (напомним, что у нас есть взаимно однозначное соответствие f между B и $B\times B$). Соответствие f можно продолжить до соответствия f' между B' и $B'\times B'$, дополнив его соответствием между C' и $(B'\times B')\setminus(B\times B)$ (эта разность состоит из трех множеств, равномощных B, так что равномощна B). В итоге мы получаем большую пару $\langle B', f'\rangle$, что противоречит утверждению леммы Цорна о максимальности. Таким образом, этот случай невозможен.

Выведем теперь некоторые следствия из доказанного утверждения.

Теорема 25. (а) Произведение двух бесконечных мощностей равно большей из них. (б) Если множество A бесконечно, то множество A^n всех последовательностей длины $n>0$, составленных из элементов A, равномощно A. (в) Если множество A бесконечно, то множество всех конечных последовательностей, составленных из элементов A, равномощно A.

Доказательство. Первое утверждение доказывается просто: если $|A| \leqslant |B|$, то

$$|B| \leqslant |A| \times |B| \leqslant |B| \times |B| = |B|.$$

Второе утверждение легко доказывается индукцией по n: если $|A^n| = |A|$, то

$$|A^{n+1}| = |A^n| \times |A| = |A| \times |A| = |A|.$$

Третье тоже просто: множество конечных последовательностей есть $1 + A + A^2 + A^3 + \ldots$; каждая из частей (кроме первой, которой можно пренебречь) равномощна A (по доказанному), и потому все вместе есть $|A| \times \aleph_0 = |A|$.

Заметим, что из последнего утверждения теоремы вытекает, что семейство всех конечных подмножеств бесконечного множества A

имеет ту же мощность, что и A (подмножеств не больше, чем конечных последовательностей и не меньше, чем одноэлементных подмножеств).

Пусть A бесконечно. тогда $|A^A| = |2^A|$.

Рассмотрим мощность $\alpha = \aleph_0 + 2^{\aleph_0} + 2^{(2^{\aleph_0})} + \cdots$ (счетная сумма). Можно показать, что α - минимальная мощность, которая больше мощностей множеств $\mathbb{N}$, $P(\mathbb{N})$, $P(P(\mathbb{N}))$, $\ldots$ Можно показать, что $\alpha^{\aleph_0} = 2^{\alpha} > \alpha$.

Теперь мы можем доказать упоминавшееся ранее утверждение о равномощности базисов.

Теорема 26. Любые два базиса в бесконечномерном векторном пространстве имеют одинаковую мощность.

Доказательство. Пусть даны два базиса - первый и второй. Для каждого вектора из первого базиса фиксируем какой-либо способ выразить его через векторы второго базиса. В этом выражении участвует конечное множество векторов второго базиса. Таким образом, есть некоторая функция, которая каждому вектору первого базиса ставит в соответствие некоторое конечное множество векторов второго. Как мы только что видели, возможных значений этой функции столько же, сколько элементов во втором базисе. Кроме того, прообраз каждого значения состоит из векторов первого базиса, выражающихся через данный (конечный) набор векторов второго, и потому конечен. Выходит, что первый базис разбит на группы, каждая группа конечна, а всего групп не больше, чем векторов во втором базисе. Поэтому мощность первого базиса не превосходит мощности второго, умноженной на $\aleph_0$ (от чего, как мы знаем, мощность бесконечного множества не меняется). Осталось провести симметричное рассуждение и сослаться на теорему Кантора-Бернштейна.

5.Ординалы

Ординалом называется порядковый тип вполне упорядоченного множества.

5.1. Арифметика ординаров

Как мы уже говорили, ординалом называется порядковый тип вполне упорядоченного множества, то есть класс всех изоморфных ему упорядоченных множеств (естественно, они будут вполне упорядоченными).

На ординалах естественно определяется линейный порядок. Чтобы сравнить два ординала α и β, возьмем их представители *A* и *B*. Применим теорему 22 и посмотрим, какой из трех случаев (*A* изоморфно начальному отрезку *B*, отличному от всего *B*; множества *A* и *B* изоморфны; *B* изоморфно начальному отрезку *A*, отличному от всего *A*) имеет место. В первом случае $\alpha < \beta$, во втором $\alpha = \beta$, в третьем $\alpha > \beta$.

Мы отвлекаемся от трудностей, связанных с основаниями теории множеств; как формально можно оправдать наши рассуждения, мы еще обсудим. Пока что отметим некоторые свойства ординалов.

- Мы определили на ординалах линейный порядок. Этот порядок будет полным: любое непустое семейство ординалов имеет наименьший элемент (теорема 13; разница лишь в том, что мы не употребляли там слова "ординал", а говорили о представителях).
- Пусть α - некоторый ординал. Рассмотрим начальный отрезок [0, α) в классе ординалов (образованный всеми ординалами, меньшими α в смысле указанного порядка). Этот отрезок упорядочен по типу α (то есть изоморфен представителям ординала α). В самом деле, пусть *A* - один из представителей ординала α. Ординалы, меньшие α, соответствуют собственным (не совпадающим с *A*) начальным отрезкам множества *A*. Такие отрезки имеют вид [0, α) и тем самым

находятся во взаимно однозначном соответствии с элементами множества *A*. (Легко проверить, что это соответствие сохраняет порядок.)

Сказанное можно переформулировать так: каждый ординал упорядочен как множество меньших ординалов. (В одном из формальных построений теории ординалов каждый ординал равен множеству всех меньших ординалов.)

- Ординал называется непредельным, если существует непосредственно предшествующий ему (в смысле указанного порядка) ординал. Если такого нет, ординал называют предельным.
- Любое ограниченное семейство ординалов имеет точную верхнюю грань (наименьший ординал, больший или равный всем ординалам семейства). В самом деле, возьмем какой- то ординал β, являющийся верхней границей. Тогда все ординалы семейства изоморфны начальным отрезкам множества *B*, представляющего ординал β. Если среди этих отрезков есть само *B*, то β будет точной верхней гранью (и наибольшим элементом семейства). Если нет, то эти отрезки имеют вид [0, *b*) для различных элементов $b \in B$. Рассмотрим множество *S* всех таких элементов *b*. Если *S* не ограничено в *B*, то β будет точной верхней гранью. Если *S* ограничено, то оно имеет точную верхнюю грань *s*, и [0, *s*) будет точной верхней гранью семейства.

Можно сказать, что семейство ординалов - это как бы универсальное вполне упорядоченное семейство; любое вполне упорядоченное множество изоморфно некоторому начальному отрезку этого семейства. Поэтому мы немедленно придем к противоречию, если захотим рассмотреть множество всех ординалов (ведь для всякого вполне упорядоченного множества есть еще большее - добавим к нему новый элемент, больший всех предыдущих). Этот парадокс называется парадоксом Бурали- Форти.

Точная верхняя грань счетного числа счетных ординалов счетна.

Как же рассуждать об ординалах, не впадая в противоречия? В принципе можно заменять утверждения об ординалах утверждениями о их представителях и воспринимать упоминания ординалов как "вольность речи". Другой подход (предложенный фон Нейманом) применяется при аксиоматическом построении теории множеств, и состоит он примерно в следующем: мы объявляем каждый ординал равным множеству всех меньших ординалов. Тогда минимальный ординал 0 (порядковый тип пустого множества) будет пустым множеством Ø, следующий за ним ординал 1 (порядковый тип одноэлементного множества) будет {0}={Ø}, затем $2 = \{0, 1\} = \{\varnothing, \{\varnothing\}\}$, $3 = \{0, 1, 2\} = \{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$, $4 = \{0, 1, 2, 3\}$ и т.д За ними следует ординал ω (порядковый тип множества натуральных чисел), равный $\{0, 1, 2, 3, \ldots\}$, потом $\omega + 1 = \{0, 1, 2, 3, \ldots, \omega\}$, потом $\omega + 2 = \{0, 1, 2, 3, \ldots, \omega, \omega + 1\}$ и т.д

Мы не будем говорить подробно об аксиоматической теории множеств Цермело-Френкеля, но два обстоятельства следует иметь в виду. Во-первых, в ней нет никаких объектов, кроме множеств, и есть аксиома экстенсиональности (или объемности), которая говорит, что два объекта, содержащие одни и те же элементы, равны. Поэтому существует лишь один объект, не содержащий элементов (пустое множество). Во-вторых, в ней есть аксиома фундирования }, которая говорит, что отношение $\in$ фундировано: во всяком множестве *X* есть элемент, являющийся $\in$- минимальным, то есть элемент $x \in X$, для которого $X \cap x = \varnothing$. Отсюда следует, что никакое множество *x* не может быть своим элементом (иначе для множества *{x}* нарушалась бы аксиома фундирования).

Не существует множеств *x*, *y*, *z*, для которых $x \in y \in z \in x$.

Философски настроенный математик объяснил бы смысл аксиомы фундирования так: множества строятся из ранее построенных множеств, начиная с пустого, и поэтому возможна индукция по построению (доказывая какое-либо свойство множеств, можно рассуждать индуктивно и предполагать, что оно верно для всех его элементов).

Теперь можно определить ординалы так. Будем говорить, что множество *x* транзитивно, если всякий элемент множества *x*является подмножеством множества *x*, то есть если из $z \in y \in x$ следует $z \in x$. Назовем ординалом транзитивное множество, всякий элемент которого транзитивен. Это требование гарантирует, что на элементах любого ординала отношение $\in$ является (строгим) частичным порядком.

Аксиома фундирования гарантирует, что частичный порядок $\in$ на любом ординале является фундированным. После этого по индукции можно доказать, что он является линейным (и, следовательно, полным).

(а) Используя определение ординала как транзитивного множества с транзитивными элементами, можно доказать, что элемент ординала есть ординал. (б) Пусть α - ординал (в смысле данного нами определения). Можно доказать, что отношение $\in$ на нем является частичным порядком. (в) Можно доказать, что для любых элементов $a, b \in \alpha$ верно ровно одно из трех соотношений: либо $a \in b$, либо $a = b$, либо $b \in a$. (Указание: используйте двойную индукцию по фундированному отношению $\in$ на α, а также аксиому экстенсиональности.) (г) Можно доказать, что один ординал изоморфен собственному начальному отрезку другого тогда и только тогда, когда является его элементом. (Таким образом, отношение < на ординалах как упорядоченных множествах совпадает с отношением принадлежности.) Можно доказать, что каждый ординал является множеством всех меньших его ординалов.

Заметим еще, что если каждый ординал есть множество всех меньших его ординалов, то точная верхняя грань множества ординалов есть их объединение.

Мы не будем подробно развивать этот подход и по-прежнему будем представлять себе ординалы как порядковые типы вполне упорядоченных множеств.

Прежде чем перейти к сложению и умножению ординалов, отметим такое свойство:

Теорема 27. Пусть *A* - подмножество вполне упорядоченного множества *B*. Тогда порядковый тип множества *A* не превосходит порядкового типа множества *B*.

Доказательство. Отметим сразу же, что равенство возможно, даже если *A* является собственным подмножеством *B*. Например, четные натуральные числа имеют тот же порядковый тип *w*, что и все натуральные числа.

Рассуждая от противного, предположим, что порядковый тип множества *A* больше. Тогда *B* изоморфно некоторому начальному отрезку множества *A*, не совпадающему со всем *A*. Пусть a_0 - верхняя граница (в *A*) этого отрезка, а $f: B \to A$ - соответствующий изоморфизм. Тогда *f* строго возрастает и потому $f(b) \geqslant b$ для всех $b \in B$ (теорема 7). В частности, $f(a_0) \geqslant a_0$, но по предположению любое значение *f*(*b*) меньше a_0 - противоречие.

Арифметика ординалов

Мы ранее определили сумму и произведение линейно упорядоченных множеств. (Напомним, что в *A*+*B* элементы *A* предшествуют элементам *B*, а в *A*×*B* мы сначала сравниваем *B* - компоненты пар, а в случае их равенства- *A* - компоненты.)

Легко проверить следующие свойства сложения:

- Сложение ассоциативно: $\alpha + (\beta + \gamma) = (\alpha + \beta) + \gamma$.
- Сложение не коммутативно: например, 1+*w*=*w*, но *w*+1≠*w*.
- Очевидно, α+0=0+ α= α.
- Сумма возрастает при росте второго аргумента: если $\beta_1 < \beta_2$, то $\alpha + \beta_1 < \alpha + \beta_2$. (В самом деле, пусть β_1 изоморфно начальному отрезку в β_2, отличному от всего β_2. Добавим к этому изоморфизму тождественное отображение на α и получим изоморфизм между α+ β_1 и начальным отрезком в α+ β_2, отличным от α+ β_2.)

- Сумма неубывает при росте первого аргумента: если $\alpha_1 < \alpha_2$, то $\alpha_1 + \beta \leqslant \alpha_2 + \beta$. (В самом деле, α_1+ β изоморфно подмножеству в α_2+ β. Это подмножество не является начальным отрезком, но мы можем воспользоваться теоремой 27.)
- Определение суммы согласовано с обозначением α+1 для следующего за α ординала. (Здесь 1 - порядковый тип одноэлементного множества.) Следующим за α+1 ординалом будет ординал $(\alpha+1)+1=\alpha+(1+1)=\alpha+2$ и т.д.
- Если α≥β, то существует единственный ординал γ, для которого β+γ= α . (В самом деле, β изоморфно начальному отрезку в α; оставшаяся часть α и будет искомым ординалом γ. Единственность следует из монотонности сложения по второму аргументу.) Заметим, что эту операцию можно называть " вычитание слева".
- " Вычитание справа", напротив, возможно не всегда. Пусть α - некоторый ординал. Тогда уравнение β+1=α (относительно β) имеет решение тогда и только тогда, когда α - непредельный ординал, (т.е. когда α имеет наибольший элемент).

Определение суммы двух ординалов в силу ассоциативности можно распространить на любое конечное число ординалов. Можно определить и сумму α_1+ α_2+… счетной последовательности ординалов (элементы α_i предшествуют элементам α_j при $i<j$; внутри каждого α_i порядок прежний). Как легко проверить, это множество действительно будет вполне упорядоченным: чтобы найти минимальный элемент в его подмножестве, рассмотрим компоненты, которые это подмножество задевает, выберем из них компоненту с наименьшим номером и воспользуемся ее полной упорядоченностью.

В этом построении можно заменить натуральные числа на элементы произвольного вполне упорядоченного множества I и определить сумму $\sum A_i$ семейства вполне упорядоченных множеств A_i, индексированного элементами I, как порядковый тип множества всех пар вида $\langle a, i\rangle$, для которых $a \in A_i$. При сравнении пар сравниваются вторые компоненты, а в случае равенства и первые (в соответствующем A_i). Если все A_i изоморфны одному и тому же множеству A, получаем уже известное нам определение произведения $A \times I$.

Теперь перейдем к умножению ординалов.

- Умножение ассоциативно: $(\alpha\beta)\gamma = \alpha(\beta\gamma)$. (В самом деле, в обоих случаях по существу получается множество троек; тройки сравниваются справа налево, пока не обнаружится различие.)
- Умножение не коммутативно: например, $2 \cdot \omega = \omega$, в то время как $\omega \cdot 2 \neq \omega$.
- Очевидно, $\alpha \cdot 0 = 0 \cdot \alpha = 0$ и $\alpha \cdot 1 = 1 \cdot \alpha = \alpha$.
- Выполняется одно из свойств дистрибутивности: $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$ (непосредственно следует из определения). Симметричное свойство выполнено не всегда: $(1 + 1) \cdot \omega = \omega \neq \omega + \omega$.
- Произведение строго возрастает при увеличении второго множителя, если первый не равен 0. (Для разнообразия выведем это из ранее доказанных свойств: если $\beta_2 > \beta_1$, то $\beta_2 = \beta_1 + \delta$, так что $\alpha\beta_2 = \alpha(\beta_1 + \delta) = \alpha\beta_1 + \alpha\delta > \alpha\beta_1$.)
- Произведение не убывает при возрастании первого множителя. (В самом деле, если $\alpha_1 < alpha_2$, то $\alpha_1\beta$ изоморфно подмножеству $\alpha_2\beta$. Это подмножество не является начальным отрезком, но можно сослаться на теорему 27.)
- Любой ординал, меньший $\alpha\beta$, однозначно представим в виде $\alpha\beta' + \alpha'$, где $\beta' < \beta$ и $\alpha' < \alpha$.

(В самом деле, пусть множества A и B упорядочены по типам α и β. Тогда $A \times B$ упорядочено по типу $\alpha\beta$. Всякий ординал, меньший $\alpha\beta$, есть начальный отрезок в $A \times B$, ограниченный некоторым элементом $\langle a, b \rangle$. Начальный отрезок $[0, \langle a, b \rangle)$ состоит из пар, у которых второй член меньше b, а также из пар, у которых второй член равен b, а первый меньше a. Отсюда следует, что этот начальный отрезок изоморфен

$A \times [0, b) + [0, a)$, так что остается положить $\beta' = [0, b)$ и $\alpha' = [0, a)$. Теперь проверим однозначность. Пусть αβ′+ α′= αβ″+ α″. Если β′=β″, то можно воспользоваться однозначностью левого вычитания и получить, что α′ = α″. Остается проверить, что β′ не может быть, скажем, меньше β″. В этом случае β″=β′+δ, и сокращая αβ′ слева, получим, что α′ =αδ+α″, что невозможно, так как левая часть меньше α, а правая часть больше или равна α.)

- Аналогичное " деление с остатком" возможно и для любых ординалов. Пусть α>0. Тогда любой ординал γ можно разделить с остатком на α, то есть представить в виде $\alpha\tau + \rho$, где ρ<α, и притом единственным образом.

  (В самом деле, существование следует из предыдущего утверждения, надо только взять достаточно большое β, чтобы αβ было больше γ, скажем, β=γ+1. Единственность доказывается так же, как и в предыдущем пункте.)

- Повторяя деление с остатком на α>0, можно построить позиционную систему счисления для ординалов: всякий ординал, меньший α^{k+1} (здесь *k* - натуральное число), однозначно представим в виде $\alpha^k\beta_k + \alpha^{k-1}\beta_{k-1} + \ldots + \alpha\beta_1 + \beta_0$, где β_k, $\ldots$, β_1, β_0 - ординалы, меньшие α.

Можно доказать, что α+β=β тогда и только тогда, когда $\alpha\omega \leqslant \beta$ (здесь α и β - ординалы).

Можно доказать, что если α+β=β + α для некоторых ординалов α и β, то найдется такой ординал γ и такие натуральные числа *m* и *n*, что α=γ *m* и β =γ*n*.

Определим операцию "замены основания" с *k*>1на *l*>*k*. Чтобы применить эту операцию к натуральному числу *n*, надо записать *n* в *k*- ичной системе счисления, а затем прочесть эту запись в *l*-ичной системе. (Очевидно, число при этом возрастет, если оно было больше или равно *k*) Возьмем произвольное число *n* и будет выполнять над ним такие операции: замена основания с 2 на 3 - вычитание единицы - замена основания с 3 на 4 - вычитание единицы - замена основания с 4 на 5 - вычитание единицы - ... Можно доказать, что рано или поздно мы получим нуль и вычесть единицу не удастся. (Указание: замените все основания на ординал *w*; получится убывающая последовательность ординалов.)

5.2. Индуктивные определения и степени

Определяется ряд теорем с подробными доказательствами, которые полностью описывают операции сложения и умножения ординалов.

Мы определили сложение и умножение ординалов с помощью явных конструкций порядка на соответствующих множествах. Вместо этого можно было бы их определить индуктивно.

Теорема 28. Сложение ординалов обладает следующими свойствами:

$$\alpha + 0 = \alpha;$$

$$\alpha + (\beta + 1) = (\alpha + \beta) + 1;$$

$$\alpha + \gamma = \sup\{\alpha + \beta \mid \beta < \gamma\} \text{ для предельного } \gamma \neq 0.$$

Эти свойства однозначно определяют операцию сложения.

Доказательство. Два первых свойства очевидны; проверим третье. Если β<γ, то α+β<α+γ, так что α+γ будет верхней границей всех сумм вида α+β при β<γ. Надо проверить, что эта граница точная. Пусть некоторый ординал τ меньше α+γ. Убедимся, что он меньше α+β для некоторого β<γ. Если τ<α, все очевидно. Если τ≥α, представим его в виде τ=α+σ. Тогда α+σ<α+γи потому σ<γ. Поскольку ординал γ предельный, σ+1также меньше γ и остается положить β=σ+1.

Указанные свойства однозначно определяют операцию сложения, так как представляют собой рекурсивное определение по β (если есть две операции сложения, обладающие этими свойствами, возьмем минимальное β, для которого они различаются и т.д).

Аналогично можно определить и умножение:

Теорема 29. Умножение ординалов обладает следующими свойствами:

$$\alpha 0 = 0;$$
$$\alpha(\beta + 1) = \alpha\beta + \alpha;$$
$$\alpha\gamma = \sup\{\alpha\beta \mid \beta < \gamma\} \text{ для предельного } \gamma \neq 0.$$

Эти свойства однозначно определяют операцию умножения.

Доказательство. Доказательство аналогично, нужно только проверить, что если $\tau<\alpha\gamma$ для предельного γ, то $\tau<\alpha\beta$ для некоторого $\beta<\gamma$. Как мы видели ранее, ординал τ имеет вид $\tau=\alpha\gamma'+\alpha'$ при $\gamma'<\gamma$; достаточно положить $\beta=\gamma'+1$.

Возникает естественное желание определить операцию возведения в степень. Мы уже по существу определили возведение в целую положительную степень (α^n есть произведение n сомножителей, равных α). Другими словами, если A упорядочено по типу α, то множество A^n последовательностей длины n с элементами из A с обратным лексикографическим порядком (сравнение справа налево) упорядочено по типу α^n.

Следующий шаг - определить α^w. Первая идея, приходящая в голову- взять множество $A^{\mathbb{N}}$ бесконечных последовательностей и определить на нем полный порядок. Но как его ввести - неясно. Поэтому можно попробовать определить возведение в степень индуктивно с помощью следующих соотношений:

$$\alpha^0 = 1;$$
$$\alpha^{\beta+1} = \alpha^\beta \cdot \alpha;$$
$$\alpha^\gamma = \sup\{\alpha^\beta \mid \beta < \gamma\} \text{ для предельного } \gamma \neq 0.$$

Теорема 8 (о трансфинитной рекурсии) гарантирует, что эти соотношения однозначно определяют некоторую операцию над ординалами, которая и называется возведением в степень.

Замечание. Тут опять мы подходим к опасной границе парадоксов и вынуждены выражаться уклончиво. На самом деле теорема о трансфинитной рекурсии говорила об определении функции на вполне упорядоченном множестве, а ординалы не образуют множества - их слишком много. Кроме того, в ней шла речь о функциях со значениями в некотором заданном множестве, которого здесь тоже нет. Подобные индуктивные определения можно корректно обосновать в теории множеств с использованием так называемой аксиомы подстановки, но мы об этом говорить не будем. Вместо этого мы дадим явное описание возведения в степень, свободное от этих проблем.

Чтобы понять смысл возведения в степень, посмотрим, как выглядит ординал α^w (для некоторого α). Пусть A - множество, упорядоченное по типу α. Ординал α^w по определению есть точная верхняя грань α^n для натуральных n. Ординал α^n есть порядковый тип множества A^n, упорядоченного в обратном лексикографическом порядке. Чтобы найти точную верхнюю грань, представим множества A^n как начальные отрезки друг друга. Например, A^2 состоит из пар $\langle a_1, a_2 \rangle$ и отождествляется с начальным отрезком в A^3, состоящим из троек $\langle a_1, a_2, 0 \rangle$. (Здесь 0 - наименьший элемент в A.) Теперь видно, что все множества A^n можно рассматривать как начальные отрезки множества A^∞, состоящего из бесконечных последовательностей $a_0, a_1, \ldots$, элементы которых принадлежат A и в которых лишь конечное число членов отлично от нуля. (Последнее требование делает корректным определение обратного лексикографического порядка - мы находим самую правую позицию, в которой последовательности различаются, и сравниваем их значения в этой позиции.) В объединении эти начальные отрезки дают все A^∞, так что это множество с описанным порядком имеет тип α^w.

Аналогичное утверждение верно и для любого показателя степени.

Пусть A и B - вполне упорядоченные множества, имеющие порядковые типы α и β. Рассмотрим множество $[B \to A]$ состоящее из отображений B в A, имеющих "конечный носитель" (равных минимальному элементу A всюду, за исключением конечного множества). Введем на $[B \to A]$ порядок: если $f_1 \neq f_2$, выберем наибольший элемент $b \in B$, для которого $f_1(b) \neq f_2(b)$ и сравним $f_1(b)$ и $f_1(b)$.

Теорема 30. Указанное правило задает полный порядок на множестве [$B \to A$] и порядковый тип этого множества есть α^{β}.

Доказательство. Нам надо проверить, что указанный порядок является полным и что выполнены требования индуктивного определения степени. Назовем носителем элемента $f \in [B \to A]$ множество тех $b \in B$, для которых $f(b)>0$ (здесь 0 обозначает наименьший элемент множества A). Назовем рангом функции f наибольший элемент носителя (по определению носитель конечен, так что наибольший элемент существует). Ранг определен для всех функций, кроме тождественно нулевой, которая является минимальным элементом множества [$B \to A$]. Чем больше ранг функции, тем больше сама функция в смысле введенного нами порядка.

Пусть порядок на [$B \to A$] не является полным и $f_0>f_1>f_2>\ldots$ - убывающая последовательность элементов [$B \to A$]. Все элементы $f_i(b)$ отличны от 0; рассмотрим их ранги. Эти ранги образуют невозрастающую последовательность, поэтому начиная с некоторого места стабилизируются (множество B вполне упорядочено). Отбросим начальный отрезок и будем считать, что с самого начала ранги всех элементов убывающей последовательности одинаковы и равны некоторому b. В соответствии с определением, значения $f_0(b)$, $f_1(b)$,… образуют невозрастающую последовательность, поэтому начиная с некоторого места стабилизируются. Отбросив начальный отрезок, будем считать, что все f_i имеют одинаковый ранг b и одинаковое значение $f_i(b)$. Тогда значения $f_i(b)$ не влияют на сравнения, и потому их можно заменить на 0. Получим убывающую последовательность элементов [$B \to A$] с рангами меньше b. Чтобы завершить рассуждение, остается сослаться на принцип индукции по множеству B.

(Более формально, рассмотрим все бесконечно убывающие последовательности. У каждой из них рассмотрим ранг первого элемента. Рассмотрим те из них, у которых этот ранг минимально возможный; пусть b - это минимальное значение. В любой такой последовательности все элементы имеют ранг b. Из всех таких последовательностей $f_0> f_1>\ldots$выберем ту, у которой значение $f_0(b)$ минимально; все следующие ее члены имеют то же значение в точке b (т.е. $f_i(b)= f_0(b)$). Заменив значение в точке b нулем, получим бесконечную убывающую последовательность из элементов меньшего ранга, что противоречит предположению.)

Теперь покажем, что такое явное определение степени согласовано с индуктивным определением. Для конечных n это очевидно. Пусть $\gamma=\beta+1$. Каково (явное) определение α^{γ}? Пусть B упорядочено по типу β. Тогда мы должны добавить к B новый наибольший элемент (обозначим его m) и рассмотреть отображения $B \cup \{m\} \to A$ с конечным носителем. Ясно, что такое отображение задается парой, состоящей из его сужения на B (которое может быть произвольным элементом множества $[B\to A]$) и значения на m. При определении порядка мы сначала сравниваем значения на m, а потом сужения на B, то есть полученное множество изоморфно $[B\to A]\times A$, что и требовалось.

Пусть теперь γ - ненулевой предельный ординал и множество C упорядочено по типу γ. Как устроено множество $[C\to A]$? Элементы, ранг которых меньше $c \in C$, образуют в нем начальный отрезок, и этот начальный отрезок изоморфен $[[0, c) \to A]$. А само множество $[C\to A]$ является объединением этих начальных отрезков (поскольку каждый элемент этого множества имеет конечный носитель) и потому его порядковый тип является точной верхней гранью их порядковых типов, что и требовалось.

Непосредственным следствием этой теоремы является такое утверждение:

Теорема 31. Если α и β - счетные ординалы, то $\alpha+\beta$, $\alpha\beta$, и α^{β} счетны.

Доказательство. Для суммы и произведения утверждение очевидно. Для степени: если мы пронумеровали все элементы вполне упорядоченных множеств A и B, то любой элемент множества $[B\to A]$ может быть задан конечным списком натуральных чисел (носитель и значения на элементах носителя), а таких списков счетное число.

Отметим важную разницу между операцией возведения ординалов в степень и ранее рассмотренными операциями сложения и умножения ординалов. Определяя сумму и произведение ординалов, мы вводили некоторый порядок на сумме и произведении соответствующих множеств (в обычном смысле), здесь же само множество $[B\to A]$ определяется с учетом порядка и отлично от A^B. (В частности, при счетных A и B множество $[B\to A]$ счетно, а A^B - нет.)

Явное описание множества $[B \to A]$ позволяет понять, как устроены его начальные отрезки, то есть какой вид имеют ординалы, меньшие α^β.

Рассмотрим некоторую функцию $f \in [B \to A]$. Пусть она отлична от нуля в точках $b_1 > b_2 > \ldots > b_k$ и принимает там значения a_1, a_2, $\cdots$, a_k. Нас интересуют все функции, меньшие функции f.

Все они равны нулю в точках, больших b_1. В самой точке b_1 они могут быть либо меньше a_1, либо равны a_1. Любая функция первого типа меньше любой функции второго типа. Функции первого типа могут принимать любые значения в точках, меньших b_1, а в точке b_1имеют значение из $[0, a_1)$. Тем самым их можно отождествить с элементами множества $[[0, b_1) \to A] \times [0, a_1)$, и при этом отождествлении сохраняется порядок.

Функции второго типа (равные a_1 в точке b_1) снова разбиваются на две категории: те, которые в точке b_2 меньше a_2 и те, которые в b_2 равны a_2. Функции первой категории отождествляются с элементами множества $[[0, b_2) \to A] \times [0, a_2)$. Функции второй категории снова разобьем на части в зависимости от их значения в точке b_3 и т.д. Таким образом, $[0, f)$ как упорядоченное множество изоморфно множеству

$$[[0, b_1) \to A] \times [0, a_1) + [[0, b_2) \to A] \times [0, a_2) + \ldots + \\ + [[0, b_k) \to A] \times [0, a_k).$$

Переходя к ординалам (начальные отрезки - это меньшие ординалы), получаем такое утверждение:

Теорема 32. Всякий ординал, меньший α^β, представляется в виде

$$\alpha^{\beta_1}\alpha_1 + \alpha^{\beta_2}\alpha_2 + \ldots + \alpha^{\beta_k}\alpha_k$$

, где $\beta > \beta_1 > \beta_2 > \ldots > \beta_k$, а $\alpha_1, \alpha_2, \ldots, \alpha_k < \alpha$. Такое представление однозначно и любая сумма указанного вида является ординалом, меньшим α^β.

Доказательство. Возможность такого представления мы уже доказали. Последнее утверждение следует из того, что любая сумма такого вида является начальным отрезком в множестве [*B*→*A*] (где *A* и *B* упорядочены по типам α и β) и разным суммам соответствуют разные начальные отрезки.

Это утверждение обобщает описанную нами ранее "позиционную систему обозначений с основанием α" для ординалов, меньших α^k; теперь вместо *k* можно использовать любой ординал.

Можно было бы сразу сказать, что элементами множества [*B*→*A*] являются формальные суммы вида

$$\alpha^{\beta_1}\alpha_1 + \alpha^{\beta_2}\alpha_2 + \ldots + \alpha^{\beta_k}\alpha_k$$

(где $\beta > \beta_1 > \ldots > \beta_k$ и $\alpha_1, \ldots, \alpha_k < \alpha$) с естественным порядком на них.

Теперь уже понятно, как устроены ординалы в последовательности

$$\omega^\omega, \omega^{(\omega^\omega)}, \ldots$$

Первый из них образован " одноэтажными" выражениями вида

$$\omega^{b_1}a_1 + \omega^{b_2}a_2 + \ldots + \omega^{b_k}a_k,$$

где a_i и b_i - натуральные числа (и $b_1 > \ldots > b_k$). Если в качестве $b_1, \ldots, b_k$ разрешить писать любые " одноэтажные" выражения указанного вида, то полученные " двухэтажные" выражения упорядочены по типу $\omega^{(\omega^\omega)}$. Разрешив в показателях двухэтажные выражения, мы получим трехэтажные выражения, которые образуют следующий ординал и т.д. Если объединить все эти множества, то есть не ограничивать число этажей (которое для каждого выражения тем не менее конечно), то получится множество, упорядоченное по типу

$$\sup(\omega, \omega^\omega, \omega^{(\omega^\omega)}, \ldots)$$

Этот ординал обозначается ε_0.

Определим для натуральных чисел операцию " тотальной замены основания *k* на *l*" (здесь *k* и *l* - натуральные числа, причем *l*>*k*) следующим образом: данное число *n* запишем в *k* - ичной системе, то есть разложим по степеням *k*, показатели степеней снова запишем в *k*-

ичной системе, новые показатели также разложим и т.д. Затем на всех уровнях заменим основание k на основание l и вычислим значение получившегося выражения. Можно доказать, что начав с любого n и выполняя последовательность операций " вычитание единицы - тотальная замена основания 2 на 3 - вычитание единицы - тотальная замена основания 3 на 4 - вычитание единицы - тотальная замена основания 4 на 5 -, $\dots$", мы рано или поздно зайдем в тупик,т.е. получится нуль и вычесть единицу будет нельзя. (Указание: заменим все основания сразу на ординал w; получится убывающая последовательность ординалов, меньших ε_0.)

5.3. Приложения ординалов

В большинстве случаев рассуждения с использованием трансфинитной индукции и ординалов можно заменить ссылкой на лемму Цорна; при этом рассуждение становится менее наглядным, но формально более простым. Тем не менее бывают ситуации, когда этого сделать не удается (по крайней мере, неясно, как бы это следовало делать), и приходится пользоваться вполне упорядоченными множествами в явном виде. В этом разделе мы приведем два подобных примера.

Первый из них касается борелевских множеств. (Для простоты мы рассматриваем подмножества действительной прямой.) Семейство подмножеств действительной прямой называется σ-алгеброй, если оно замкнуто относительно конечных и счетных пересечений и объединений, а также относительно перехода к дополнению. (Это означает, что вместе с каждым множеством A это семейство содержит его дополнение $\mathbb{R} \setminus A$, и вместе с любыми множествами A_0, A_1, $\dots$ семейство содержит их объединение $A_0 \cup A_1 \cup \dots$и пересечение $A_0 \cap A_1 \cap \dots$) Пример: семейство $P(\mathbb{R})$ всех подмножеств прямой, очевидно, является σ - алгеброй.

Теорема 33. Существует минимальная σ-алгебра, содержащая все отрезки $[a, b]$ на прямой.

Доказательство. Формально можно рассуждать так: рассмотрим все возможные σ - алгебры, содержащие отрезки. Их пересечение будет σ - алгеброй, и тоже будет содержать все отрезки. (Вообще пересечение любого семейства σ - алгебр будет σ - алгеброй - это очевидное следствие определения.) Эта σ - алгебра и будет искомой.

Множества, входящие в эту минимальную σ - алгебру, называют борелевскими.

Всякое открытое и всякое замкнутое подмножество прямой является борелевским. (Указание: открытое множество есть объединение содержащихся в нем отрезков с рациональными концами.)

Прообраз любого борелевского множества при непрерывном отображении является борелевским множеством.

Пусть f_0, f_1, $\ldots$ - последовательность непрерывных функций с действительными аргументами и значениями. Можно доказать, что множество точек *x*, в которых последовательность $f_0(x)$, $f_1(x)$, $\ldots$ имеет предел, является борелевским.

Борелевские множества играют важную роль в дескриптивной теории множеств. Но мы хотим лишь продемонстрировать использование трансфинитной индукции (вряд ли легко заменяемой на использование леммы Цорна) на примере следующей теоремы:

Теорема 34. Семейство всех борелевских множеств имеет мощность континуума.

Доказательство. Класс борелевских множеств можно строить постепенно. Начнем с отрезков и дополнений к отрезкам. На следующем шаге рассмотрим всевозможные счетные пересечения и объединения уже построенных множеств (отрезков и дополнений к ним). При этом получатся (среди прочего) все открытые и все замкнутые подмножества прямой.

Далее можно рассмотреть счетные объединения и пересечения уже построенных множеств и т.д.

Более формально, пусть $\mathcal{B}_0$- семейство множеств, состоящее из всех отрезков и дополнений к ним. Определим $\mathcal{B}_{i+1}$по индукции как семейство множеств, являющихся счетными объединениями или пересечениями множеств из $\mathcal{B}_i$.

Все семейства $\mathcal{B}_i$ состоят из борелевских множеств (поскольку счетное объединение или пересечение борелевских множеств является борелевским). Исчерпывают ли они все борелевские множества? Вообще говоря, нет: если мы возьмем по одному множеству из каждого класса $\mathcal{B}_i$ для всех i=0, 1, 2,…, и рассмотрим их счетное пересечение, то оно вполне может не принадлежать ни одному из классов. Поэтому мы рассмотрим класс $\mathcal{B}_\omega$, представляющий собой объединение всех $\mathcal{B}_i$ по всем натуральным i, затем $\mathcal{B}_{\omega+1}$, $\mathcal{B}_{\omega+2}$ и т.д. Объединение этой последовательности классов естественно назвать $\mathcal{B}_{\omega 2}$ и продолжить построение.

Дадим формальное определение $\mathcal{B}_\alpha$ для любого ординала α. Это делается с помощью трансфинитной рекурсии. Именно, при α=β+1 элементами класса $\mathcal{B}_\alpha$ будут счетные объединения и пересечения множеств из класса $\mathcal{B}_\beta$. Если α - предельный ординал, отличный от 0, то класс $\mathcal{B}_\alpha$ представляет собой объединение всех $\mathcal{B}_\beta$ по всем β<α. (Класс $\mathcal{B}_0$ мы уже определили.)

Из определения следует, что $\mathcal{B}_\alpha \subset \mathcal{B}_\beta$ при α<β, так что мы получаем возрастающую цепь классов. Каждый класс замкнут относительно перехода к дополнению (для начального класса мы об этом позаботились, далее по индукции). Все классы $\mathcal{B}_\alpha$ содержатся в классе борелевских множеств, так как мы применяем лишь операции счетного объединения и пересечения, относительно которых класс борелевских множеств замкнут.

Возникает вопрос: как далеко нужно продолжать эту конструкцию? Оказывается, что достаточно дойти до первого несчетного ординала.

Пусть $\aleph_1$ - наименьший несчетный ординал. (Это - стандартное для него обозначение.) Другими словами, $\aleph_1$ есть семейство всех счетных ординалов, упорядоченных отношением < на ординалах.

Лемма. Класс $\mathcal{B}_{\aleph_1}$ замкнут относительно счетных объединений и пересечений и потому содержит все борелевские множества.

Доказательство леммы. Пусть имеется счетная последовательность множеств $B_0, B_1, \ldots$, принадлежащих $\mathcal{B}_{\aleph_1}$. Ординал $\aleph_1$ - предельный, и класс $\mathcal{B}_{\aleph_1}$ является объединением меньших классов. Поэтому каждое из множеств B_i принадлежит какому-то классу $\mathcal{B}_{\alpha_i}$, где α_i - некоторый ординал, меньший $\aleph_1$, -е конечный или счетный ординал. Положим $\beta = \sup_i \alpha_i$. Ординал β есть точная верхняя грань счетного числа счетных ординалов и потому счетен. В самом деле, рассмотрим ординалы α_i как начальные отрезки в каком-то большем ординале (например, в $\aleph_1$); их точная верхняя грань будет объединением счетного числа счетных начальных отрезков и потому будет счетным ординалом.

Теперь первое утверждение леммы очевидно: все B_i лежат в $\mathcal{B}_\beta$, а потому их объединение (или пересечение) лежит в $\mathcal{B}_{\beta+1}$ и тем более в $\mathcal{B}_{\aleph_1}$ (поскольку β+1 есть счетный ординал и меньше $\aleph_1$).

Таким образом, класс $\mathcal{B}_{\aleph_1}$ является σ - алгеброй, содержащей отрезки, и потому содержит все борелевские множества. Лемма доказана.

Как мы уже отмечали, все классы $\mathcal{B}_\alpha$ состоят из борелевских множеств, так что класс $\mathcal{B}_{\aleph_1}$ совпадает с классом всех борелевских множеств.

Что можно сказать про мощность классов? Класс $\mathcal{B}_0$ имеет мощность континуума (отрезки задаются своими концами). Если класс $\mathcal{B}_\alpha$ имеет мощность континуума, то и следующий класс $\mathcal{B}_{\alpha+1}$ имеет мощность континуума (каждый его элемент задается счетной последовательностью элементов предшествующего класса, а $\mathfrak{c}^{\aleph_0} = \mathfrak{c}$). Каждый предельный класс есть объединение предыдущих, и пока мы не выходим за пределы счетных ординалов, объединение это будет счетно, а $\mathfrak{c}\aleph_0 = \mathfrak{c}$, так что мы не выходим за пределы континуума. Наконец, $\mathcal{B}_{\aleph_1}$ есть объединение несчетного числа предыдущих классов (а именно, $\aleph_1$ классов), но так как $\aleph_1 \leqslant \mathfrak{c}$, то $\mathfrak{c}\aleph_1 = \mathfrak{c}$.

Таким образом, класс $\mathcal{B}_{\aleph_1}$, он же класс всех борелевских множеств,

имеет мощность континуума.
Обычно построение борелевских множеств начинается немного иначе. Именно, на нижнем уровне рассматриваются два класса: открытые и замкнутые множества. На следующем уровне находятся классы F_σ (счетные объединения замкнутых множеств) и G_δ (счетные пересечения открытых множеств). Еще на уровень выше лежат счетные пересечения множеств из F_σ и счетные объединения множеств из G_δ, и т.д. Такой подход является более естественным с точки зрения топологии, поскольку отрезки на прямой ничем не замечательны. Можно проверить, что разница между таким подходом и нашим определением невелика.

(Докажите, что пересечение двух F_σ- множеств является F_σ-множеством (и вообще классы F_σ, G_δ, а также классы следующих уровней, замкнуты относительно конечных объединений и пересечений).)

Можно доказать, что F_σ- и G_δ- множества лежат в классе B_2 в соответствии с нашей классификацией.

Можно доказать, что всякое множество класса B_2 отличается от некоторого F_σ- или G_δ- множества не более чем на счетное множество.

Можно доказать, что всякое множество класса B_3 является счетным пересечением F_σ- множеств или счетным объединением G_δ-множеств и что аналогичное утверждение верно для более высоких уровней нашей иерархии.
Можно доказать, что существует открытое множество на плоскости, среди вертикальных сечений которого встречаются все открытые подмножества прямой. Можно доказать что существует G_δ-множество на плоскости, среди сечений которого встречаются все G_δ - подмножества прямой. Можно доказать аналогичные утверждения для следующих уровней.

Существует G_δ- множество, не являющееся F_σ- множеством. Существует счетное объединение G_δ- множеств, не являющееся счетным пересечением F_σ- множеств и т.д. (Указание: воспользуйтесь предыдущей задачей.)
Ординалы часто появляются при классификации элементов того или иного множества по " рангам". Например, можно классифицировать

элементы фундированного множества.

Теорема 35. Пусть X- фундированное множество. Тогда существует и единственна функция rk, определенная на X и принимающая значения в классе ординалов, для которой

$$\mathrm{rk}(x) = \min\{\alpha \mid \alpha > \mathrm{rk}(y) \text{ для любого } y < x\}$$

(при любом $x \in X$).

Доказательство. Определим множество X_α рекурсией по ординалу α: X_α состоит из всех элементов $x \in X$, для которых все меньшие их (в X) элементы принадлежат X_β с меньшими индексами β:

$$x \in X_\alpha \;\Leftrightarrow\; (\forall y < x)\,(\exists \beta < \alpha)\,(y \in X_\beta).$$

Заметим, что здесь (как и в формулировке теоремы) знак < используется в двух разных смыслах: как порядок на X и как порядок на ординалах.

Очевидно, что с ростом α множество X_α растет (точнее, не убывает). Докажем, что при достаточно большом α множество X_α покрывает все X. Если это не так, то из $\beta<\gamma$ следует $X_\beta \subsetneq X_\gamma$ (произвольный минимальный элемент, не лежащий в X_β, принадлежит X_γ). Поэтому отображение $\alpha \mapsto X_\alpha$ будет инъекцией, что невозможно (возьмем ординал, по мощности больший $P(X)$; предшествующих ему ординалов уже слишком много).

Теперь определим rk(x) как минимальное α, при котором $x \in X_\alpha$. Если rk(x)=α и $y<x$, то rk(y)<α . (В самом деле, по определению X_α из $x \in X_\alpha$ и $y<x$ следует, что $y \in X_\beta$ при некотором $\beta<\alpha$.) Наоборот, если для некоторого ординала γвыполнено неравенство rk(y)<γ при всех $y<x$, то rk(x)$\leq\gamma$. В самом деле, тогда любой элемент $y<x$ принадлежит некоторому X_β с $\beta<\gamma$ (положим β=rk(y)) и потому $x \in X_\gamma$ и rk(x)$\leq\gamma$.

Итак, построенная нами функция rk обладает требуемым свойством. Единственность доказать совсем легко: если есть две такие функции, рассмотрим минимальную точку в X, на которой они различаются, и сразу же получим противоречие.

В частности, счетные ординалы можно использовать для классификации деревьев, в которых нет бесконечных путей. Мы будем рассматривать корневые деревья с конечным или счетным ветвлением (у каждой вершины конечное или счетное число сыновей), в которых нет бесконечной ветви (последовательности вершин, в которых каждая есть сын предыдущей).

Формально такое дерево можно определить как подмножество *T* множества $\mathbb{N}^*$ конечных последовательностей натуральных чисел, замкнутое относительно взятия префикса (если последовательность принадлежит *T*, то любой ее начальный отрезок принадлежит *T*). Элементы множества *T* мы называем вершинами дерева; вершина *y* есть сын вершины *x*, если *y* получается из *x* приписыванием справа какого-то одного числа. Вершина *y* является потомком вершины *x*, если *y* получается добавлением к *x* одного или нескольких чисел.

Мы говорим, что в дереве *T* нет бесконечной ветви, если не существует бесконечной последовательности натуральных чисел, все начала которой принадлежат *T*. В этом случае отношение порядка

$$y < x \;\Leftrightarrow\; y \text{ есть потомок } x$$

фундировано и можно применить предыдущую теорему, определив ранги всех вершин дерева *T*. Ранг его корня (последовательности длины 0) и будем называть рангом дерева.

Теорема 36.

(а) Ранг любого дерева (описанного вида) является счетным ординалом.

(б) Всякий счетный ординал является рангом некоторого дерева.

Доказательство. (а) Пусть ранг некоторого дерева, то есть ранг его корня, является несчетным ординалом. Тогда ранг одного из сыновей корня также несчетен. (В самом деле, точная верхняя грань счетного множества счетных ординалов является счетным ординалом; это становится ясным, если рассматривать эти ординалы как начальные отрезки большего - тогда точная верхняя грань будет объединением.) У этого сына в свою очередь есть сын несчетного ранга и т.д. Этот процесс не может оборваться, и мы получаем бесконечную ветвь в противоречии с предположением.

(б) Это утверждение доказывается индукцией: пусть α - наименьший счетный ординал, для которого такого дерева нет. Тогда для всех меньших ординалов деревья есть. Возьмем эти деревья и сделаем их поддеревьями с общим корнем (их корни станут сыновьями этого общего корня). Новое дерево также имеет счетное ветвление и ранг его корня равен α.

Пусть имеется счетное дерево, не имеющее бесконечных ветвей. Предположим, что в каждом его листе находится отрезок или дополнение до отрезка, а в каждой внутренней вершине стоит знак пересечения или объединения. Как сопоставить такому дереву некоторое борелевское множество? (Указание: покажите, что в каждой вершине можно единственным образом написать некоторое множество, согласованное с пометками.) Покажите, что все борелевские множества могут быть получены таким способом.

Деревья с пометками, рассмотренные в этой задаче, представляют собой как бы бесконечные формулы, составленные из отрезков и дополнений к ним с помощью операций счетного объединения и пересечения (конечные деревья соответствовали бы конечным формулам). Можно условно сказать, что борелевские множества - это множества, выражающиеся с помощью таких формул.

Можно доказать, что семейство борелевских множеств имеет мощность континнума, используя " бесконечные формулы" - размеченные деревья, в которых нет бесконечных ветвей. (Это доказательство обходится без ординалов, трансфинитной индукции и даже леммы Цорна - хотя и использует аксиому выбора.)

В заключение приведем скорее забавный, чем важный, пример использования трансфинитной рекурсии и ординалов.

Теорема 37. Существует множество точек на плоскости, которое пересекается с каждой прямой ровно в двух точках.

Две параллельные прямые почти что удовлетворяют этому требованию (исключением являются лишь параллельные им прямые). Но избавиться от этого исключения не так просто.

Доказательство. Требования к множеству можно сформулировать так: никакие три точки не лежат на одной прямой, но любая прямая пересекает его не менее чем в двух точках.

Будем строить это множество трансфинитной рекурсией. Пусть α - минимальный ординал, имеющий мощность континуума. (Если континуум - гипотеза верна, то он совпадает с $\aleph_1$, но это нам не важно.) Тогда множество всех меньших ординалов можно поставить во взаимно однозначное соответствие с множеством всех прямых на плоскости. Пусть l_β - прямая, соответствующая ординалу $\beta<\alpha$.

Для каждого $\beta<\alpha$ построим множество M_β, в котором никакие три точки не лежат на одной прямой, следующим образом. Объединим все построенные ранее множества M_γ при всех $\gamma<\beta$. Могут ли в этом множестве (обозначим его T) какие-то три точки лежать на одной прямой? Если да, то эти точки берутся из каких-то множеств M_{γ_1}, M_{γ_2}, M_{γ_3}; возьмем наибольший из ординалов γ_1, γ_2, γ_3; в соответствующем множестве будут три точки, лежащие на одной прямой, что противоречит предположению индукции.

Посмотрим, во скольких точках пересекает прямая l_β множество T. Таких точек (по доказанному) не больше двух. Если их ровно две, то все хорошо и мы новых точек не добавляем, считая, что $M_\beta=T$. Если их меньше, то мы должны добавить новые точки (до двух), но только так, чтобы при этом не образовалось трех точек, лежащих на одной прямой. Другими словами, нельзя добавлять точки, которые лежат на пересечении l_β с прямыми, проходящими через пары уже имеющихся точек.

Сколько таких прямых (то есть сколько пар уже имеющихся точек)? По построению видно, что все уже имеющиеся точки лежат по две на каждой прямой l_γ при $\gamma<\beta$. (Строго говоря, это следует включить в индуктивное предположение.) Таким образом, множество T по мощности есть $2\beta=\beta$, а пар точек не больше $\beta^2=\beta<c$. Поэтому запрещенные точки составляют лишь малую (по мощности) часть прямой l_β, и можно выбрать две разрешенные точки.

Теперь осталось объединить множества M_β для всех ординалов $\beta<\alpha$ и получить искомое множество. (По условию три точки на одной прямой в нем появиться не могут, а всякая прямая будет рано или поздно рассмотрена и две точки на ней будут обеспечены.)

Найдите ошибку в следующем " опровержении гипотезы континуума": пусть $\aleph_1 = \mathfrak{c}$. Упорядочим отрезок [0, 1] по типу $\aleph_1$. Рассмотрим функцию двух переменных, равную единице на паре (x, y), если $x<y$ (в смысле этого порядка), и нулю в остальных случаях. Тогда при

фиксированном *x*функкия $y \mapsto f(x,y)$ равна единице везде, кроме счетного множества, и потому интегрируема и $\int f(x,y)\,dy = 1$ при любом *x*. С другой стороны, функция $x \mapsto f(x,y)$ равна нулю всюду, кроме счетного множества, так что $\int f(x,y)\,dx = 0$. Получаем противоречие с теоремой Фубини, которая утверждает, что

$$\int_0^1 \left(\int_0^1 f(x,y)\,dy \right) dx = \int_0^1 \left(\int f(x,y)\,dx \right) dy$$

5.4. Решетка

Используя понятие частично упорядоченного множества, определим понятие «решетка». *Решеткой* называется частично упорядоченное множество $\langle X, \leq \rangle$, любые два элемента x_i, x_j которого имеют наибольшую нижнюю грань, или *пересечение* $x_i \cap x_j$ и наименьшую верхнюю грань, или *объединение* $x_i \cup x_j$. Очевидно, что упорядоченное множество $\overline{X}$, двойственное решетке X, является решеткой, в которой пересечение и объединение меняются ролями.

Упорядоченное множество, в котором любое подмножество имеет наибольшую нижнюю и наименьшую верхнюю грани, называется *полной решеткой.* Очевидно, что если все цепи в решетке конечны, то всякое подмножество в решетке имеет наименьшую верхнюю и наибольшую нижнюю грань.

Найдем в качестве упражнения пересечение и объединение некоторых элементов решетки, определенной диаграммой Хассе H:

$$\{y\} \cup \{x\} = \{y, x\}, \quad 1 \cup \varnothing = 1.$$

$$\{y\} \cap \{a, y\} = \{y\}, \quad \{y, x\} \cap \{a\} = \varnothing.$$

$$\{y, x\} \cap \{a, x\} = \{x\}, \quad \{y\} \cup \{a, x\} = 1.$$

Решетку можно определить и как алгебру $A = \langle X, \cup, \cap \rangle$, сигнатура которой имеет следующие свойства:

идемпотентности

$$x \cup x = x, \; x \cap x = x; \tag{5.1}$$

коммутативности

$$x_i \cup x_j = x_j \cup x_i\,,\ x_i \cap x_j = x_j \cap x_i; \qquad (5.2)$$

ассоциативности

$$(x_i \cup x_j) \cup x_k = x_i \cup (x_j \cup x_k); \qquad (5.3)$$

$$(x_i \cap x_j) \cap x_k = x_i \cap (x_j \cap x_k);$$

поглощения

$$x_i \cup (x_i \cap x_j) = x_i, \quad x_i \cap (x_i \cup x_j) = x_i\,. \qquad (5.4)$$

Здесь $\cup$ — операция взятия *наименьшей верхней* грани элементов *(объединение);* $\cap$ - операция взятия *наибольшей нижней* грани элементов *(пересечение).*

Покажем, что оба определения понятия *решетка* равносильны. Предположим, что дана решетка, которая определена через частично упорядоченное множество. Тогда выполнение свойств идемпотентности и коммутативности очевидно. Проверим, выполняется ли свойство ассоциативности, например, для объединения $\cup$. Так как $\cup$ — наименьшая верхняя грань, то

$$(x_i \cup x_j) \cup x_k \ge x_i \cup x_j \ge x_i;$$

$$(x_i \cup x_j) \cup x_k \ge x_i \cup x_j \ge x_j;$$

$$(x_i \cup x_j) \cup x_k \ge x_k\,.$$

Аналогично имеем:

$$(x_i \cup x_j) \cup x_j \ge x_k;$$

$$(x_i \cup x_j) \cup x_k \ge x_i \cup (x_j \cup x_k);$$

$$(x_i \cup x_j) \cup x_k \ge (x_i \cup x_i) \cup x_k.$$

Свойство ассоциативности доказано.

Докажем справедливость свойства поглощения:

$$x_i \cap (x_i \cap x_j) \ge x_i,$$

так как результат объединения - наименьшая верхняя грань; $x_i \le x_i$, $x_i \ge x_i \cap x_j$, поскольку $x_i \cap x_j$ - наибольшая нижняя грань. Тогда $x_i \ge x_j \cup (x_i \cap x_j)$, так как $\cup$ - наименьшая верхняя грань. Согласно принципу двойственности, утверждение справедливо и для пересечения.

Пусть теперь решетка определена как алгебра с операциями $\cap$ и $\cup$, удовлетворяющими свойствам (5.1) - (5.4). Предварительно заметим, что если $x_i, x_j \in X$, то равенства

$$x_i \cap x_j = x_i, \quad x_i \cup x_j = x_j \qquad (5.5)$$

одновременно могут выполняться или не выполняться. Действительно, если $x_i \cap x_j = x_i$, то, согласно свойствам (5.4) и (5.2),

$$x_i \cup x_j = (x_i \cap x_j) \cup x_j = x_j\,,$$

если $x_i \cup x_j = x_j$, то на основании (3.4)

$$x_i \cap x_j = x_i \cap (x_i \cup x_j) = x_i.$$

Если для элементов x_i и x_j имеют место равенства (5.5), то положим $x_i \le x_j$. Таким образом, в множество X введена частичная упорядоченность. Действительно, согласно свойству (5.1), $x_i \le x_j$.

Далее, если $x_i \le x_j$ и $x_j \le x_k$, т.е. $x_i \cap x_j = x_i$; $x_j \cap x_k = x_j$, то, согласно свойству (5.3),

$$x_i \cap x_k = (x_i \cap x_j) \cap x_k = x_i \cap (x_j \cap x_k) = x_i \cap x_j = x_i, \text{ т.е. } x_i \le x_k.$$

Наконец, если $x_i \le x_j$ и $x_j \le x_i$, т.е. $x_i \cap x_j = x_i$; $x_j \cap x_i = x_j$, то, согласно свойству (5.2), $x_i = x_j$.

Покажем, что выполняется условие наибольшей нижней грани. Из равенства

$$(x_i \cap x_j) \cap x_i = x_i \cap (x_i \cap x_j) = (x_i \cap x_j) \cap x_j = x_i \cap x_j$$

следует, что $x_i \cap x_j \le x_i$.

Аналогично имеем $x_i \cap x_j \le x_j$.

Если же в X взят произвольный элемент x_α, удовлетворяющий условиям $x_\alpha \le x_i$, $x_\alpha \le x_j$, т.е. $x_\alpha \cap x_i = x_\alpha$ и $x_\alpha \cap x_j = x_\alpha$, то

$$x_\alpha \cap (x_i \cap x_j) = (x_\alpha \cap x_i) \cap x_j = x_\alpha \cap x_j = x_\alpha,$$

откуда $x_\alpha \le x_i \cap x_j$. Следовательно, элемент $x_i \cap x_j$ есть наибольшей нижней гранью.

Согласно принццпу двойственности, получаем, что элемент $x_i \cup x_j$ в рассматриваемой алгебре является наименьшей верхней гранью.

В качестве упражнения вычислим значение некоторой формулы F, рассматривая решетку как алгебру. Имеем

$$F(a,b) = ((a \cap a \cup b)) \cap ((a \cup (a \cap b)) \cap (a \cup b))) \cup 0 \cup a =$$
$$= (a \cap (a \cap (a \cup b))) \cup 0 \cup a = (a \cap a) \cup 0 \cup a = a \cup 0 \cup a = a \cup a = a.$$

При упрощении $F(a,b)$ были использованы свойства (5.4), (5.1) и (5.2). Здесь и ниже 0 и 1 в решетке будем называть *структурными нулем* и *единицей.*

Определим в решетке свойство дистрибутивности.

Подрешеткой A' решетки A называется подмножество решетки A, которое вместе с каждой парой элементов x_i, x_j из A содержит также $x_i \square x_j$ и $x_i \cap x_j$.

Интервалом I, определенным элементами x_α и x_β в решетке A, называется подрешетка A' решетки A с наибольшим элементом x_β и наименьшим элементом x_α:

$$I = [x_\alpha, x_\beta] = \{x_i \in A' / x_\alpha \leq x_i \leq x_\beta\}.$$

В решетке A со структурными нулем и единицей два элемента x_α и x_β называются *дополнительными,* если

$$x_\alpha \cap x_\beta = 0, \quad x_\alpha \cup x_\beta = 1.$$

Элемент $\overline{x}$, дополнительный к x, называется также *дополнением* элемента x в решетке A. Два элемента, обладающие общим дополнением в решетке A, называются *связанными* в A.

Важным классом решеток являются дистрибутивные решетки. Решетка A называется *дистрибутивной,* если она удовлетворяет следующим тождествам:

$$(x_i \cup x_j) \cap x_k = x_i \cap x_k \cup x_j \cap x_k$$

и

$$x_k \cap (x_i \cup x_j) = x_k \cap x_i \cup x_k \cap x_j$$

для всех $x_i, x_j, x_k \in X$.

Согласно свойству коммутативности пересечения, для определения дистрибутивных решеток достаточно выполнения одного из этих тождеств.

Приведем критерий дистрибутивности решетки: *решетка A дистрибутивна тогда и только тогда, когда в каждом интервале I решетки A любые два связанных в I элемента равны.*

Этот критерий можно выразить в более удобной для вычислений форме, если показать структуру подрешеток, наличие которых выводит решетку из класса дистрибутивных.

Введем понятие дедекиндовых (модулярных) решеток. Решетка A называется *дедекиндовой* тогда и только тогда, когда

$$(x_i \cup x_j) \cap x_k = (x_j \cap x_k) \cup x_i$$

для всех таких $x_i, x_j, x_k \in A$, что $x_j \leq x_k$.

Критерий дедекиндовости решетки: *решетка A дедекиндова тогда и только тогда, когда она не содержит подрешетки, изоморфной решетке A_m* (рис. 3.10, *а*).

Решетка A_m содержит один элемент нулевой высоты, два элемента единичной высоты и по одному элементу, высота которых 2 и 3.

Используя критерий модулярности решеток, сформулируем критерий дистрибутивности в более удобной для вычислений форме: *решетка дистрибутивна тогда и только тогда, когда она не содержит подрешетки, изоморфной A_m, т.е. является дедекиндовой, и не содержит подрешетки, изоморфной подрешетке A_g* (рис. 5.1, *б).* Решетка A_g содержит три цепи длины 2, состоящие из одного элемента нулевой высоты, трех элементов единичной высоты и одного элемента высоты 2.

Решетка *А,* задаваемая диаграммой Хассе *Н*, является не только дедекиндовой, но и дистрибутивной.

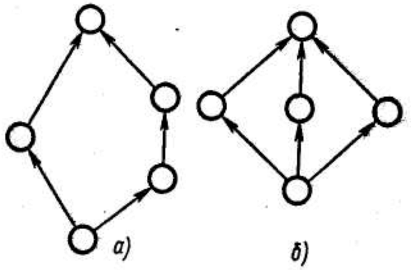


Рис. 5.1.

В решетке A со структурными нулем и единицей, в которой каждый элемент x обладает дополнением $\overline{x}$, можно считать заданной унарной операцией $f(x) = \overline{x}$. Решетка A называется *решеткой с дополнениями,* если она обладает структурным нулем и такой унарной операцией $f_1(x) = \overline{x}$, что

$$\overline{x} = x; \tag{5.6}$$

$$\overline{x_i \cup x_j} = \overline{x}_i \cap \overline{x}_j; \tag{5.7}$$

$$x \cap \overline{x} = 0. \tag{5.8}$$

Согласно (5.6) и (5.7), одна из операций $\cap$, $\cup$ может быть выражена через другую. Следовательно, решетку с дополнениями можно определить как алгебру, сигнатура которой состоит из $\cup$, $^{-}$.

Отметим некоторые следствия тождеств (5.6) — (5.8). Имеем $0 \leq x$ для всех $x \in X$, следовательно, $0 \cap x = 0$.

Если положить $1=\overline{0}$, а $0 \cap x = 0$ и $0 \cup x = x$ подставить в (5.7), то получим $1 \cap x = x$, $1 \cup x = 1$. Следовательно, 1 — наибольший элемент решетки, т.е. является структурной единицей.

Согласно тождествам (5.8) и (5.7),

$$x \cup \overline{x} = 1.$$

Дистрибутивная решетка с дополнениями называется *булевой алгеброй.*

Изоморфизмом η между алгебрами $A_1=<X_1, S_1>$ и $A_2=<X_2, S_2>$ называется взаимно однозначное соответствие между элементами носителей и сигнатур такое, что

$$f_i(x_{a_1}, x_{a_2}, ..., x_{a_{k-1}}) = x_{a_k} \leftrightarrow \eta\left(f_i\right)(\eta(x_{a_1}), \eta(x_{a_2}), ..., \eta(x_{a_{k-1}})) =$$
$$= \eta(x_{a_k}),$$
$$x_{a_j} \in X_1,\ \eta(x_{a_j}) \in X_2, \quad j = 1,2,...,k, \quad f_i \in S_1,\ \eta\left(f\right)_i \in S_2.$$

Алгебры, между которыми существует изоморфизм, называются *изоморфными.* Все законы алгебры A_1 справедливы и в изоморфной ей алгебре A_2.

Теорема. (теорема Стоуна). *Булева алгебра изоморфна алгебре Кантора.*

Для рассмотренных алгебр имеется следующий изоморфизм:

$$a \cup b \leftrightarrow M_a \cup M_b,\ a \cap b \leftrightarrow M_a \cap M_b,\ \overline{a} \leftrightarrow M_a,$$

где в левой части выражений — теоретико-решетчатые, в правой — теоретико-множественные операции. Эти операции имеют одно и то же написание, поэтому, чтобы их различать, аргументы теоретико-решетчатых операций будем обозначать малыми буквами, а аргументы теоретико-множественных операций — прописными буквами латинского алфавита.

На рис. 5.2 изображена дистрибутивная решетка с дополнениями, в которой результаты действия теоретико-решетчатых операций объединения, пересечение и дополнения определены в таблицах (табл. 5.1, *а, б, в).*

Элементам α и $\overline{\alpha}$ соответствуют максимально (в содержании длины соединяющей их цепи) удаленные вершины диаграммы *H.*

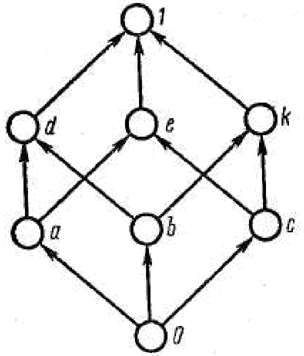


Рис. 5.2.

Таблица 5.1, *а* α □β

α	β							
	0	*a*	*b*	*c*	*d*	*e*	*k*	1
0	0	*a*	*b*	*c*	*d*	*e*	*k*	1
a	*a*	*a*	*d*	*e*	*d*	*e*	1	1
b	*b*	*d*	*b*	*k*	*d*	1	*k*	1
c	*c*	*e*	*k*	*c*	1	*e*	*k*	1
d	*d*	*d*	*d*	1	*d*	1	1	1
e	*e*	*e*	1	*e*	1	*e*	1	1
k	*k*	1	*k*	*k*	1	1	*k*	1
1	1	1	1	1	1	1	1	1

Таблица 5.1, *б* α I β

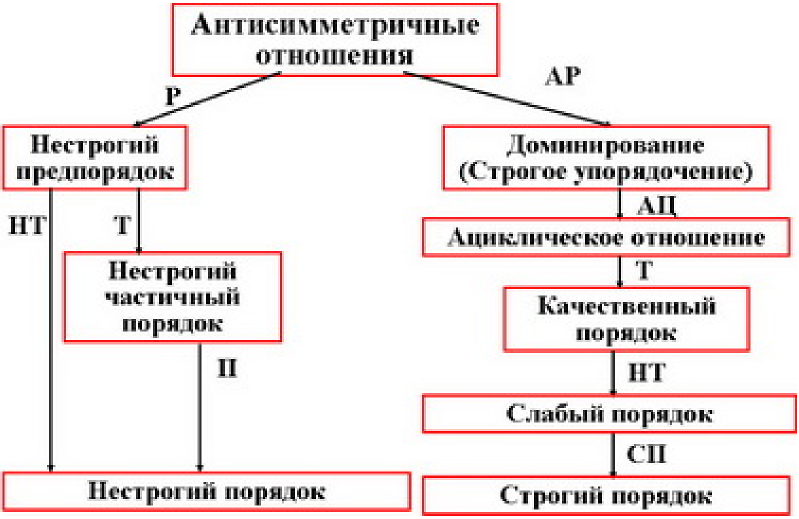
α	β							
	0	*a*	*b*	*c*	*d*	*e*	*k*	1
0	0	0	0	0	0	0	0	0
a	0	*a*	0	0	*a*	*a*	0	*a*
b	0	0	*b*	0	*b*	0	*b*	*b*
c	0	0	0	*c*	0	*c*	*c*	*c*
d	0	*a*	*b*	0	*d*	*a*	*b*	*d*
e	0	*a*	0	*c*	*a*	*e*	*c*	*e*
k	0	0	*b*	*c*	*b*	*c*	*k*	*k*
1	0	*a*	*b*	*c*	*d*	*e*	*k*	1

Таблица 5.1, *в*

α	$\overline{\alpha}$	Комментарии
0	1	0 □ 1=1, 0 I 1=0
a	*k*	*a* U *k*=1, *a* I *k*=0
b	*e*	*b* U *e*=1, *b* I *e*=0
c	*d*	*c* U *d*=1, *c* I *d*=0
e	*b*	*e* U *b*=1, *e* I *b*=0
k	*a*	*k* U *a*=1, *k* I *a*=0
1	0	1 U 0=1, 1 I 0=0

5.5. Отношения доминирования

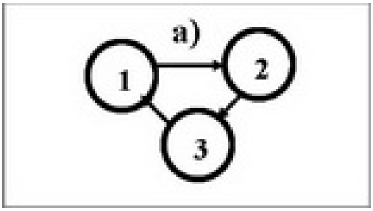
Классификация антисимметричных отношений

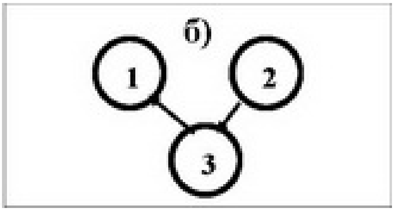


Отношения доминирования (строгого упорядочения)

Доминированием (строгим упорядочением) называется антисимметричное отношение, обладающее свойством антирефлексивности. То есть доминированием является любое асимметричное отношение.

Говорят, что *x* доминирует *y*, если *x* в каком либо смысле превосходит *y*. Так *x* может быть командой, победившей команду *y*, лицом, являющимся родителем лица у и т.д.

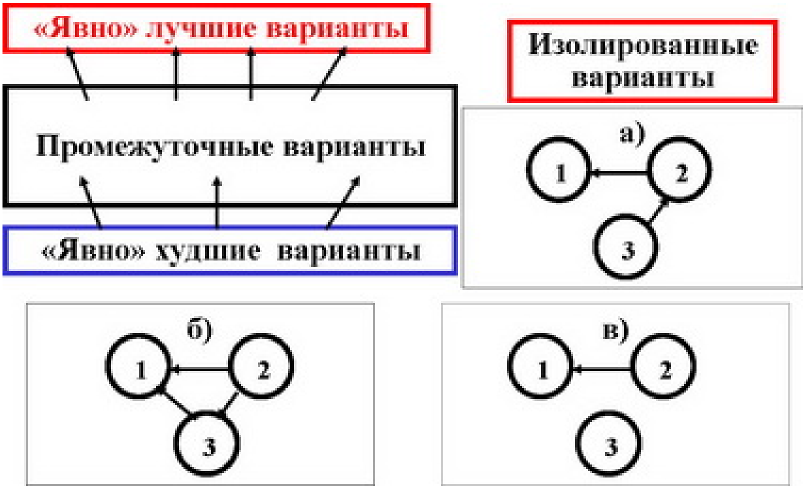




Свойства отношений доминирования

1. Пересечение отношений доминирования остается отношением доминирования.
2. Граф отношения доминирования не содержит ни петель, ни колец , но может содержать циклы большей длины, чем 2.
3. Важнейшим видом отношений доминирования являются ациклические отношения.
4. Например, ациклическими отношениями являются строгий порядок и его обобщения (слабый порядок, качественный порядок).
5. Двойственными к отношениям доминирования являются отношения нестрогого упорядочения.

Структура графов ациклических отношений

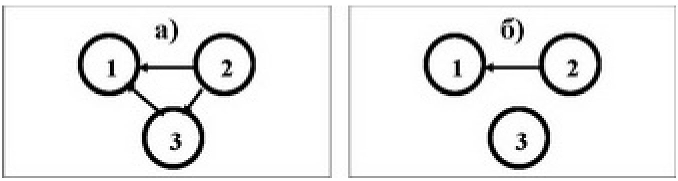


Отношения качественного порядка

Качественный порядок является транзитивным отношением доминирования.

Качественным порядком называется отношение, имеющее свойства антирефлексивности, асимметричности и транзитивности.

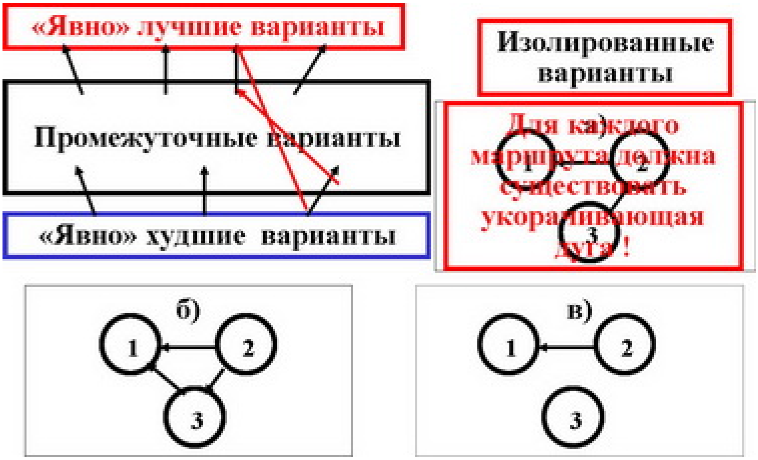
Свойство асимметричности можно исключить из определения качественного порядка, так как его выполнение является следствием выполнения свойств транзитивности и антирефлексивности.



Свойства отношений качественного порядка

1. Инверсия качественного порядка остается качественным порядком.
2. Пересечение качественных порядков является отношением качественного порядка.
3. Двойственное отношение обладает свойствами полносвязности и негатранзитивности, поэтому является отношением нестрогого качественного порядка.
4. Из качественного порядка может быть получен нестрогий частичный порядок путем объединения качественного порядка с диагональю.

Структура графов отношений качественного порядка

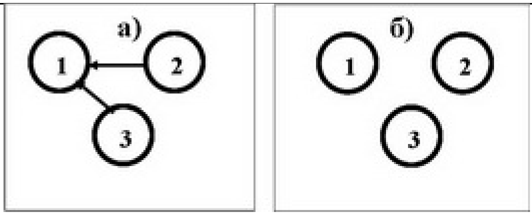


Отношения слабого порядка

Слабым порядком называется любое асимметричное негатранзитивное отношение.

Другими словами, отношение слабого порядка является частным случаем отношения доминирования, при котором дополнительно требуется негатранзитивность.

Бинарное отношение "быть студентом более старшего курса" на множестве студентов некоторого факультета.



Свойства отношений слабого порядка

1. Слабый порядок является транзитивным отношением (так как из выполнения свойств асимметричности и негатранзитивности следует выполнение свойства транзитивности).
2. Слабый порядок является двойственным к отношению нестрогого слабого порядка.
3. Если есть хотя бы одна какая-либо связь, то никакой из вариантов не может быть совсем изолирован от других.
4. Слабый порядок можно представить в виде множества частично упорядоченных слоев, каждый из которых содержит только изолированные друг от друга варианты.
5. Из каждого варианта нижних слоев обязательно существует дуга ко всем вариантам вышележащих слоев.

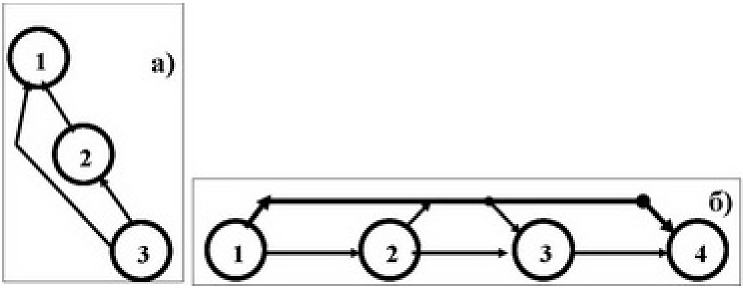
Структура графов отношений слабого порядка

Отношения строгого порядка

Строгим порядком (строгим предпочтением, сильным порядком, строгим линейным порядком) называется антирефлексивное, транзитивное, слабосвязное бинарное отношение.

Строгий порядок является частным случаем слабого порядка (строгого частичного предпочтения) с дополнительным условием слабосвязности.

Пример: Отношение "строго меньше" на множестве целых чисел.



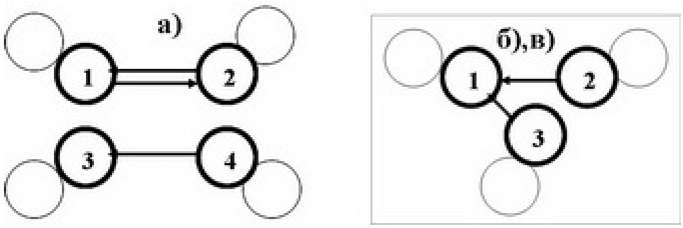
Классификация рефлексивных отношений

Отношения квазипорядка

Эти бинарные отношения позволяют сравнивать элементы некоторого множества, но не по сходству, а путем расположения элементов групп в некотором порядке, т.е. путем частичного упорядочивания.

Квазипорядком (нестрогим частичным предпочтением) называется рефлексивное и транзитивное бинарное отношение.

Пример: "быть братом" (Иван-Петр, Андрей-Анна)



Свойства квазипорядков

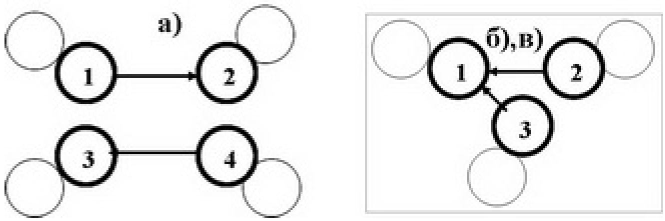
1. Пересечение квазипорядков остается квазипорядком.
2. Симметричная часть квазипорядка обладает свойствами рефлексивности, симметричности и транзитивности и поэтому является отношением эквивалентности. $R^c=R/\backslash R^{инв}$
3. С помощью этого пересечения можно выделить группы эквивалентных между собой вариантов, тогда между выделенными группами может быть установлено отношение нестрогого частичного порядка, порожденное исходным отношением.
4. Асимметричная часть квазипорядка является транзитивным и антирефлексивным отношением = качественный порядок.

Отношения нестрогого частичного порядка

Отношением нестрогого частичного порядка называется отношение, имеющее свойства рефлексивности, антисимметричности и транзитивности.

Нестрогий частичный порядок является антисимметричным квазипорядком

Пример: отношение "быть частью", определенное для множеств (и их подмножеств)



Свойства нестрогих частичных порядков

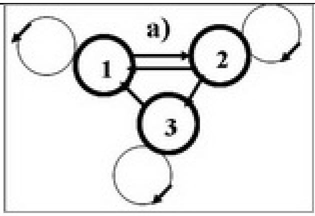
1. Пересечение нестрогих частичных порядков остается нестрогим частичным порядком .
2. Симметричная часть нестрогого частичного порядка является диагональю.
3. Асимметричная часть нестрогого частичного порядка является (строгим) качественным порядком.
4. В теории интеллектуальных систем важную роль играют частично упорядоченные множества – домены вместе с определенными на них отношениями нестрогого частичного порядка.
5. Частично упорядоченные множества с дополнительным свойством существования у каждой пары элементов верхней и нижней граней называются решетками. Частным случаем решеток являются булевы алгебры.

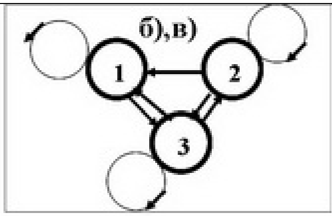
Отношения нестрогого упорядочения

Нестрогим упорядочением называется рефлексивное отношение, обладающее свойством слабосвязности.

Нестрогое упорядочение можно определить также как полносвязное отношение.

Отношение нестрогого упорядочения можно представить как результат объединения некоторых отношений толерантности и доминирования.





Свойства отношений нестрогого частичного упорядочения

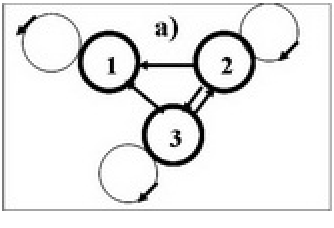
1. Пересечение и объединение полносвязных отношений остается полносвязным отношением.
2. Симметричная часть нестрогого частичного упорядочения является толерантностью.
3. Асимметричная часть нестрогого частичного упорядочения является доминированием.
4. Для полносвязных отношений необходимым условием транзитивности является негатранзитивность отношения.
5. Для полносвязных отношений свойство транзитивности является достаточным условием негатранзитивности отношения.

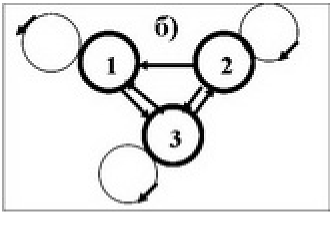
Отношения нестрогого качественного порядка

Бинарное отношение R называется нестрогим качественным порядком, если оно негатранзитивно и полносвязно.

Нестрогий качественный порядок является негатранзитивным нестрогим упорядочиванием.

Отношение нестрогого качественного порядка можно представить как результат объединения некоторых отношений толерантности и качественного порядка.





Свойства отношений нестрогого качественного порядка

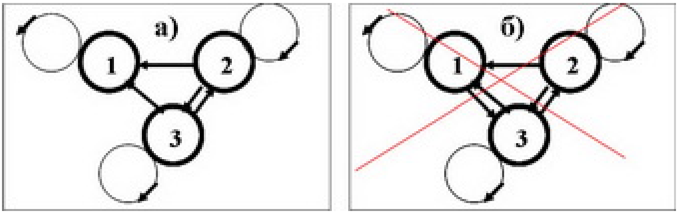
1. Симметричная часть нестрогого качественного порядка является толерантностью. НТ?
2. Асимметричная часть нестрогого качественного порядка транзитивна, поэтому является отношением качественного порядка.
3. Таким образом, отношение нестрогого качественного порядка можно представить как результат объединения отношений толерантности и качественного порядка, порожденных исходным отношением.
4. Двойственное отношение обладает свойствами асимметричности и транзитивности поэтому является отношением качественного порядка.

Отношения нестрогого слабого порядка

Нестрогим слабым порядком называется полносвязное транзитивное и негатранзитивное отношение.

Нестрогим слабым порядком называется полносвязное транзитивное отношение.

Нестрогим слабым порядком называется транзитивное нестрогое упорядочение.



Свойства отношений нестрогого слабого порядка

1. Симметричная часть нестрогого слабого порядка является эквивалентностью.
2. Асимметричная часть R^{ac} нестрогого слабого порядка транзитивна, поэтому является отношением качественного порядка.
3. Таким образом, отношение нестрогого слабого порядка можно представить как результат объединения отношений эквивалентности и

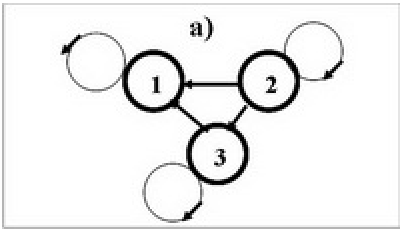
слабого порядка, порожденных исходным отношением.
4. Нестрогий слабый порядок можно представить в виде множества частично упорядоченных слоев, каждый из которых является классом эквивалентности.

Отношения нестрогого (линейного) порядка

Нестрогим порядком (нестрогим линейным порядком) называется антисимметричное, транзитивное, полносвязное бинарное отношение.

Нестрогим порядком называется антисимметричный нестрогий слабый порядок.

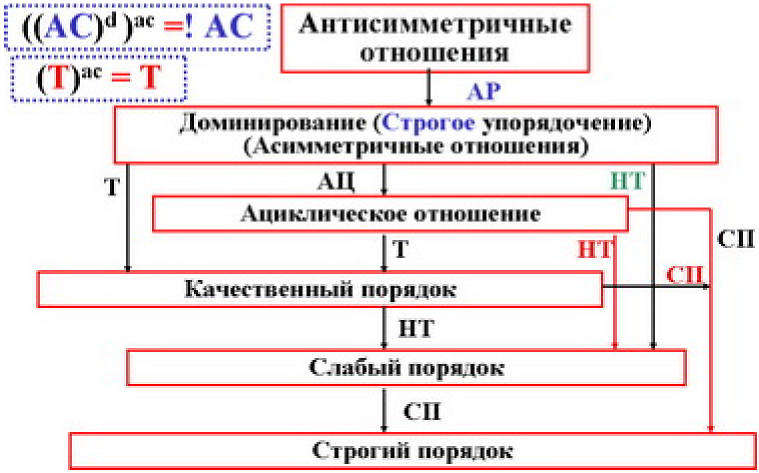
Нестрогим порядком называется антисимметричное нестрогое упорядочение.



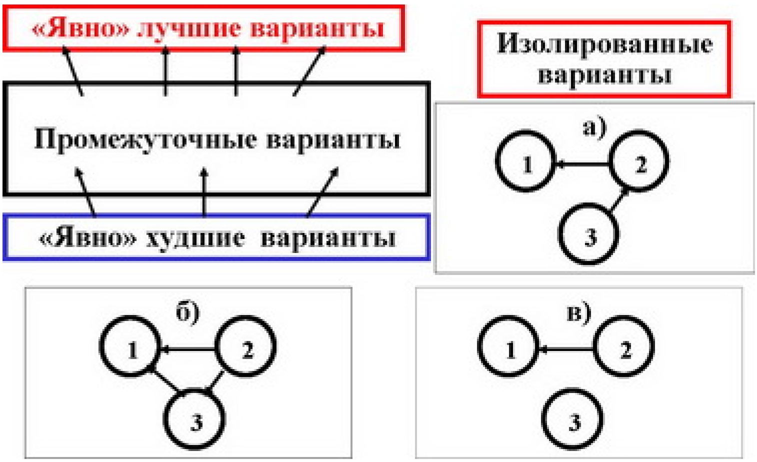
Свойства отношений нестрогого линейного порядка

1. Симметричная часть нестрогого порядка является диагональю.
2. Асимметричная часть R^{ac} нестрогого порядка транзитивна и слабосвязна, поэтому является отношением строгого порядка.
3. Двойственное отношение обладает свойствами асимметричности, негатранзитивности и слабосвязности поэтому является отношением строгого порядка. Кроме того оно совпадает с R^{ac}.
4. Таким образом, отношение нестрогого порядка можно представить как результат объединения диагонали и строгого порядка, порожденных исходным отношением.

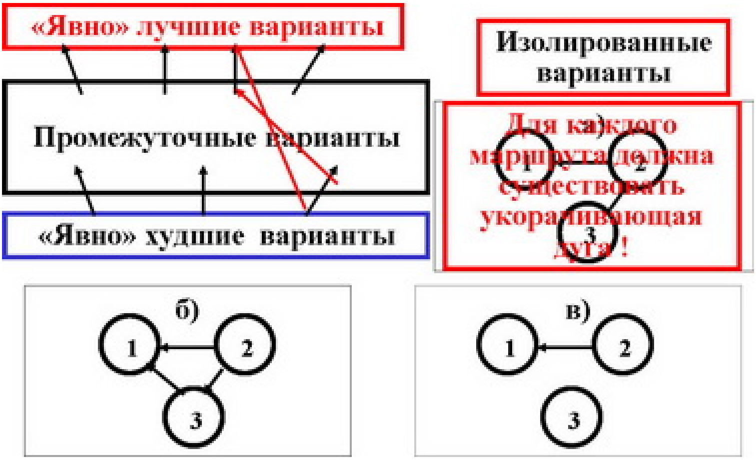
Классификация антисимметричных отношений



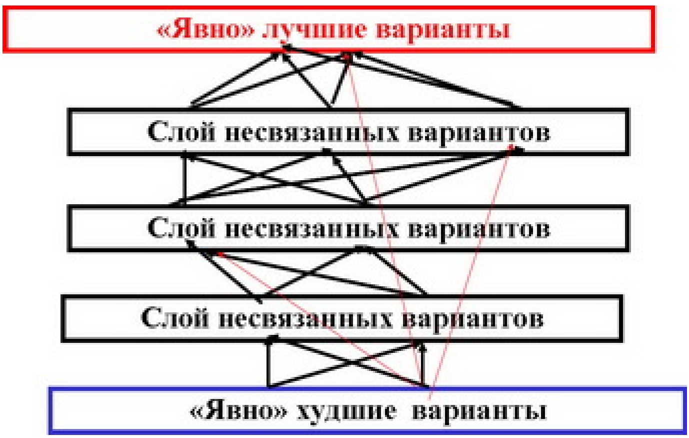
Структура графов ациклических отношений



Структура графов отношений качественного порядка



Структура графов отношений слабого порядка

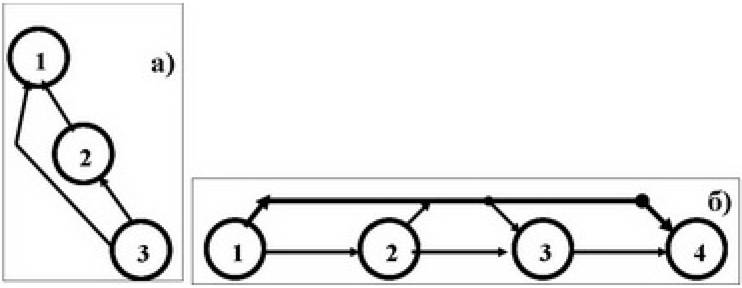


Отношения строгого порядка

Строгим порядком (строгим предпочтением, сильным порядком, строгим линейным порядком) называется антирефлексивное, транзитивное, слабосвязное бинарное отношение.

Строгий порядок является частным случаем слабого порядка (строгого частичного предпочтения) с дополнительным условием слабосвязности.

Пример: Отношение "строго меньше" на множестве целых чисел.



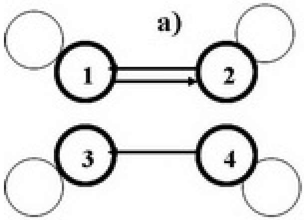
Классификация рефлексивных отношений

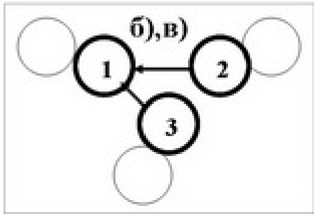
Отношения квазипорядка

Эти бинарные отношения позволяют сравнивать элементы некоторого множества, но не по сходству, а путем расположения элементов групп в некотором порядке, т.е. путем частичного упорядочивания.

Квазипорядком (нестрогим частичным предпочтением) называется рефлексивное и транзитивное бинарное отношение.

Пример: "быть братом" (Иван-Петр, Андрей-Анна)





Свойства квазипорядков

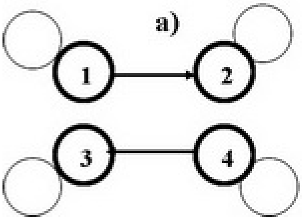
1. Пересечение квазипорядков остается квазипорядком.
2. Симметричная часть квазипорядка обладает свойствами рефлексивности, симметричности и транзитивности и поэтому является отношением эквивалентности. $R^c = R \wedge R^{инв}$
3. С помощью этого пересечения можно выделить группы эквивалентных между собой вариантов, тогда между выделенными группами может быть установлено отношение нестрогого частичного порядка, порожденное исходным отношением.
4. Асимметричная часть квазипорядка является транзитивным и антирефлексивным отношением = качественный порядок.

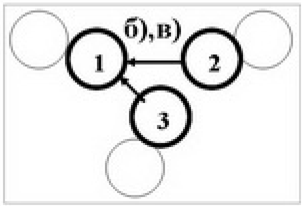
Отношения нестрогого частичного порядка

Отношением нестрогого частичного порядка называется отношение, имеющее свойства рефлексивности, антисимметричности и транзитивности.

Нестрогий частичный порядок является антисимметричным квазипорядком

Пример: отношение "быть частью", определенное для множеств (и их подмножеств)





Свойства нестрогих частичных порядков

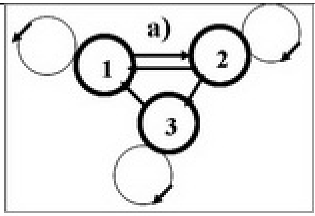
1. Пересечение нестрогих частичных порядков остается нестрогим частичным порядком .
2. Симметричная часть нестрогого частичного порядка является диагональю.
3. Асимметричная часть нестрогого частичного порядка является (строгим) качественным порядком.
4. В теории интеллектуальных систем важную роль играют частично упорядоченные множества – домены вместе с определенными на них отношениями нестрогого частичного порядка.
5. Частично упорядоченные множества с дополнительным свойством существования у каждой пары элементов верхней и нижней граней называются решетками. Частным случаем решеток являются булевы алгебры.

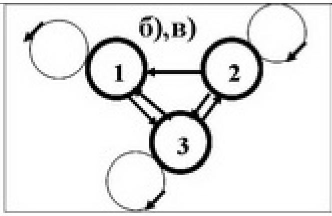
Отношения нестрогого упорядочения

Нестрогим упорядочением называется рефлексивное отношение, обладающее свойством слабосвязности.

Нестрогое упорядочение можно определить также как полносвязное отношение.

Отношение нестрогого упорядочения можно представить как результат объединения некоторых отношений толерантности и доминирования.





Свойства отношений нестрогого частичного упорядочения

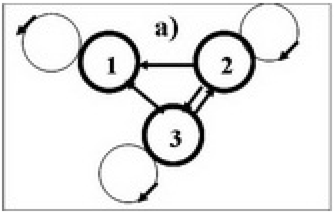
1. Пересечение и объединение полносвязных отношений остается полносвязным отношением.
2. Симметричная часть нестрогого частичного упорядочения является толерантностью.
3. Асимметричная часть нестрогого частичного упорядочения является доминированием.
4. Для полносвязных отношений необходимым условием транзитивности является негатранзитивность отношения.
5. Для полносвязных отношений свойство транзитивности является достаточным условием негатранзитивности отношения.

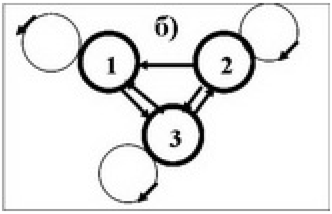
Отношения нестрогого качественного порядка

Бинарное отношение R называется нестрогим качественным порядком, если оно негатранзитивно и полносвязно.

Нестрогий качественный порядок является негатранзитивным нестрогим упорядочиванием.

Отношение нестрогого качественного порядка можно представить как результат объединения некоторых отношений толерантности и качественного порядка.





Свойства отношений нестрогого качественного порядка

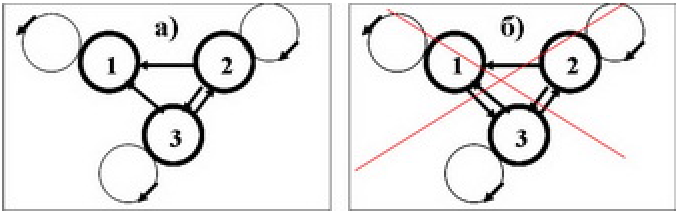
1. Симметричная часть нестрогого качественного порядка является толерантностью. НТ?
2. Асимметричная часть нестрогого качественного порядка транзитивна, поэтому является отношением качественного порядка.
3. Таким образом, отношение нестрогого качественного порядка можно представить как результат объединения отношений толерантности и качественного порядка, порожденных исходным отношением.
4. Двойственное отношение обладает свойствами асимметричности и транзитивности поэтому является отношением качественного порядка.

Отношения нестрогого слабого порядка

Нестрогим слабым порядком называется полносвязное транзитивное и негатранзитивное отношение.

Нестрогим слабым порядком называется полносвязное транзитивное отношение.

Нестрогим слабым порядком называется транзитивное нестрогое упорядочение.



Свойства отношений нестрогого слабого порядка

1. Симметричная часть нестрогого слабого порядка является эквивалентностью.
2. Асимметричная часть R^{ac} нестрогого слабого порядка транзитивна, поэтому является отношением качественного порядка.
3. Таким образом, отношение нестрогого слабого порядка можно представить как результат объединения отношений эквивалентности и

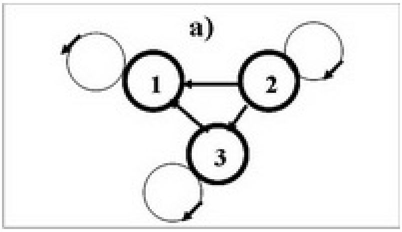
слабого порядка, порожденных исходным отношением.
4. Нестрогий слабый порядок можно представить в виде множества частично упорядоченных слоев, каждый из которых является классом эквивалентности.

Отношения нестрогого (линейного) порядка

Нестрогим порядком (нестрогим линейным порядком) называется антисимметричное, транзитивное, полносвязное бинарное отношение (8).

Нестрогим порядком называется антисимметричный нестрогий слабый порядок.

Нестрогим порядком называется антисимметричное нестрогое упорядочение.



Свойства отношений нестрогого линейного порядка

1. Симметричная часть нестрогого порядка является диагональю.
2. Асимметричная часть R^{ac} нестрогого порядка транзитивна и слабосвязна, поэтому является отношением строгого порядка.
3. Двойственное отношение обладает свойствами асимметричности, негатранзитивности и слабосвязности поэтому является отношением строгого порядка. Кроме того оно совпадает с R^{ac}.
4. Таким образом, отношение нестрогого порядка можно представить как результат объединения диагонали и строгого порядка, порожденных исходным отношением.

Двойственность отношений строгого и нестрогого порядка

Свойства	Отношения	Двойств. свойства	Двойств. отношения
Р + СП= =П		АР+ АнтиС = АС	
5. Нестрогое упорядочение		d	9. Доминирование
+НТ		+Т	
6. Нестрогий качественный порядок		d	10. Качественный порядок
+Т		+НТ	
7. Нестрогий слабый порядок		d	11. Слабый порядок
+АнтиС		+СП	
8. Нестрогий порядок		d	12. Строгий порядок

Обзор свойств различных видов отношений

Свойства: Виды отношений	Р	АР	С	АнтС	АС	Т	НТ	П	СП	АЦ
1.Эквивалентность	Р		С			Т				
2.Толерантность	Р		С							
3.Квазипорядок	Р					Т				
4.Нестрогий частичн. порядок	Р			АнтС		Т				
5.Нестрогое упорядочение	Р							П	СП	
6.Нестрогий качествен. порядок	Р						НТ	П	СП	
7.Нестрогий слабый порядок	Р					Т	НТ	П	СП	
8.Нестрогий порядок	Р			АнтС		Т	НТ	П	СП	
9.Доминирование		АР			АС					
10.Качественный порядок		АР			АС	Т				АЦ
11. Слабый порядок		АР			АС	Т	НТ			АЦ
12. Строгий порядок		АР			АС	Т	НТ		СП	АЦ

6. Сходство и толерантность

6.1. От сходства к толерантности

Ранее мы подробно обсудили содержательный смысл отношения одинаковости объектов. Не менее важной является ситуация, когда нам приходится устанавливать сходство объектов. Если одинаковость объектов обозначает их полную взаимозаменимость в некоторой ситуации, то сходство — это частичная взаимозаменимость, т. е. возможность взаимной замены с некоторыми (допустимыми в данной ситуации) потерями, с допустимым риском.

Например, два новых «Мерседеса» одного выпуска и цвета с точки зрения покупателя вполне одинаковы и, стало быть, взаимозамеиимы.

Но два «Мерседеса» разного выпуска (или новый и старый «Мерседеса» одного выпуска) только похожи. При отсутствии необходимого выбора один может заменить другим, если покупатель готов согласиться с подобной заменой.

Двое близнецов бывают настолько одинаковыми, что без всякого риска могут сдавать экзамены друг за друга. Если два студента только похожи, то такая жульническая проделка, хотя и осуществима, но рискована.

Рисунок голландского художника Эсхера (рис. 6.1) подсказывает читателю, что накапливание несущественных различий у сходных объектов может приводить к совершенно непохожим объектам.

Если для объектов указано только сходство, то невозможно их разбить на четкие классы так, что внутри класса объекты похожи, а между объектами разных классов сходства нет. В случае сходства возникает размытая ситуация без четких границ. Каждый элемент множества несет определенную информацию о похожих на него элементах. Но не всю информацию, как в случае одинаковых элементов. Здесь уже нет дилеммы: «Все или ничего» или «Полная информация — отсутствие информации». Здесь возможны разные степени информации, которую один элемент содержит относительно другого.

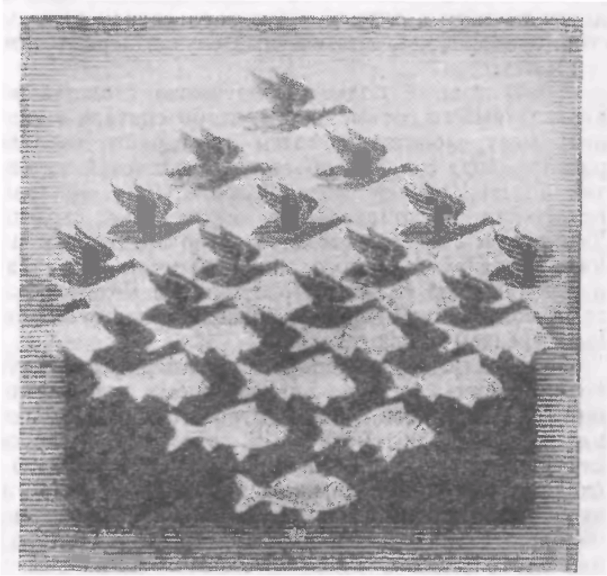


Рис. 6.1. Небо и вода (гравюра М.К. Эсхера)

Превосходная степень от сходства — неразличимость, а вовсе не одинаковость, как может показаться на первый взгляд. Одинаковость — свойство качественно иное. Дело в том, что неразличимые объекты (так же, как и сходные) не разбиваются, вообще говоря, на классы так, чтобы в каждом классе элементы не различались, а элементы разных классов заведомо различались.

В самом деле. Возьмем множество точек на плоскости. Пусть величина d лежит ниже порогаразрешимости глаза, т. е. d — такое расстояние, при котором точки, находящиеся на этом расстоянии, неразличимы зрительно (при выбранном удалении плоскости от наблюдателя). Возьмем теперь n точек, лежащих на одной прямой и отстоящих (каждая — от соседних) на расстоянии d. Каждая пара соседних точек неразличима, но если n достаточно велико, то первая и последняя точки будут отстоять друг от друга на метр и заведомо будут различимы. Разумеется, одинаковость есть частный случай неразличимости и сходства.

Традиционный подход к изучению сходства или неразличимости состоит в том, чтобы сначала определить меру сходства, а затем исследовать взаимное расположение сходных объектов. Английский математик Зиман, изучая модели зрительного аппарата, предложил аксиоматическое определение сходства. Тем самым свойства сходства стало возможным изучать независимо от того, как конкретно оно задано в той или иной ситуации: расстоянием между объектами, совпадением каких-то признаков или субъективным мнением наблюдателя.

Так же, как переход от расплывчатого понятия «одинаковость» к точно определенному типу отношений сопровождался введением нового термина «эквивалентность», математическое отношение, соответствующее нашему интуитивному представлению о сходстве или неразличимости, получило у Зимана название «толерантность». Иначе говоря, толерантность является экспликацией понятия сходства или неразличимости.

Введем следующее

Определение 6.1. Отношение A на множестве M называется *толерантностью* (или *отношением толерантности),* если оно рефлексивно и симметрично.

Естественность такого определения видна из того, что всякий объект заведомо неразличим сам с собой и, уж подавно, похож на себя (это и выражает рефлексивность отношения). Ясно также, что два объекта сходны или не сходны независимо от порядка, в котором мы их рассматриваем. Это обстоятельство выражается симметричностью отношения толерантности. Из примера со зрительной неразличимостью видно, что транзитивность сходства (толерантности) отнюдь не обязательна. Ясно также, что поскольку одинаковость есть частный случай сходства, то эквивалентность должна быть частным случаем толерантности. Сравнивая соответствующие определения мы легко убеждаемся, что так оно и есть. Эквивалентность есть тот частный случай толерантности, когда кроме симметричности и рефлексивности выполняется еще и транзитивность

Рассмотрим теперь серию примеров, где сходство (толерантность) задается разными способами.

Пример 1. Множество M состоит из четырехбуквенных русских слов — нарицательных существительных в именительном падеже. Будем называть такие слова *сходными,* если они отличаются не более чем на одну букву. Известная задача «Превращение мухи в слона» в точных терминах формулируется так: Найти такую последовательность слов, начинающуюся словом «муха» и кончающуюся словом «слон», любые два соседних слова в которой сходны (в смысле только что данного определения).

Приведем решение этой задачи, которая может служить своеобразным образцом студенческого фольклора:

Муха — мура — тура — тара — кара — каре — кафе — кафр — каюр — каюк — крюк — крок — срок — сток — стон — слон.

Очевидно, что в этой задаче самое трудное — найти переход к иной расстановке гласных и согласных (кафе — кафр — каюр — каюк — крюк).

Пример 2. На рис. 6.2 изображены геральдические звери и существа. Между ними существуют разные признаки сходства. В частности, никто нам не мешает принять следующее определение сходства, которое, во всяком случае, ничем не хуже других:

Змея, гидра и дракон сходны как пресмыкающиеся. Гидра, кентавр и вепрь участвуют в мифах о Геракле. Единорог и кентавр сходны очевидным образом: оба суть мифические варианты коня. Единорог и двуглавый орел — мифические существа, изображенные на государственных гербах. Орел и альконоста (женщина-птица) принадлежат к классу птиц, альконоста и дракон сходны потому, что имеют крылья. Именно это отношение сходства выразил художник на рис. 6.2: изображения сходных существ заключены в соприкасающихся кубах.

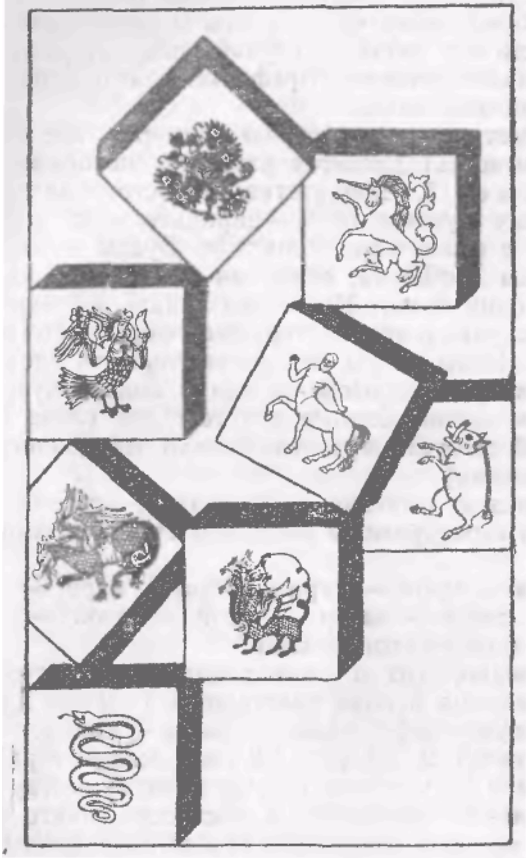


Рис. 6.2. Сходство геральдических существ.

Пример 3. А на рис. 6.3 тем же способом изображена другая группа геральдических существ.

Рыбы и дельфин принадлежат водному царству. Дельфин, изображенный на этом рисунке, внешне сходен с обычным конем (другим оправданием этого сходства может послужить отрывок из В. Брюсова: «Предадим молитве душу, а тебя из мглы пучин тихо вынесет на сушу на спине своей дельфин».) Конь, пегас и единорог сходны чисто анатомически. Единорог, пегас и двуглавый орел образуют группу мифических существ. Сова, орел, двуглавый орел и пегас имеют крылья и этим схожи.

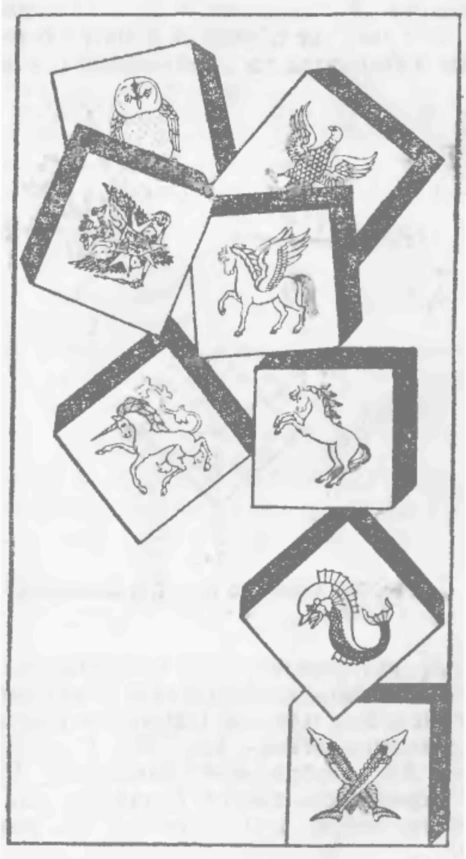


Рис. 6.3. Сходство геральдических существ.

Следующая группа примеров носит более академический характер.

Пример 4. Пусть p — натуральное число. Обозначим через S_p совокупность всех непустых подмножеств множества $\{1, 2, 3, ..., p\}$. Два таких подмножества объявим *толерантными,* если у них есть хотя бы один общий элемент. Законность такого определения очевидна: рефлексивность и симметричность данного отношения легко проверяются.

Множество S_p называется $(p - 1)$-*мерным симплексом.* Это понятие обобщает понятия отрезка, треугольника и тетраэдра на многомерный случай. Числа 1, 2, ..., p интерпретируются как *вершины* симплекса.

Двухэлементные подмножества — как *ребра,* трехэлементные — как *плоские* (двумерные) *грани,* k -элементные подмножества — как (k — 1)*-мерные грани.* На рис. 6.4 изображены симплексы S_1, S_2, S_3 и S_4.

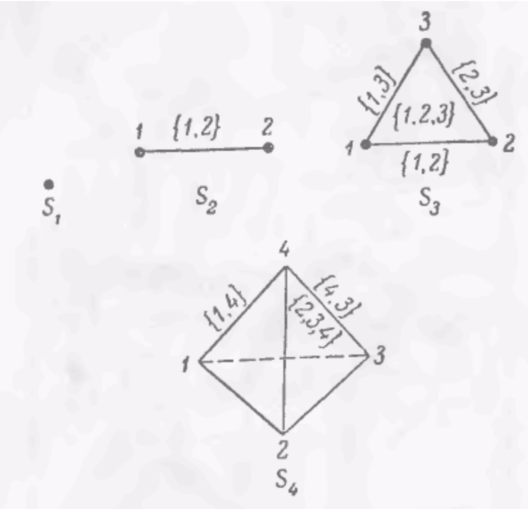


Рис. 6.4. Сходство граней в симплексе.

Толерантность граней симплекса S_p означает их геометрическую инцидентность — наличие общих вершин.

Число всех элементов из S_p равно 2^p—1.

Мы можем элементы множества S_p изобразить в виде вершин графа, тогда ребра, как обычно, будут изображать выполнение соответствующего соотношения. На рис. 6.5 дано изображение для S_3.

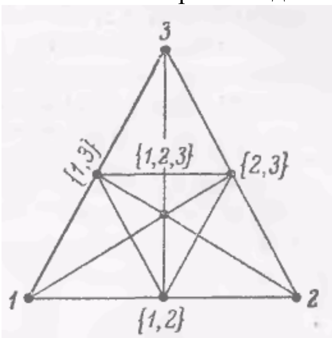


Рис. 6.5.

Нам будет удобно использовать

Определение 6.2. Множество M с заданны и на нем отношением толерантности τ называется *пространством толерантности.* Таким образом, пространство толерантности есть пара $<M, \tau>$.

Пример 5. Пространство толерантности S_p допускает изящное обобщение на бесконечный случай. Пусть H — произвольное множество. Обозначим через S_H совокупность всех непустых подмножеств множества H. Толерантность τ на S_H задается условием: $X\tau Y$, если $X \cap Y \neq \varnothing$. Симметричность и рефлексивность этого отношения очевидны. Пространство толерантности S_H будет играть в дальнейшем особую роль — роль «универсального» пространства толерантности.

Пример 6. Пусть p — натуральное число. Обозначим через B_p множество всех *двоичных кортежей* длины p. Таким образом, любой элемент $x \in B_p$ имеет вид $x = \langle \xi_1, \xi_2, \ldots, \xi_p \rangle$, где ξ = 0 или 1. Толерантность τ в B_p задается правилом: если $x = \langle \xi_1, \xi_2, \ldots, \xi_p \rangle$ и $y = \langle \eta_1, \eta_2, \ldots, \eta_p \rangle$, то $x\tau y$ означает, что хотя бы при каком-нибудь i: $\xi_i = \eta_i$. Иначеговоря, толерантность двух элементов: $x\tau y$ — означает, что у них есть хотя бы одна общая компонента. Общее количество элементов в B_p равно, очевидно, 2^p. Для каждого $x = \langle \xi_1, \xi_2, \ldots, \xi_p \rangle$ имеется ровно один не толерантный к нему элемент $y = \langle 1 - \xi_1, 1 - \xi_2, \ldots, 1 - \xi_p \rangle$, у которого отличны от x все компоненты.

Для тех, кто отчетливо понимает, что компоненты кортежа длины p — это координаты точки в p-мерном пространстве, будет ясно, что B_p состоит из всех вершин p-меркого единичного куба (рис. 6.6; при изображении пространства B_3 мы опустили в графе все диагональные связи: чтобы изобразить все толерантности между элементами, следовало бы провести все диагонали на гранях куба.)

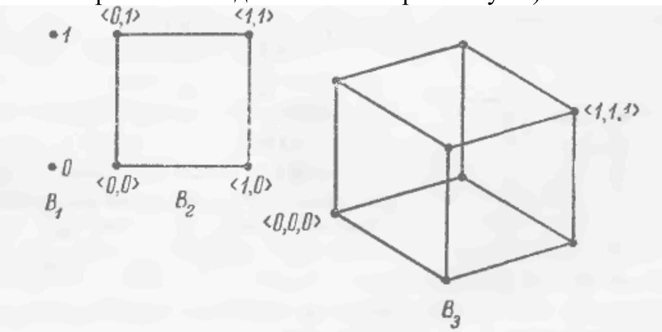


Рис. 6.6.

Пример 7. Простым обобщением пространства B_p является пространство толерантности $B^m{}_p$, где компоненты ξ_i принимают любые целые значения от 0 до m— 1, а толерантность определяется как совпадение хотя бы одной компоненты. Очевидно, $B_p = B^2{}_p$.

Пример 8. Следующее обобщение состоит в том, чтобы рассмотреть пространство толерантности B_p, компоненты элементов которого принимают любые вещественные значения.

В частности, $B^\infty{}_2$ есть множество всех пар вида $x = \langle \xi_1, \xi_2 \rangle$, где ξ_1 и ξ_2 — произвольные вещественные числа. Элементы пространства $B^\infty{}_2$ можно изображать точками на плоскости, если ξ_1 и ξ_2 интерпретировать как декартовы координаты. Толерантность двух точек означает совпадение хотя бы одной координаты. Стало быть, две толерантные точки всегда находятся либо на общей вертикали, либо на общей горизонтали. На рис. 6.7 изображена координатная плоскость, на которой точки x_1 и x_2 толерантны, а также толерантны x_2 и x_3.

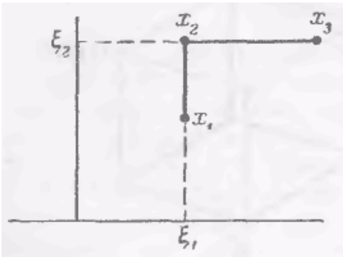


Рис. 6.7.

Точки x_1 и x_3 уже не являются толерантными.

При других p пространство $B^\infty{}_p$ можно интерпретировать как множество точек p-мерного пространства. Но интересней другая интерпретация пространства $B^\infty{}_p$. Каждый кортеж $x = \langle \xi_1, \xi_2, \dots, \xi_p \rangle \in B_p^\infty$ можно считать числовой функцией, заданной на множестве $\{1, 2, \dots, p\}$: каждому числу $j \quad (1 \leqslant j \leqslant p)$ функция x сопоставляет число ξ_j. Толерантность двух функций x и y означает, что они хотя бы в одной точке принимают одинаковое значение (рис. 6.8).

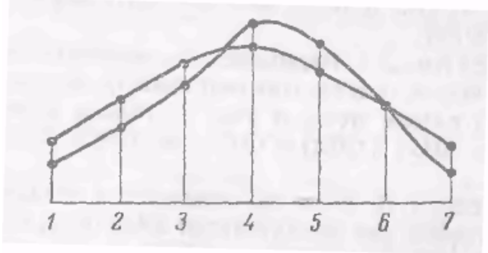


Рис. 6.8. От p-мерных векторов к функциям.

Пример 9. Возьмем теперь произвольное множество M (для наглядности можно представить себе отрезок на прямой). Пространство толерантности $B^{\infty}{}_p$ состоит из всех *числовых функций,* определенных на этом множестве (то есть функций, которые каждому элементу из M сопоставляют некоторое число.) Две функции объявляются толерантными, если хотя бы на одном элементе из M эти функции принимают одно и то же значение (если, другими словами, графики этих функций пересекаются).

Так как $B^{\infty}{}_p$ можно рассматривать как совокупность точек p-мерного пространства, то $B^{\infty}{}_M$ — совокупность всех функций на некотором бесконечном множестве — естественно считать типовым бесконечномерным пространством. (Эта идея — считать совокупность функций обобщением многомерного пространства — лежит в основе важного раздела математики, так называемого функционального анализа.)

Существует еще один важный способ задания отношении толерантности. Рассмотрим соответствие

$$\varphi\colon\ M \to L.$$

Множество всех образов элемента x при соответствии φ (т. е. множество элементов, соответствующих элементу x при соответствии φ) мы обозначим $\Phi(x)$. Отношение A_φ на множестве M задается условием: $xA_\varphi y$, если у элементов x и y существует общий образ,

т. е. если $\Phi(x) \cap \Phi(y) \neq \varnothing$.

Установим основные свойства отношения A_φ:

Свойство 1. Отношение A_φ всегда симметрично. Это следует просто из того, что $\Phi(x) \cap \Phi(y) = \Phi(y) \cap \Phi(x)$.

Свойство 2. Отношение A_φ рефлексивно тогда и только тогда, когда соответствие φ определено на всем M. В самом деле, в этом и только этом случае множество $\Phi(x) \cap \Phi(x) = \Phi(x)$ не пусто для любого $x \in M$.

Свойство 3. Если на элементе x отношение A_φ не рефлексивно (не выполняется $xA_\varphi x$ или, что то же, $\Phi(x) = \varnothing$), то соотношение $xA_\varphi y$ не выполнено ни для какого y, так как $\Phi(x) \cap \Phi(y) = \varnothing \cap \Phi(y) = \varnothing.$

Это свойство имеет простой геометрический смысл: если вершина x в графе, изображающем A_φ, не имеет петли, то она не связана ни с какой другой вершиной. Иначе говоря, для отношений типа A_φ нерефлексивность может быть только такого типа: если на элементе x отношение A_φ не рефлексивно, то этот элемент уже ни с кем не вступает в отношение.

Свойство 4. Если соответствие $\varphi\colon\ M \to L$ является функцией, т. е. для любого $x \in M$ $\Phi(x)$ состоит не более чем из одного элемента (в этом случае $xA_\varphi y$ равносильно или $\Phi(x) = \Phi(y)$ или $\varphi(x)= \varphi(y)$), то отношение A_φ транзитивно.

Действительно, пусть $xA_\varphi y$ и $yA_\varphi z.$ Это значит, что $\varphi(x) = \varphi(y)$ и $\varphi(y) = \varphi(z).$ Следовательно, $\varphi(x)= \varphi(z)$, т.е. $xA_\varphi z$.

Из этих свойств следует, что всюду определенное соответствие $\varphi\colon\ M \to L$ определяет на M симметричное и рефлексивное отношение A_φ, т. е. толерантность. В 6.3 мы увидим, что любое отношение толерантности может быть определено как отношение A_φ по некоторому соответствию φ (теорема 6.4).

Если соответствие ф вдобавок является функцией, то отношение A_φ — эквивалентность. Ранее мы убедились, что любое отношение эквивалентности может быть определено как A_φ, где φ есть отображение множества M на некоторое множество L.

Конец этого параграфа не имеет прямого отношения к изучаемому понятию толерантности. Однако описываемые далее типы отношений естественно возникают в ряде ситуаций и заслуживают некоторого рассмотрения. Но их роль не настолько велика, чтобы посвящать им отдельную главу или параграф.

Всякое транзитивное и симметричное отношение A на множестве M можно представить как отношение типа A_φ. Для доказательства этого утверждения нужно вспомнить следующее свойство отношений, одновременно транзитивных и симметричных: если существует y такое, что xAy, то выполнено xAx (xAy влечет yAx, а отсюда, по транзитивности, вытекает xAx). Таким образом, элементы, на которых отношение A не рефлексивно, ни с кем не связаны отношением A. Возьмем теперь подмножество M_0 множества M, состоящее из всех *рефлексивных элементов* (таких, что xAx выполнено). Тогда на M_0 мы имеем отношение эквивалентности. Множество классов эквивалентности обозначим через L. Определим теперь функцию $\varphi\colon\ M \to L$ условием: если $x \in M_0$, то $\varphi(x)$ есть класс эквивалентности, содержащий x если x не входит в M_0, то $\varphi(x)$ не определено. Отношение

A_φ, определенное по этому φ, на M_0 совпадает с сужением отношения A на M_0, а, при $x \in M \setminus M_0$, $\Phi(x) = \varnothing$ и $xA_\varphi y$ не выполнено ни для какого y.

Когда отношение A транзитивно и симметрично, то нерефлексивность может быть только типа, описанного свойством 3.

Если же A только симметрично, то элемент может быть не рефлексивным, но вступать в отношения с другими элементами. Поэтому далеко не все симметричные отношения могут быть представлены в форме A_φ. Легко показать, что симметричное отношение представляется в форме A_φ, если оно обладает свойством 3.

Однако есть другой способ задавать отношение через соответствие φ. Пусть задано соответствие $\varphi\colon M \to L$, а отношение B_φ на M задается условием: $xB_\varphi y$, если множества образов $\Phi(x)$ и $\Phi(y)$ имеют ровно один общий элемент.

Разница между A_φ н B_φ состоит содержательно в том, что A_φ есть отношение «иметь хотя бы один общий признак», а B_φ — «иметь ровно один общий признак». Нетрудно заметить, что B_φ — обязательно симметричное отношение. Если соответствие $\varphi\colon M \to L$ является функцией, то $A_\varphi = B_\varphi$. Смысл предыдущего определения показывает

Теорема 6.1. *Пусть отношение B на множестве M симметрично и антирефлексивно. Тогда существует соответствие* $\varphi\colon M \to L$ *такое, что* $B = B_\varphi$.

Доказательство. Рассмотрим граф, изображающий отношение B (причем, как мы уже уславливались, поскольку B симметрично, элементы x и y, находящиеся в отношении B, мы будем вместо двух стрелок соединять одним ребром.). Пусть L — множество всех его вершин и ребер. Пусть соответствие $\varphi\colon M \to L$ определяется следующим образом.

Если $x \in M$ — изолированная вершина *(x* не связан отношением B ни с кем), то соответствие φ на x не определено, т. е. элементу x ничего не сопоставляется. Если x — неизолированная вершина, то $\Phi(x)$ состоит из всех ребер, содержащих вершину x и самой вершины x. Ясно, что $\Phi(x)$ содержит в этом случае более одного элемента (самое вершину и еще хотя бы одно ребро, так как вершина не изолирована). Таким образом, $xB_\varphi x$ всегда не выполнено. Пусть теперь имеет место соотношение xBy. Это значит, что $x \neq y$ и вершины x и y графа соединены общим ребром (единственным). Это ребро является динственным общим элементом множеств $\Phi(x)$ и $\Phi(y)$, т. е. $xB_\varphi y$ также выполнено. С другой стороны, выполнение соотношения $xB_\varphi y$ означает, что $x \neq y$ и у вершин x и y есть общее ребро. Стало быть, выполнено соотношение xBy. Теорема доказана.

Итак, симметричное и антирефлексивное отношение B на множестве M всегда можно описать, задав на M такого систему признаков, что соотношение xBy будет выполнено в том и только том случае, когда x и y имеют ровно один общий признак.

Примером симметричного и антирефлексивного отношения служит отношение «рифмоваться» на множестве русских слов. Очевидно, что, если слово x рифмуется со словом y, то и y рифмуется с x. По традициям русского стихосложения рифмовать слово с самим собой не полагается, т. е. это отношение естественно считать антирефлексивным. Заметим, что уже отсюда следует нетранзитивность отношения «рифмоваться». Действительно, из транзитивности и симметричности вытекало бы, что для всякого слова x, имеющего хоть какую-то рифму y, выполнено: «x рифмуется с x». Впрочем, для современных рифм легко указать цепочку слов, в которой каждые соседние слова рифмуются, а первое и последнее совершенно не созвучны: пли — пари — тариф — рифм — ритм.

6.2. Операции над толерантностями

Алгебраические свойства операций над толерантностями сравнительно просты. Большую часть этих свойств мы фактически уже получили ранее. Тем не менее мы систематизируем имеющиеся у нас сведения, добавив необходимые новые.

Лемма 6.1. *Если A — толерантность, B — эквивалентность и* $A \subseteq B$, то $\hat{A} \subseteq B$.

Доказательство получается применением транзитивного замыкания к обеим частям включения $A \subseteq B$.

Смысл этой леммы в том, что транзитивное замыкание $\hat{A}$ отношения толерантности A есть минимальная эквивалентность, включающая эту толерантность.

Из приведенных ранее лемм и следствия следует, что если отношения A и B суть толерантности, то толерантностями также являются и следующие отношения: $A \cup B$, $A \cap B$, A^{-1} и $\hat{A}$.

Из приведенных ранее лемм следует

Теорема 6.2. *Для того чтобы произведение AB отношений толерантности A и B было толерантностью, необходимо и достаточно, чтобы A и B коммутировали. В этом случае* $AB = A \circ B$.

Симметризованное произведение AB толерантностей A и B всегда будет толерантностью. В самом деле, рефлексивность вытекает из Из приведенной ранее леммы. Симметричность симметризовашюго произведения $A \circ B$ следует из того, что

$$(A \circ B)^{-1} = (AB \cup BA)^{-1} = (AB)^{-1} \cup (BA)^{-1} =$$
$$= B^{-1}A^{-1} \cup A^{-1}B^{-1} = BA \cup AB = AB \cup BA = A \circ B.$$

Можно ввести еще один вариант симметризованного произведения: $A * B = AB \cap BA.$ Легко показать, что $A * B$ будет толерантностью, если A и B — толерантности.

Полезно еще заметить, что для любого рефлексивного отношения A отношения $A \cup A^{-1}$, $A \cap A^{-1}$, $A \circ A^{-1}$ будут толерантностями.

6.3. Классы толерантности

Мы займемся здесь изучением структуры пространств толерантности и попробуем различными способами представить, как устроены произвольные пространства толерантности. Содержательно, общий результат состоит в том, что любое отношение толерантности может быть задано набором признаков так, что толерантные элементы суть те, которые имеют общие признаки.

Охарактеризовать некоторую совокупность объектов *признаками* означает, строго говоря, следующее. Возьмем множество M всех этих объектов и множество N всех возможных признаков. Установим теперь соответствие

$$\varphi: M \to N, »$$

сопоставляющее каждому объекту из M все те признаки, которыми он обладает. Наоборот, любое соответствие $\psi: M \to N$ можно содержательно интерпретировать как присвоение некоторым объектам (элементам множества M) некоторых признаков (элементов из N).

Итак, строгое понятие «соответствие» позволяет придать точный смысл обиходному выражению «иметь признаки». Ранее мы показали, что всякое всюду определенное на M соответствие φ задает на множестве M отношение толерантности A_φ, определяемое как совпадение хотя бы одного признака (наличие общего признака).

Мы покажем, что любое отношение толерантности можно задать таким образом. Более того, существует некоторая каноническая совокупность признаков, которая строится по данному отношению толерантности независимо от способа его конкретного задания.

Отношение толерантности A_φ на множестве M может быть определено также на языке покрытий. (Система множеств П называется *покрытием* множества *M,* если $\bigcup_{A\in\Pi} A \supseteq M$.) В самом деле. Пусть $\varphi\colon M \to N$ — всюду определенное соответствие. Сопоставим каждому «признаку» $\xi \in N$ множество $M(\xi)$ всех элементов из *M,* обладающих признаком ξ, т. е. множество $\varphi^{-1}(\{\xi\})$. Система всех множеств $M(\xi)$ образует покрытие множества

$$M\colon M = \bigcup_{\xi} M(\xi),$$

поскольку любой элемент $x \in M$ входит в некоторое $M(\xi)$. Легко видеть, что $xA_\varphi y$ тогда и только тогда, когда существует такой признак ξ, что $x \in M(\xi)$ и $y \in M(\xi)$. Таким образом, толерантность A_φ может быть задана так: $xA_\varphi y$, если x и y принадлежат некоторому общему классу покрытия $\{M(\xi)\}$. В этом параграфе мы построим каноническое покрытие пространства толерантности.

Теоремы этого параграфа — хороший пример так называемых классификационных теорем, когда объекты, заданные абстрактными аксиомами, «материализуются» в виде конкретной и обозримой конструкции. Перейдем к формальным построениям. Пусть задано пространство толерантности $\langle M, \tau\rangle$.

Определение 6.3. Множество $L \subseteq M$ называется *предклассом* в $\langle M, \tau\rangle$, если любые два его элемента x и y толерантны, т. е. для них выполнено соотношение $x\tau y$.

Л е м м а 6.2. *Для того чтобы два элемента x и y были толерантны, необходимо и достаточно, чтобы существовал предкласс L, содержащий оба эти элемента.*

Доказательство. Если x и y лежат в предклассе *L,* то по определению предкласса выполнено соотношение $x\tau y$. Если $x\tau y$, то множество $\{x, y\}$ само образует предкласс, так как, кроме исходного соотношения, выполнены также соотношения $x\tau x$, $y\tau y$ и $y\tau x$.

Определение 6.4. Множество $K \subseteq M$ называется *классом толерантности* в $\langle M, \tau\rangle$, если *K* есть максимальный предкласс.

Это значит, что любое множество $R \supset K$ уже не является предклассом. Или, иначе, для всякого элемента $z \in M$, не входящего в *K,* существует элемент $x \in K$, не толерантный к *z*.

Лемма 6.3 (о пополнении предклассов). *Всякий предкласс содержится хотя бы в одном классе K.*

Доказательство мы проведем лишь для случая, когда само множество *M* конечно, так как в общем случае необходимо использовать так называемую трансфинитную индукцию.

Итак, пусть L — предкласс. Если само L есть класс, то лемма доказана. Если L — не класс, то в множестве $M \setminus L$ существует элемент z, толерантный ко всякому элементу из L. Добавим такой элемент z к L, т. е. рассмотрим множество $L_1 = L \cup \{z\}$. Тогда $L_1 \supset L$ и L_1 снова является предклассом. Либо L_1 — класс, либо мы продолжаем дальше этот процесс расширения предкласса до класса. Поскольку множество M конечно, через конечное число шагов наше построение класса закончится. Лемма доказана.

Следствие. Всякий элемент $x \in M$ содержится в некотором классе, т. е. *система классов толеранстности образует покрытие множества M.*

В самом деле, в силу рефлексивности, $x\tau x$ и множество $\{x\}$, состоящее из одного элемента x, образует предкласс.

Из лемм 6.2, 6.3 непосредственно вытекает

Лемма 6.4. *Для того чтобы два элемента x и y были толерантны, необходимо и достаточно, чтобы существовал класс, содержащий оба эти элемента.*

Теперь у нас все подготовлено к тому, чтобы сформулировать и доказать основную классификационную теорему. Напомним еще раз определение пространства толерантности S_H. ОНО СОСТОИТ ИЗ всех непустых подмножеств множества H. Подмножества считаются толерантными в S_H, если их пересечение непусто.

Теорема 6.3. *Пусть $\langle M, \tau\rangle$— произвольное пространство толерантности, а H — множество всех его классов толерантности. Тогда существует отображение*

$$\varphi:\ M \to S_H \tag{6.1}$$

такое, что элементы из M толерантны в том и только том случае, когда толерантны их образы в S_H.

Доказательство. Выберем в качестве φ отображение, которое каждому элементу $x \in M$ сопоставляет множество $H(x)$, состоящее из всех содержащих его классов. По следствию из леммы 6.3 $H(x) \neq \varnothing$ для любого x. По лемме 6.4 отношение $x\tau y$ выполнено в том и только том случае, когда $H(x) \cap H(y) \neq \varnothing$, т. е. $H(x)$ и $H(y)$ содержат общий класс.

Если M конечно, то количество всех его подмножеств конечно и, стало быть, конечно пространство S_H. Поэтому вместо отображения (6.1) можно взять отображение $\varphi:\ M \to S_p$, где p — число классов толерантности в $\langle M, \tau\rangle$, которое каждому элементу x сопоставляет множество номеров, содержащих его классов:

$$x \to \{n_1,\ n_2,\ \ldots,\ n_k\} \tag{6.2}$$

(здесь $n_i \leq p$). Толерантность элементов x и y означает, что среди номеров, сопоставленных элементам x и y согласно (6.2), есть хотя бы один общий. Иными словами, x и y имеют общий числовой признак.
Рассмотрим в качестве примера множество геральдических существ на рис. 6.9.

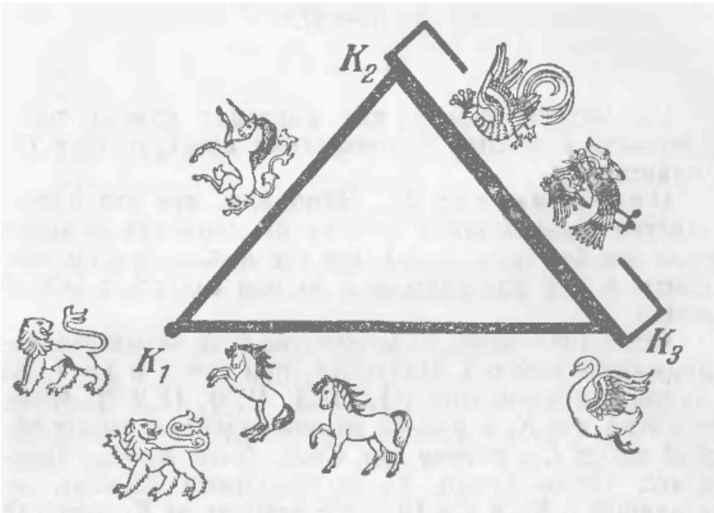


Рис. 6.9. Группировка по классам толерантности.

Будем полагать толерантными те из них, которые имеют общим один из следующих признаков: 1) быть млекопитающим; 2) быть мифическим существом; 3) быть птицей. Легко видеть, что все существа, обладающие одним из этих признаков, толерантны меж собой и тем самым образуют предкласс. Можно проверить, что для множества существ на рис. 6.9 эти предклассы суть классы толерантности. Львы и кони обладают только первым признаком и им соответствует вершина K_1 трехмерного симплекса (попросту говоря, треугольника). Единорог обладает первым и вторым признаком и поэтому изображен на ребре K_1K_2. Алькоността и райская птица обладают вторым и третьим признаком. Им соответствует ребро K_2K_3. Лебедь помещен в вершину K_3, поскольку он обладает только третьим из признаков.
Рассмотрим теперь всюду определенное соответствие

$$\varphi:\ M \rightarrow H,$$

которое каждому $x \in M$ сопоставляет все классы, в которые он входит. Из леммы 6.4 вытекает, что $x\tau y$ равносильно тому, что у x и y имеется общий образ в H. Тем самым доказана анонсированная в 6.1.
Теорема 6.4. (Л. Кальмар —С. Якубович). *Произвольное отношение толерантности τ на множестве М можно задать как отношение A_φ с помощью некоторого всюду определенного соответствия*

$$\varphi: M \to H.$$

* * *

Рассмотрим теперь, как выглядят классы толерантности в некоторых конкретных пространствах толерантности.
Пространство S_p. Напомним, что это пространство толерантности состоит из множеств номеров вида $x = \{n_1, n_2, \ldots, n_k\}$, где все $n_i \leqslant p$, причем элементы x и y толерантны, если они содержат общий номер.
Обозначим через K_i множество всех элементов, содержащих номер L. Например, при $p = 3$ и $i= 1$, K_1 состоит из элементов {1}, {1,2}, {1,3}, {1,2,3}. Ясно, что если $x \in K_i$ и $y \in K_i$, то они заведомо имеют общий номер i, и потому $x\tau y$. Стало быть, K_i есть предкласс. Пусть теперь z — произвольный элемент, не входящий в K_i а $x = \{i\}$ — тот элемент из K_i который имеет единственный номер i. Ясно, что $x\tau z$ не выполнено, поскольку z не содержит номера i, а x содержит только этот номер. Значит, предкласс K_i нельзя расширить и справедлива
Лемма 6.5. *Множество K_i является классом толерантности.*
Так как K_i состоит из всех множеств вида $\{i, n_1, ..., n_k\}$, то число элементов множества K_i i равно 2^{p-1} — число всех подмножеств множества из оставшихся p — 1 номеров. Геометрически K_i составляет совокупность всех граней (любых размерностей) симплекса, содержащих i-ю вершину.
Фактически найденных классов K_i достаточно, чтобы задать толерантность в S_p. Точный смысл этого утверждения состоит в том, что соотношение $x\tau y$ выполняется тогда и только тогда, когда существует класс K_i, содержащий одновременно $x\tau y$. Действительно, если $x\tau y$, то x и y содержат некоторый общий номер i, и тем самым входят в класс K_i. Обратное столь же очевидно. Таким образом, лемма 6.4 допускает для пространства S_p уточнение. Для проверки толерантности достаточно ограничиться проверкой вхождения в один из классов K_i. Меньшим запасом классов ограничиться уже нельзя, поскольку толерантность элементов $[i]$ и $\{i, j\}$ определяется их вхождением именно в класс K_i (см. ниже лемму 6.6). Однако в S_p кроме K_i

есть еще классы толерантности — в указанном смысле избыточные. Так, в S_3 множество {{1,2}, {2,3}, {3, 1}, {1,2,3}} образует класс. (Это проверяется непосредственным перебором.) Ясно, что этот класс не совпадает ни с одним K_i, так как не содержит элементов вида $\{i\}$. Замеченный факт о существовании «необходимых» и «избыточных» классов естественным образом приводит к понятию базиса.

Определение 3.5. Совокупность $H_B=\{K^1, K^2, \ldots\}$ классов в пространстве толерантности $\langle M, \tau\rangle$ называется *базисом,* если 1) для всякой толерантной пары x и y существует класс $K^i \in H_B$, содержащий оба этих элемента: $x \in K^i$, $y \in K^i$; 2) удаление из H_B хотя бы одного класса приводит к потере этого свойства, т. е. для всякого $K^i \in H_B$ существует толерантная пара x, y, для которой K^i является единственным общим классом толерантности в H_B.

Замечание. Произвольная система классов толерантности, обладающая свойством 1) из определения 6.5, содержит базис. Чтобы выделить этот базис, достаточно последовательно удалить «лишние» классы. Правда, для проведения этой процедуры в случае бесконечной системы классов понадобится уже упоминавшаяся трансфинитная индукция.

В качестве исходной системы можно выбрать все множество классов. Отсюда следует существование базиса в любом пространстве толерантности.

Используя понятие базиса, мы сформулируем следующее утверждение:

Теорема 6.3'. *Пусть* $\langle M, \tau\rangle$ *— произвольное пространство толерантности, а* H_B *— базис. Тогда сущевует отображение* $\varphi\colon M \to S_{H_B}$ *такое, что элементы из М толерантны в том и только том случае, когда толерантны их образы в* S_{H_B}.

Эта теорема доказывается почти буквальным повторением доказательства теоремы 6.3, а смысл ее состоит в том, что любое пространство толерантности реализуется (с точностью до склеек) как система множеств классов из базиса с естественной толерантностью типа S_{H_B}.

Выше мы показали, что в пространстве толерантности S_p набор классов $K_1, K_2, \ldots, K_p$ образует базис, не совпадающий с совокупностью всех классов. С. М. Якубович описала все классы толерантности в S_p. Мы не будем приводить здесь это описание, а только установим одно простое свойство этих классов.

Лемма 6.6. *Если К— класс толерантности в* S_p*, содержащий элемент* $\{i\}$*, то* $K = K_i$.

Действительно, все элементы, толерантные к $\{i\}$, обязаны содержать номер i в своем наборе. Значит, $K \subseteq K_i$. Но K есть класс, т. е. по определению не может целиком содержаться в другом классе. Значит, $K = K_i$. Отсюда сразу получается

Лемма 6.7 (С. М. Якубович). *В пространстве S_p существует единственный базис:* $\{K_1, K_2, \ldots, K_p\}$.

Доказательство. Пусть H_B — базис в S_P. Тогда в нем обязан существовать класс, содержащий элемент $\{i\}$. По предыдущей лемме таким классом может быть только K_i. Стало быть, базис H_B должен содержать все классы $K_1, K_2, \ldots, K_p$. Но они уже сами образуют базис, т. е. $H_B = \{K_1, K_2, \ldots, K_p\}$.

В силу определения базиса толерантность в S_P можно задать (как она, впрочем, и задана заранее) только p признаками, соответствующими p базисным классам $K_1, K_2, \ldots, K_p$. При этом не надо думать, в какие паразитические классы входит еще каждый элемент.

Итак, в пространстве S_p остальные классы играют чисто паразитическую роль, не участвуя ни в одном базисе. Вообще говоря, существуют пространства толерантности с неединственным базисом. Такой пример легче всего построить геометрически. Заметим, что в графе, изображающем множество с некоторым отношением толерантности, класс толерантности образует максимальный полный подграф в том смысле, что все вершины, входящие в один класс толерантности, соединены в графе ребрами (поскольку класс является предклассом), а любая другая вершина не соединена ребром хотя бы с одной вершиной из данного класса. На рис. 6.5 легко выделяются группы вершин, образующие максимальные полные подграфы в S.

Это {{1}, {1, 3}, {1, 2}, 1, 2 3}}; {{2}, {2,3}, {1,2}, {1,2,3}}; {{3}, {1,3}, {2,3} {1 2,3}}, соответствующие базисным классам K_1, K_2, K_3, и группа {{1, 3}, {2, 3}, {1, 2}, {1. 2, 3}}, образующая паразитический класс.

Граф на рис. 6.10 изображает бесконечное пространство толерантности — правильную треугольную решетку в которой соседние узлы толерантны между собой.

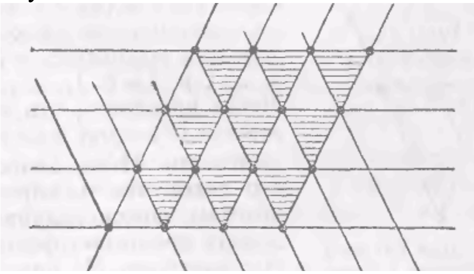


Рис. 6.10. Два базиса.

Классом здесь будет каждый треугольник. Один базис H_B образуют все зачерненные треугольники, а другой H_B - все белые треугольники. Действительно, каждое ребро (т. е. каждая пара *x, y* различных толерантных элементов) принадлежит двум треугольникам — светлому и темному. Поэтому для толерантности пары *x* и *y* необходимо и достаточно чтобы эта пара входила в общий темный (светлый) треугольник.
Пространство толерантности на рис. 6.11 образует конечную вырезку из предыдущего.

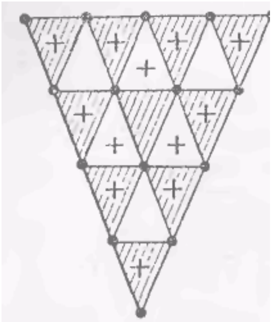


Рис. 6.11. Два базиса с разным числом классов.

В нем есть очевидный базис H^1_B, состоящий из всех затемненных треугольников — всего 10 классов, но в нем можно обнаружить и другой базис H_B, состоящий из всех треугольников, отмеченных крестиками. Этот базис состоит уже из 12 классов. Таким образом, число классов в базисе не инвариантно относительно выбора базиса. Предоставляем читателю проверить, что в примере на рис. 6.11 имеется только два указанных базиса.
Пространство B^m_p. Определение этого пространства дано в 6.1. Оно состоит из целочисленных кортежей $x = \langle \xi_1, \xi_2, \ldots, \xi_p \rangle$ длины *p,* где $0 \leqslant \xi_i \leqslant m-1$.
Обозначим через K^j_i множество, состоящее из всех элементов, для которых $\xi_i = j$ $(i = 1, 2, \ldots, p;$ $j = 0, 1, \ldots, m-1)$
Легко проверить, что эти множества образуют классы толерантности. Итак, класс K^j_i — это совокупность кортежей, у которых фиксированная координата принимает фиксированное значение. Из определения толерантности в B^m_p сразу следует, что классы K^j_i образуют базис. Общее количество этих классов равно *pm,* а каждый

класс содержит m^{p-1} элементов. Менее очевиден тот факт, что в B^m_p существуют и другие классы толерантности).

* * *

Когда отношение толерантности оказывается транзитивным, т. е. превращается в свой частный случай — в отношение эквивалентности, то классы толерантности превращаются, очевидно, в классы эквивалентности. Так как классы эквивалентности не пересекаются, справедлива

Лемма 6.8. *Отношение толерантности х является отношением эквивалентности тогда и только тогда, когда классы толерантности не пересекаются друг с другом.*

Вернемся теперь к изучению отображения φ, построенного в процессе доказательства теоремы 6.3. и выясним, какие элементы из *М* имеют одинаковый образ при отображении φ, т. е. отчего φ бывает не инъективным.

Определение 6.6. Пусть $\langle M, \tau\rangle$— пространство толерантности. Множество $L \subseteq M$ называется *ядром,* если существует такая совокупность классов K^1, K^2, ..., что L есть совокупность всех элементов из *М,* каждый из которых входит во все эти и только эти классы.

Ядра — это прообразы при отображении φ. Действительно, ядро $Я(K^1,\ K^2,\ \ldots)$ состоит из всех тех элементов *х,* для которых образ φ(*х*) есть именно это множество классов толерантности: $\{K^1,\ K^2,\ \ldots\}$.

Отсюда ясно, что непустые ядра образуют разбиение множества *М* и тем самым задают отношение эквивалентности. Мы попробуем разобраться, как это отношение связано с исходной толерантностью.

Пусть задано пространство толерантности $\langle M,\ \tau\rangle$. Далее мы будем обозначать через T_x множество всех элементов, толерантных к *х.* Отношение θ на *М* определим условием

$$x\theta y, \quad \text{если} \quad T_x = T_y. \tag{6.3}$$

Иначе говоря, *х*θ*у* означает, что *х* и *у* толерантны к одним и тем же элементам.

Лемма 6.9. *Для того чтобы выполнялось соотношение х*θ*у, необходимо и достаточно, чтобы х и у лежали в одном и том же ядре* $Я(K^1,\ K^2,\ \ldots)$.

Доказательство. Пусть *х* и *у* принадлежат ядру $Я(K^1,\ K^2,\ \ldots)$. По лемме 6.4 множество T_x состоит из всех элементов, входящих хотя бы в один из классов $K^1, K^2, \ldots: T_x = K^1 \cup K^2 \cup\ \ldots$ Но то же самое справедливо и для множества T_y, т. е. $T_x = T_v$ или $x\theta y$. Обратно. Предположим, что $x\theta y$, и пусть x принадлежит некоторому классу K. Тогда для любого $z \in K$ будет выполнено соотношение $x\tau z$. В силу (6.3) выполнено и $y\tau z$. Значит, $y \in K$ (поскольку K—максимальный предкласс). Аналогично показывается, что всякий класс, содержащий y, содержит одновременно x. Итак, x и y принадлежат одной и той же совокупности классов, а значит, и общему ядру. Лемма доказана. Отсюда вытекает важное

Следствие. *Отношение* θ *есть эквивалентность, а непустые ядра служат для* θ *классами эквивалентности.*

Отметим очевидное включение

$$Я\,(K^1, K^2, \ldots) \subseteq K^1 \cap K^2 \cap \ \ldots \tag{6.4}$$

В случае эквивалентности классы не пересекаются и каждое ядро совпадает со своим классом толерантности:

$$Я\,(K) = K,$$

и, кроме того, для любого $x \in Я\,(K)$

$$T_x = Я\,(K).$$

Заметим, что при обобщении понятия эквивалентности — переходе к толерантности — понятие класса эквивалентности расщепляется на два разных понятия — класс толерантности и ядро. Это довольно часто встречающаяся в математике ситуация — расщепление понятий при переходе от частного понятия к общему.

Определение 6.7. Пространство толерантности $\langle M, \tau \rangle$ называется *безъядерным,* если каждое ядро состоит не более чем из одного элемента.

Примером безъядерного пространства может служить пространство, изображенное на рис. 6.10. Каждая точка принадлежит ровно шести треугольникам — классам толерантности. Одной шестерке примыкающих треугольников соответствует ровно одна точка — определяемое этими классами ядро. Любой другой совокупности треугольников соответствует пустое ядро. Для безъядерных пространств толерантности основная классификационная теорема (теорема 6.3) может быть уточнена так:

Теорема 6.3". *Пусть* $\langle M, \tau \rangle$ *— безъядерное пространство толерантности, а H — множество всех его классов толерантности. Тогда существует инъективное отображение*

$$\varphi\colon\ M \to S_H$$

такое, что элементы из M толерантны в том и только в том случае, когда толерантны их образы в S_H.

Для конечных пространств с нетривиальными ядрами можно применить тот же прием, который был уже использован для задания признаками эквивалентности. А именно, выберем в каждом ядре свою нумерацию. Сопоставим каждому элементу x конечного пространства $\langle M, \tau \rangle$ набор номеров

$$x \to (n_0;\ n_1,\ n_2,\ \ldots,\ n_k),$$

где n_1, n_2,…,n_k — те же самые номера, что и в (6.2), а n_0 — номер элемента в своем ядре. Ясно, что элемент однозначно определяется целочисленными признаками $n_0;\ n_1,\ n_2,\ \ldots,\ n_k,$ а толерантность пары определяется совпадением одного из признаков n_1, n_2,…,n_k.

Пусть теперь $\langle M, \tau \rangle$— произвольное пространство толерантности. Обозначим через $M^Я$ множество его ядер и определим толерантность ядер $Я_1$ и $Я_2$ условием: $Я_1 \tau_Я Я_2,$ если существуют представители $x_1 \in Я_1$ и $x_2 \in Я_2,$ толерантные в $\langle M, \tau \rangle.$ Так как элементы одного ядра имеют общее множество толерантных с ними элементов, то из $Я_1 \tau_Я Я_2$ следует, что для любого $x_1 \in Я_1$ и любого $x_2 \in Я_2$ выполнено $x_1 \tau x_2$. Иначе говоря, если $x_1 \tau x_2,\ x'_1 \theta x_1,\ \ x'_2 \theta x_2,$ то также $x'_1 \tau x'_2.$ Мы получили новое пространство $\langle M^Я, \tau_Я \rangle.$ Можно убедиться, что оно является безъядерным. Ясно также, что $x \tau y$ равносильно $Я(x) \tau_Я Я(y),$ где $Я(x)$ и $Я(y)$ — содержащие эти элементы ядра.

Теперь заметим, что ядра можно было бы определять не с помощью полного запаса классов, а только с помощью классов, принадлежащих некоторому базису H_B. Пусть $\{K'_1,\ K'_2,\ \ldots\}$ — некоторая совокупность классов из базиса H_B. *Ядром* $Я\ (K'_1,\ K'_2,\ \ldots)$ *относительно базиса* H_B мы назовем совокупность всех элементов из M, каждый из которых входит во все эти классы и не входит ни в какие другие классы из данного базиса H_B (ср. с определением 6.6). Верна следующая

Лемма 6.10. *Разбиение множества M на ядра относительно базиса H_B совпадает с разбиением множества M на обычные ядра.*

Доказательство. Буквально повторяя доказательство леммы 6.9, мы получим, что ядра, определенные по бааису H_B, суть классы эквивалентности по θ. Стало быть, они совпадают с исходными ядрами.

Рассмотрим еще раз пространство $B^m{}_p$. Легко видеть, что $K_i^j \cap K_i^k = \varnothing$ при $j \neq k.$ (Один и тот же набор $\langle \xi_1,\ \xi_2,\ \ldots,\ \xi_p \rangle$ не может иметь два разных значения координаты ξ_i) Каждый элемент $x = \langle \xi_1,\ \xi_2,\ \ldots\ \ \ldots,\ \xi_p \rangle$ входит ровно в p классов: $K_1^{\xi_1},\ K_2^{\xi_2},\ \ldots,\ K_p^{\xi_p}.$ Таким образом, все непустые базисные ядра здесь имеют вид $Я\left(K_1^{\xi_1},\ K_2^{\xi_2},\ \ldots,\ K_p^{\xi_p}\right)$ и состоят ровно из одного элемента: пространство $B^m{}_p$ безъядерно. Заметим, что в случае пространства толерантности $B^m{}_p$ включение (6.4) обращается в равенство

$$Я\left(K_1^{\xi_1},\ K_2^{\xi_2},\ \ldots,\ K_p^{\xi_p}\right)=K_1^{\xi_1}\cap K_2^{\xi_2}\cap\ \ldots\ K_p^{\xi_p}.$$

В некоторых случаях оказывается полезной

Теорема 6.5. *Если пространство толерантности* $\langle M,\tau\rangle$ *имеет конечный базис* H_B, *то совокупность всех классов толерантности в* $\langle M,\tau\rangle$ *конечна.*

Доказательство. В силу леммы 6.10 число ядер конечно, т. е. конечно пространство ядер $\langle M^Я,\ \tau_Я\rangle$. Значит, $\langle M^Я,\ \tau_Я\rangle$ имеет конечное число классов толерантности. Но так как $x\tau y$ равносильно $Я(x)\,\tau_Я Я(y)$, то каждый класс толерантности в $\langle M,\ \tau\rangle$ есть объединение ядер, образующих соответствующий класс толерантности в $\langle M^Я,\ \tau_Я\rangle$. Таким образом, совокупность всех классов толерантности в $\langle M,\tau\rangle$ конечна.

Обратим внимание, что ни в формулировке теоремы, ни в ее доказательстве не предполагается, что $\langle M,\tau\rangle$ конечно. Оно и фактически может быть бесконечным за счет бесконечности ядер.

6.4. Дальнейшее исследование структуры толерантностей

Рассмотрим множество M и его покрытие П. Пару $\langle M,\Pi\rangle$ мы будем далее называть *картой.*

Произвольная карта $\langle M,\Pi\rangle$ позволяет ввести на множестве M отношение толерантности τ, определенное условием: $x\tau y$, если существует такое $A\in\Pi$, что одновременно $x\in A$ и $y\in A$. Так определенную толерантность τ мы назовем *толерантностью, порожденную картой* $\langle M,\Pi\rangle$. Очевидно, каждое $A\in\Pi$ является предклассом порожденной толерантности τ. Если $\langle M,\tau\rangle$— пространство толерантности и H — множество всех классов толерантности в этом пространстве, то, в силу леммы 6.4, толерантность, порожденная картой $\langle M,H\rangle$, совпадает с исходной толерантностью τ. Аналогичное утверждение справедливо и для произвольного базиса H_B в пространстве $\langle M,\tau\rangle$.

Карта (M, П) называется *канонической,* если каждый элемент A покрытия П оказывается классом толерантности, порожденной исходной картой $\langle M, \Pi \rangle$. Легко видеть, что если карта $\langle M, \Pi \rangle$ является канонической, то П содержит некоторый базис H_B порожденной толерантности: $\Pi \supseteq H_B$.

На рис. 6.12 слева изображена некоторая карта $\langle M, \Pi \rangle$, а справа — система классов порожденной толерантности (впрочем, в данном случае эта система состоит из одного класса).

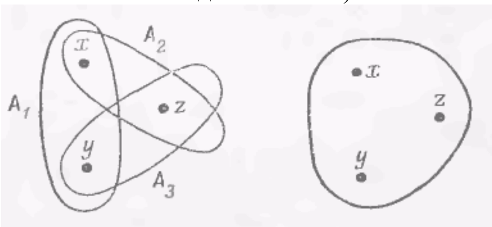


Рис. 6.12.

Этот пример показывает, в частности, существование неканонических карт.

Каждая карта $\langle M, \Pi \rangle$ естественным образом приводит ко всюду определенному соответствию

$$\psi:\ M \to \Pi, \qquad (6.5).$$

которое каждому элементу $x \in M$ сопоставляет все те $A \in \Pi$, для которых $x \in A$. Наоборот, если дано некоторое всюду определенное соответствие $\psi':\ M \to L$, то оно порождает покрытие П множества *M*, состоящее из прообразов элементов из *L* при соответствии ψ′. Таким образом, $A \in \Pi$ тогда и только тогда, когда существует такое $\xi \in L$, что *A* есть множество элементов из *M*, которым соответствие ψ′ сопоставляет ξ. Обозначим для дальнейшего прообраз элемента $\xi \in L$ при соответствии ψ′ через $M(\xi)$. По соответствию (6.5) можно построить отображение

$$\varphi:\ M \to S_\Pi, \qquad (6.6)$$

которое каждому элементу $x \in M$ сопоставляет непустое множество элементов $A \in \Pi$, для которых $x \in A$.

С помощью отображения (6.6) толерантность τ, порожденная исходной картой $\langle M, \Pi \rangle$, выражается условием: *xτy,* если $\varphi(x) \cap \varphi(y) \neq \varnothing$. Можно ввести еще и отношение θ_Π, определяемое условием: $x\theta_\Pi y$, если $\varphi(x) = \varphi(y)$. Отношение θ_Π, очевидно, является эквивалентностью.

В соответствии с ранее принятым словоупотреблением мы будем говорить, что отображение φ сопоставляет элементу *x* множество $\varphi(x) \subseteq \Pi$ его признаков. Тем самым множество П будет интерпретироваться как множество признаков, заданных для объектов из множества *М*. Признаками каждого элемента $x \in M$ являются те множества $A \in \Pi$, для которых $x \in A$. Таким образом, любая карта $\langle M, \Pi \rangle$ есть способ описания системы признаков, заданных на множестве *М.* Высказывание «элемент *х* обладает признаком *А*» равносильно включению $x \in A$. Классы порожденной толерантности называются *каноническими признаками.* Канонические признаки определяются самой толерантностью, а не способом ее задания.

Интересно посмотреть на примерах, как канонические признаки выражаются через исходные признаки карты.

В примере на рис. 6.12 имеем

$$K = A_1 \cup A_2 \cup A_3.$$

В примере на рис. 6.13, *а* изображено соответствие: ψ': $M \to L$, где $L = \{\xi_1, \xi_2, \xi_3\}$, а $M = \{x, y, z, u\}$. На рис. 6.13,*б* изображены классы порожденной толерантности.

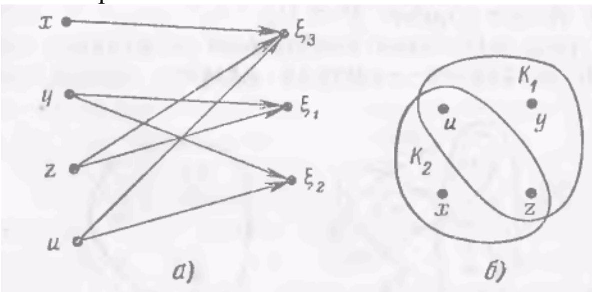


Рис. 6.13.

Легко проверить, что

$$K_1 = M(\xi_1) \cup M(\xi_2), \quad K_2 = M(\xi_3).$$

На рис. 6.14 исходная карта уже является канонической. Но если взять каноническую карту $\langle M, H \rangle$ с полным набором классов толерантности, то получим, что

$$K_4 = (A_1 \cap A_2) \cup (A_2 \cap A_3) \cup (A_3 \cap A_1).$$

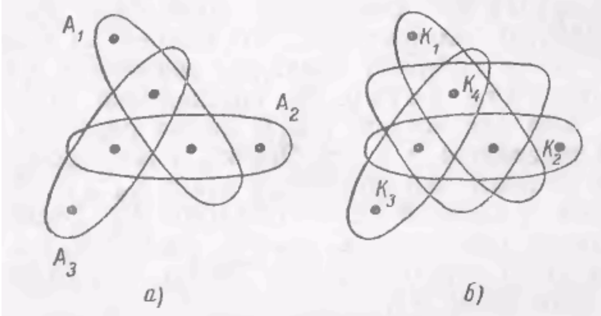


Рис. 6.14.

Мы изучим далее, каким образом и всегда ли канонические признаки могут быть выражены через исходные. Ответ на поставленный вопрос дает

Теорема 6.6. *Для произвольной карты* $\langle M, \Pi \rangle$ *любой класс порожденной толерантности К всегда может быть выражен через элементы покрытия* П *с помощью операций пересечения и объединения.*

Доказательство. Рассмотрим некоторый класс толерантности *К*. Пусть $x \in K$. По определению класса, для всякого $y \in K$, $x\tau y$, а по определению толерантности существует признак $A_{xy} \in \Pi$ такой, что $x \in A_{xy}$ и $y \in A_{xy}$. Тогда 1) $x \in \bigcap_{y \in K} A_{xy}$; 2) $\bigcap_{y \in K} A_{xy} \subseteq K$.

Действительно, 1) следует из того, что $x \in A_{xy}$ для всех признаков A_{xy}, а 2) следует из того, что всякий *z*, принадлежащий A_{xy}, толерантен к *у*. Поскольку *у* — произвольный элемент из *К*, по свойству максимальности класса $z \in K$. Отсюда вытекает, что

$$K = \bigcup_{x \in K} \bigcap_{y \in K} A_{xy}, \tag{6.7}$$

что доказывает теорему.

Подчеркнем, что канонические признаки определяются через исходные без перехода к дополнениям. О связи между исходными и каноническими признаками говорит также

Теорема 6.7. *Существует такой базис классов порожденной толерантности, что каждый из классов этого базиса содержит некоторое множество* $A \in \Pi$.

Доказательство. По определению толерантности в *М* для всякого $A \in \Pi$ любая пара $x \in A$ и $y \in A$ толерантна, т. е. $x\tau y$. Значит, *А* есть предкласс. Тогда по лемме 6.3 существует класс $K_A \supseteq A$. Выберем для каждого *А* одни из классов K_A. Очевидно, выбранная совокупность классов удовлетворяет условию 1) из определения 6.5. Значит, она содержит некоторый базис H_B.

Следствие. *Когда М конечно, то существует базис классов толерантности, число классов в котором не превышает количества исходных признаков.*

В самом деле. Каждому исходному признаку $A \in \Pi$ мы поставили в соответствие некоторый класс K_A. Таким образом, множество этих классов $\{K_A\}$ содержит не больше элементов, чем число признаков A. Выбирая в $\{K_A\}$ базис, мы можем только уменьшить число классов.

Рассмотрим исходную карту $\langle M, \Pi \rangle$ и полученную из нее каноническую карту $\langle M, H_B \rangle$, где H_B — базис.

Как мы уже отмечали, отношения толерантности, задаваемые на множестве объектов M обеими картами, совпадают.

Несколько иначе обстоит дело с отношением эквивалентности θ_Π, задаваемым на M с помощью определения, приведенного в начале параграфа. Пусть θ_Π— отношение эквивалентности, заданное исходным множеством признаков П, а θ — отношение эквивалентности, заданное по (6.3). Как показывает пример на рис. 6.12, отношения θ_Π и θ могут и не совпадать. Именно, для этого примера отношение θ_Π выполненотолько для совпадающих объектов, так как каждому объекту соответствует различный набор исходных признаков. Отношение θ, напротив, выполнено для любой пары объектов.

В общем случае справедлива

Теорема 6.8. *Если выполнено соотношение $x\theta_\Pi y$, то выполнено и соотношение* $x\theta y$, *т. е.* $\theta_\Pi \subseteq \theta$.

Доказательство. Если $x\theta_\Pi y$, то совокупности исходных признаков φ(x) и φ(y), выполненных для x и y, совпадают. Это означает, что, для каждого элемента покрытия A, x и y одновременно содержатся или не содержатся в A. Из теоремы 6.6 (см., в частности, (6.7)) вытекает, что для каждого класса толерантности x и y одновременно содержатся или не содержатся в нем. Таким образом, x и y имеют одинаковые наборы канонических признаков, т. е. $x\theta y$. Теорема доказана.

Следующая теорема, принадлежащая С. М. Якубович, дает условия того, что некоторое множество $A \in \Pi$ является классом толерантности, т. е. того, что некоторый признак является каноническим.

Теорема 6.9. *Пусть имеется карта* $\langle M, \Pi \rangle$. *Для того чтобы элемент покрытия* $A \in \Pi$ *являлся классом порожденной* *толерантности* τ, *необходимо и достаточно, чтобы для любого подмножества* $\Pi_0 \subseteq \Pi$, *из* $A \subseteq \bigcup_{B \in \Pi_0} B$ *следовало бы* $\bigcap_{B \in \Pi_0} B \subseteq A$.

Доказательство. Сначала предположим, что множество $A \in \Pi$ не является классом толерантности.

Так как *A* является предклассом, то единственная причина, по которой *A* может не быть классом, состоит в том, что существует *z*, не входящий в *A* и толерантный ко всем элементам $x \in A$. Значит, для всякого $x \in A$ существует множество $B_x \in \Pi$, содержащее *x* и *z*. Таким образом, множества B_x образуют покрытие множества A: $A \subseteq \bigcup_{x \in A} B_x$. Но все B_x содержат элемент *z*, не входящий в *A*. Следовательно, пересечение $\bigcap_{x \in A} B_x$ не содержится в *A*. Итак, мы доказали достаточность условия, указанного в теореме 6.9. Докажем теперь необходимость. Пусть существует такое подмножество $\Pi_0 \subseteq \Pi$, что $A \subseteq \bigcup_{B \in \Pi_0} B$, но $\bigcap_{B \in \Pi_0} B \not\subseteq A$.

Значит, существует элемент *z*, не входящий в *A*, но входящий во все $B \in \Pi_0$. Этот элемент толерантен ко всем $x \in A$. Значит, *A* не является максимальным предклассом, т. е. не является классом толерантности. Теорема доказана.

Предоставляем читателю применить теорему 6.9 к примерам на рис. 6.12, 6.13, 6.14.

Рассмотрим еще так называемые сопряженные и производные пространства толерантности.

Пусть $\langle M, \tau \rangle$ — произвольное пространство толерантности, и пусть H_0 — некоторая совокупность классов толерантности. Множество H_0 естественным образом превращается в пространство толерантности $\langle H_0, \tau^* \rangle$ при помощи следующего определения: $K\tau^*K'$, если $K \cap K' \neq \varnothing$.

Определение 6.8. Если H_0 совпадает с множеством *H* всех классов, то пространство $\langle H, \tau^* \rangle$ называется *сопряженным* к $\langle M, \tau \rangle$ и обозначается $\langle M^*, \tau^* \rangle$ (таким образом, *H* = *M*)*.

Рассмотрим несколько примеров.

Если τ — полное отношение, то сопряженное пространство состоит из одного элемента.

В пространстве S_p элемент $x_0 = \{1, 2, \ldots, p\}$, содержащий все числа, толерантен ко всем элементам и, стало быть, входит во все классы толерантности. Значит, в пространстве $\langle S_p^*, \tau^* \rangle$ τ^* — полное отношение.

На рис. 6.15 изображен линейный граф из 7 вершин.

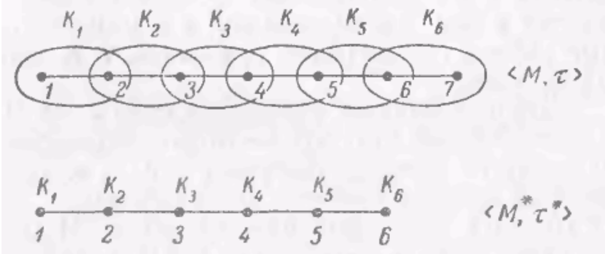


Рис. 6.15. Пространство, сопряженное к линейному.

Классами толерантности являются «ребра», а толерантны классы, соответствующие смежным ребрам. Ясно, что для линейного графа из k вершин сопряженным является линейный граф из k — 1 вершин.
На рис. 6.16 изображен циклический граф. Сопряженным к нему будет циклический граф из того же числа вершин (если количество вершин исходного графа было больше трех).

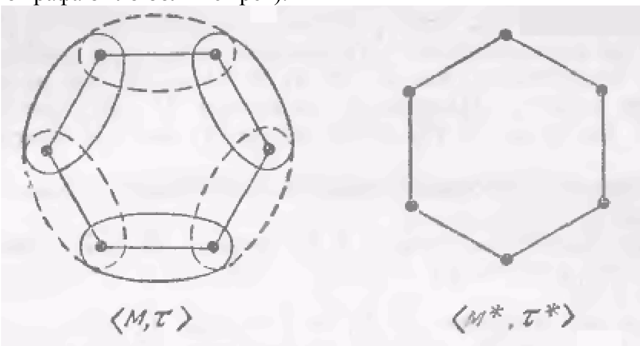


Рис. 6.16. Пространство, сопряженное к циклическому.

На рис. 6.17 изображено пространство толерантности $\langle M, \tau\rangle$, состоящее из двух циклов, зацепленных в одной точке. Сопряженное пространство $\langle M^*, \tau^*\rangle$ состоит из таких же циклов с белее сложным зацеплением.

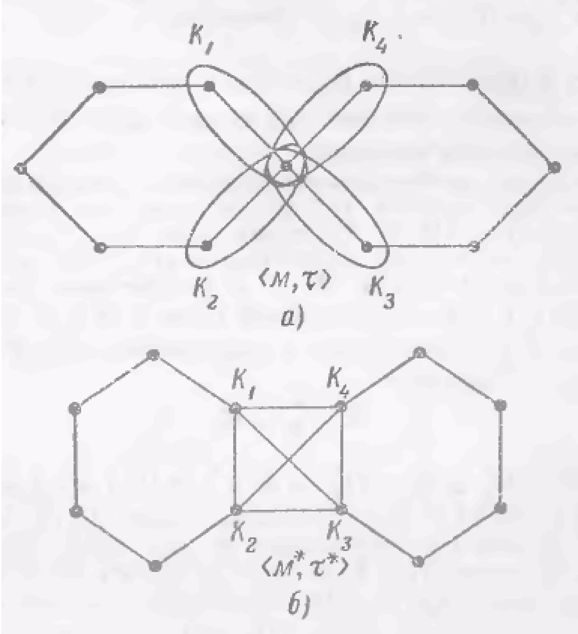


Рис. 6.17. Два зацепленных цикла и сопряженное пространство.

Но сопряженное к последнему пространство $\langle M^{**}, \tau^{**} \rangle$ по существу совпадает с исходным пространством $\langle M, \tau \rangle$. Проверку этого факта мы предоставляем читателю.

Определение 6.9. Пусть H_B — базис. Тогда пространство $\langle H_B, \tau^* \rangle$ называется *сопряженным* к $\langle M, \tau \rangle$ *относительно данного базиса* H_B.

Определение 6.10. Второе сопряженное пространство относительно некоторого базиса H_B в $\langle M, \tau \rangle$ и базиса H_B^* в $\langle H_B, \tau^* \rangle$ называется *производным* от исходного пространства толерантности $\langle M, \tau \rangle$.

Итак, производное пространство толерантности $\langle M', \tau' \rangle$ определяется не однозначно, а с точностью до выбора базисов. Этот произвол исключается, когда $\langle M, \tau \rangle$ и $\langle H_B, \tau^* \rangle$ имеют по единственному базису. (Например, когда все H образует базис в $\langle M, \tau \rangle$, и базис в $\langle H, \tau^* \rangle$ тоже содержит все соответствующие классы.)

Рассмотрим несколько примеров, понятных из предыдущих иллюстраций:

1. Для линейного графа с k вершинами $(k \geqslant 3)$ производное пространство есть также линейный граф, но с k — 2 вершинами (см. рис. 6.15).

2. Для циклического графа нз k вершин $(k \geqslant 4)$ производное пространство толерантности «совпадает» с исходным (см. рис. 6.16).

3. Для зацепленных циклических графов (см. рис. 6.17) производное пространство «совпадает» с исходным пространством.

4. Для пространства S_p производное S'_p состоит из одного элемента.

5. Если в пространстве $B^m{}_1$ выбрать канонический базис $\{k^j{}_i\}$, то $(B^m_p)'$ «устроено» так же, как и само $B^m{}_1$. Проверку этого факта предоставляем читателю.

Приведенные выше примеры наводят на мысль о том, что производное пространство $\langle M', \tau' \rangle$ устроено как «часть» исходного пространства $\langle M, \tau \rangle$. На самом деле это не совсем так.

Точную формулировку соответствующего факта составляет

Теорема 6.10. *Если* $\langle M, \tau \rangle$ *— произвольное пространство толерантности, а* H_B *—произвольный базис в нем, то существуют такой базис* H_B *в сопряженном пространстве* $\langle H_B, \tau^* \rangle$ *и такое инъективное отображение*

$$\delta\colon\ H^*_B \to M,$$

что при $K^*_1 \in H^*_B$ *и* $K^*_2 \in H^*_B$ *из* $\delta(K^*_1)\,\tau\,\delta(K^*_2)$ *следует* $K^*_1 \tau^{**} K^*_2$.

Доказательство. Обозначим через $H_B(x)$ множество классов из базиса H_B, содержащих x. Для любых классов K_1 и K_2 из $H_B(x)$ имеем $K_1 \cap K_2 \neq \varnothing$, т. е. $K_1 \tau^* K_2$. Итак, множества $H_B(x)$ суть предклассы в $\langle H_B, \tau^* \rangle$. Значит, для всякого $x \in M$ существует класс K^*_x в $\langle H_B, \tau^* \rangle$, для которого $H_B(x) \subseteq K^*_x$.

Зафиксируем для каждого x некоторый класс K^*_x и множество этих классов $\{K^*_x\}$ обозначим через $\mathfrak{H}$. Мы имеет теперь сюръективное отображение

$$f\colon\ M \to \mathfrak{H},$$

которое каждому $x \in M$ сопоставляет класс $K^*_x \in \mathfrak{H}$. Покажем, что $\mathfrak{H}$ содержит некоторый базис H^*_B. Действительно, если $K_1 \tau^* K_2$, то существует $x \in M$, содержащийся в K_1 и K_2. Тогда K_1 и K_2 содержатся в $H_B(x)$, а значит, $K_1 \in K^*_x$ и $K_2 \in K^*_x$. Теперь для каждого $K^* \in H^*_B$ выберем ровно один элемент $x \in M$, для которого $f(x)=K^*$. Множество таких элементов обозначим через M_1. Ясно, что $M_1 \subseteq M$ и возникающее при этом сюръективное отображение множества M_1, на H^*_B инъективно. Тогда обратное к нему отображение

$$f^{-1}\colon\ H^*_B \to M_1$$

инъективно отображает H_B^* на подмножество M_1 множества M. Поэтому его можно рассматривать как инъективное (но уже не сюръективное в общем случае) отображение

$$\delta\colon H_B^* \to M.$$

Пусть теперь $K_x^* \in H_B^*$ и $K_y^* \in H_B^*$, где $x = \delta\left(K_x^*\right)$ и $y = \delta\left(K_y^*\right)$ и $x\tau y$. Тогда существует класс K, содержащий x и y. Значит, $H_B(x) \cap H_B(y) \neq \varnothing$. Но из $K_x^* \supseteq H_B(x)$ и $K_y^* \supseteq H_B(y)$ следует, что $K_x^* \cap K_y^* \neq \varnothing$, т. е. $K_x^* \tau^{**} K_y^*$. Теорема доказана.

Отсюда для конечных множеств M следует, что с какого-то номера должна наступить стабилизация и последовательные производные не будут по существу отличаться.

С. М. Якубович доказала, что для любого $\langle \widehat{M}, \widehat{\tau} \rangle$ существует «первообразное» $\langle M, \tau \rangle$ такое, что $\langle M', \tau' \rangle$ «совпадает» с $\langle \widehat{M}, \widehat{\tau} \rangle$.

7. Упорядоченность

7.1. Еще раз о порядке

В этой главе мы рассмотрим некоторые вопросы упорядоченности, которые мы не рассматривали ранее. Речь идет о ситуациях, когда объекты некоторого множества соотносятся по взаимному старшинству, по важности, по «первичности» и т. д. Подобные отношения, по-видимому, не симметричны. Мы начнем с обсуждения содержательных примеров, чтобы понять, какие свойства этих отношений являются настолько существенными и общими, что их следует включить в аксиоматическое определение интересующего нас типа отношений.

Простейшим примером могут служить целые числа. Для любых двух различных целых чисел мы умеем определять, какое из них больше другого. Это случай, когда все объекты строго расставлены по величине.

Вообще говоря, далеко не всегда все объекты можно сравнить друг с другом. Изображения на рис. 7.1 являются геральдическими символами.

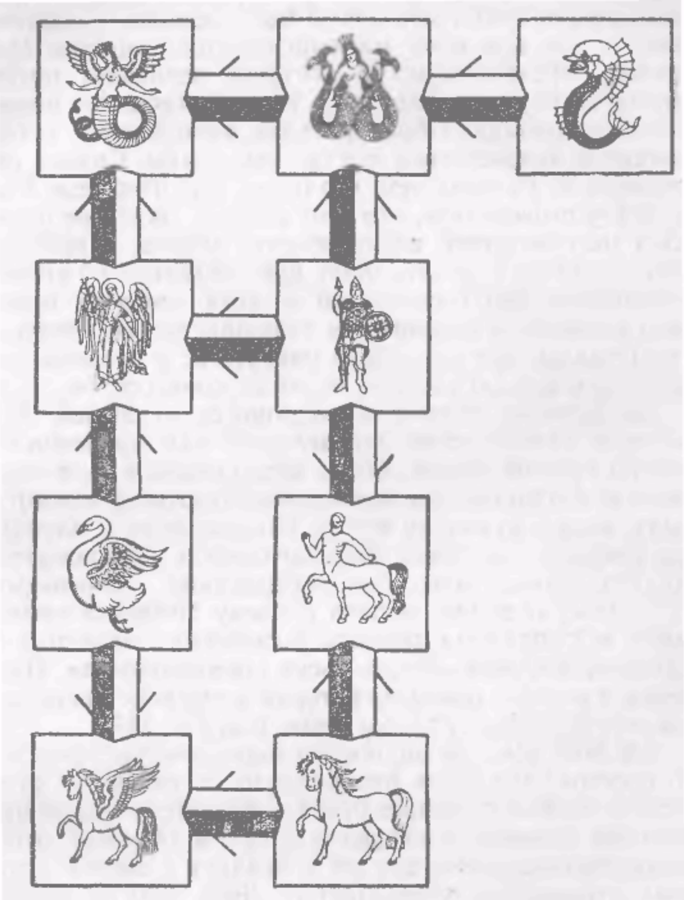


Рис. 7.1. Взаимоотношения мифических образов.

Мы упорядочили эти изображения, пытаясь представить, как могло быть создано представление о том или ином мифологическом существе. Например, представление о кентавре возникает путем смешения образов человека и коня. Пегас несет черты птицы и лошади. Образ русалки явно возник путем придания человеческой фигуре черт рыбы. Сирена отличается от русалки тем, что имеет еще и крылья. Надо сразу оговориться, что этот рисунок никак не отражает исторического возникновения мифов, а призван только иллюстрировать наше представление об упорядоченности. Ясно одно, что в этом примере имеет смысл говорить о взаимном старшинстве

(«первичности») только для некоторых пар. Пегас и русалка, например, в данной системе никак не соотносятся.
Отношение старшинства по вхождению можно определить на множестве структурных формул органической химии. Пусть M— некоторое множество, а 2^M — множество всех его подмножеств. Включение $M_1 \subseteq M_2$ является соотношением, устанавливающим порядок на 2^M.
Упорядоченность на множестве B^m_p всех кортежей длины p, состоящих из целых чисел от 0 до m—1, можно определить следующим образом. Будем говорить, что кортеж $\langle \xi_1, \xi_2, \ldots, \xi_p \rangle$ *старше* кортежа $\langle \eta_1, \eta_2, \ldots, \eta_p \rangle$, если каждая координата первого кортежа не меньше соответствующей координаты второго кортежа: $\xi_i \geqslant \eta_i$, и хотя бы одна из координат при этом фактически больше своей одноименной. Например, в B^4_4 кортеж (1,0,3,2) старше кортежа (1,0,2,2), но несравним с кортежем (1, 1,0,0).
Заметим, что мы всегда имели возможность двоякого введения упорядочения. От нас зависело выбрать, считаем ли мы каждый объект подчиненным самому себе (как в случае нестрогого неравенства ≤ или нестрогого включения $\subseteq$), или, наоборот, считаем, что объект не может быть старше самого себя (как в случае строгих неравенств < и включений $\subset$). Поэтому нам придется ввести два варианта аксиоматических определений — для строгой и нестрогой упорядоченности. Впрочем, как мы увидим, строгая и нестрогая упорядоченность весьма просто связаны между собой.
Сначала мы разберем случай со строгой упорядоченностью. Мы примем за основу

Определение 7.1. Отношение A на множестве M называется *отношением строгого порядка* (или *строгим порядком),* если оно антирефлексивно и транзитивно.

Примерами строгого порядка могут, очевидно, служить отношение < для целых или вещественных чисел и отношение включения $\subset$ для множеств.

Теорема 7.1. *Если отношение A есть отношение строгого порядка, то оно асимметрично.*

Доказательство. Предположим противное. Пусть $A \cap A^{-1}$ непусто, т. е. существует пара элементов $\langle x, y \rangle$ из множества M таких, что одновременно xAy и $xA^{-1}y$. Иначе говоря, xAy и yAx. По транзитивности отсюда вытекает xAx, что противоречит антирефлексивности.

Итак, отношение строгого порядка на множестве M обладает следующими свойствами:

1) ни для какого $x \in M$ не выполнено xAx;

2) если xAy и yAz, то выполнено xAz, и
3) если выполнено xAy, то невозможно yAx.

Первые два свойства образуют определение строгого порядка, а третье из них следует.

Если A — отношение строгого порядка, то граф отношения A не содержит контуров (*контур* (в ориентированном графе) — это такая последовательность вершин $x_0, x_1, x_2, \ldots, x_n$, что $x_n = x_0$ и от вершины x_i к вершине x_{i+1} проходит стрелка. Частным случаем контура является петля ($n = 1$).). Обратно. Пусть мы имеем граф без контуров. Определим на множестве M вершин этого графа отношение A: xAy, если существует путь по направлению стрелок, ведущий из x в y. Легко видеть, что ввиду отсутствия контуров отношение A является отношением строгого порядка.

Множество M с заданным на нем отношением строгого порядка A, т. е. пару $\langle M, A \rangle$, естественно называть *упорядоченным множеством.*

Определение 7.2. Отношение строгого порядка A называется *совершенным строгим порядком,* если для всякой пары не совпадающих элементов x и y из M верно либо xAy, либо yAx.

В силу теоремы 7.1 последние два соотношения не могут выполняться одновременно. Таким образом, если на множестве M задано отношение совер шенного строгого порядка A, то на множестве M^2 всех пар возникает разбиение на три класса: класс пар вида $\langle x, x \rangle$, класс пар $\langle x, y \rangle$ таких, что xAy, и класс пар $\langle x, y \rangle$ таких, что yAx.

Например, если множество M — прямая линия с отношением $<$, то M^2— это плоскость пар $\langle x, y \rangle$. Класс пар вида $\langle x, x \rangle$— это диагональная прямая $y = x$, класс пар $\langle x, y \rangle$ таких, что $x < y$, состоит из точек, лежащих выше диагонали, а класс пар $\langle x, y \rangle$ таких, что $y < x$,— из точек, лежащих ниже диагонали (рис. 7.2).

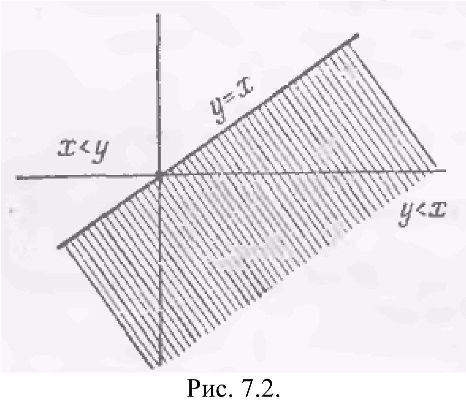


Рис. 7.2.

Опишем структуру конечных множеств с совершенным строгим порядком.

Теорема 7.2. *Пусть дано отношение совершенного строгого порядка < на конечном множестве М. Тогда на М можно выбрать такую нумерацию* $M = \{x_1, x_2, \ldots, x_n\}$, *что соотношение* $x_i < x_j$ *будет выполняться в том и только том случае, когда i < j.*

Предварительно установим, что справедлива

Лемма 7.1. *Если на конечном (непустом) множестве М задан совершенный строгий порядок <, то существует единственный элемент* $x \in M$ *такой, что для всякого у из М, не совпадающего с х, выполнено соотношение х < у.*

(Элемент *х,* обладающий указанным свойством, называется *наименьшим элементом* в упорядоченном множестве *(М,* <)).

Доказательство леммы. Возьмем произвольный элемент $y_0 \in M.$ Если y_0 — наименьший, то существование искомого элемента доказано. Если нет, то поскольку < — совершенный строгий порядок, существует такой элемент $y_1 \neq y_0$, что $y_1 < y_0$. Опять-таки либо y_1 — наименьший, либо существует $y_2 \neq y_1$ такой, что $y_2 < y_1$. Будем продолжать этот процесс. Предположим, что уже выбрано *п* + 1 элементов, для которых

$$y_n < y_{n-1},\ y_{n-1} < y_{n-2},\ \ldots,\ y_1 < y_0.$$

В силу транзитивности ясно, что $y_i < y_j$ при $i > j$. Значит, в силу антирефлексивности, все выбранные элементы попарно не равны. Стало быть, ввиду конечности множества *М* процесс выбора должен оборваться на некотором конечном шаге. Элемент y_n, выбранный на последнем шаге, будет, очевидно, искомым. Итак, для любого $z \neq y_n$ выполнено $y_n < z$. Покажем, что этот элемент единствен. В самом деле, пусть существует другой элемент y_n' такой, что, для всякого $z \neq y_n',\ \ y_n' < z.$ Тогда одновременно выполняется $y_n < y_n'$ и $y_n' < y_n$, что невозможно в виду асимметричности. Лемма доказана.

Заметим, что если на *М* задан совершенный строгий порядок, то на любом непустом подмножестве *Q* множества *М* естественно возникает совершенный строгий порядок, и, стало быть, в *Q* (если оно конечно) существует единственный наименьший элемент.

Теперь перейдем к доказательству теоремы.

Пусть x_1 — наименьший элемент во множестве *М,* выбранный согласно лемме 7.1. Обозначим через M_1 множество $M \setminus \{x_1\}$. Обозначим через x_2 наименьший лемент множества M_1. Ясно, что $x_1 < x_2$. Выкинем из M_1

элемент x_2 и оставшееся множество обозначим через M_2. Его наименьший элемент x_3 удовлетворяет условию: $x_2 < x_3$. Процедура нумерации уже ясна: перебирая по указанному методу последовательно все элементы из *M,* мы их выстроим в цепочку:

$$x_1 < x_2 < \ldots < x_p,$$

где *р* — количество элементов в *М.* В силу транзитивности и асимметричности ясно, что $x_i < x_j$ в том и только том случае, когда $i < j$. Теорема доказана.

Эта теорема в сущности означает, что любой совершенный строгий порядок на конечном множестве *М* равносилен обычному порядку на некотором отрезке натурального ряда.

Рассмотрим некоторое множество *М* из каких-то твердых тел (предметов). Будем говорить, что $x<y$, если предмет *x* весит меньше предмета *у.* Это — довольно типичный пример определения порядка. Опишем теперь соответствующий общий прием.

Пусть на множестве *М* определена инъективная функция

$$f\colon\ M \to \mathbf{R},$$

принимающая вещественные числовые значения (**R** — множество вещественных чисел). Зададим отношение < на *М* условием: $x < y$, если $f(x) < f(y)$. Так определенное отношение < аитирефлексивно, так как не может быть $f(x) < f(x)$. Транзитивность отношения < столь же очевидна. Наконец, для любой пары различных элементов *x, y* из *М* верно либо $f(x) < f(y)$, либо $f(y) < f(x)$, так как *f* — инъекция. Значит, порядок < является совершенным. Функция *f* взаимно-однозначно отображает наше множество *М* на некоторое подмножество множества **R** вещественных чисел, так что соотношение $x < y$ для любых элементов множества *М* равносильно неравенству $f(x) < f(y)$. Например, когда функция *f* сопоставляет предмету *x* его вес *f(x),* мы получаем описанный выше порядок.

Если порядок на конечном множестве *М* не является совершенным, то, очевидно, элементы этого множества нельзя перенумеровать так, чтобы большим номерам соответствовали старшие элементы.

Определение 7.3. Пусть на множестве *М* задано отношение строгого порядка < Тогда элемент $x \in M$ называется *минимальным (максимальным)* в упорядоченном множестве $\langle M, < \rangle$, если не существует никакого элемента *у*, для которого $y < x$ (соответственно $y > x$).

Если, как обычно, в случае $x < y$ проводить стрелку от *x* к *у,* то в графе отношения минимальный элемент — это тот, в который не входят стрелки, а максимальный— из которого не выходят стрелки.

В случае совершенного строгого порядка минимальный элемент *x* обладает тем дополнительным свойством, что для всякого $y \neq x$

выполнено $x < y$. Тем самым для случая совершенных порядков понятие минимального элемента совпадает с понятием наименьшего элемента. В общем случае может оказаться так, что элемент x минимален, но не находится в соотношении $x < y$ с какими-то иными элементами. Элементы x и y называют *сравнимыми* в данном упорядоченном множестве $\langle M, < \rangle$, если $x < y$, или $x = y$, или $y < x$.

Определение 7.4. Пусть на множестве M задано отношение строгого порядка <. Подмножество $Q \subseteq M$ называется *максимальным совершенным,* если 1) отношение < задает на Q совершенный строгий порядок и 2) на любом подмножестве R_1 множества M таком, что $R_1 \supset Q$, отношение < уже не является совершенным строгим порядком.

Теорема 7.3 (Хаусдорф). *Пусть* $\langle M, < \rangle$ *— упорядоченное множество. Для любого элемента* $y \in M$ *существует максимальное совершенное подмножество Q множества M, содержащее y.*

Доказательство. Мы проведем доказательство для конечного множества M. Однако — с помощью аксиомы Цермело — это доказательство может быть проведено и для бесконечных множеств. Пусть множество Q_1 состоит из исходного элемента y. Очевидно, отношение < на Q_1 является совершенным строгим порядком (график отношения < на Q_1 пуст). Если Q_1 уже является максимальным совершенным, то теорема доказана. Предположим, что мы построили множество Q_n, на котором отношение < является совершенным строгим порядком. Если оно максимально, то теорема доказана. Если нет, то существует некоторый элемент из M, сравнимый со всеми элементами из Q_n Присоединив его к Q_n, мы получим множество $Q_{n+1} \supset Q_n$ с совершенным строгим порядком. Из-за конечности самого M этот процесс оборвется на конечном шаге, и мы получим искомое максимальное совершенное множество $Q \ni y$.

Теорема 7.4. *Если < — отношение строгого порядка на конечном множестве M, то для любого элемента* $y \in M$ *существует минимальный элемент* $x \in M$ *такой, что* $x < y$ *или* $x = y$.

Доказательство. Если y — минимальный элемент, то $x = y$. В противном случае существует такой элемент z, что $z < y$. Если z— минимальный элемент, то $x = z$. В противном случае существует такой элемент u, что $u < z$, и т. д. Поскольку M — конечное множество, через конечное число шагов наша «убывающая цепочка» $y > z > u > \ldots$ оборвется на искомом элементе. Теорема доказана.

В этой теореме конечность множества M уже существенна, так как, например, во множестве всех целых чисел, упорядоченных по возрастанию, нет минимального элемента. Однако существует класс отношений порядка на бесконечных множествах, для которого теорема о существовании минимальных элементов тоже может быть доказана.

Следующий кусок написан для читателей, знакомых с элементами теоретико-множественной топологии.

Пусть на множестве M заданы отношение строгого порядка < и некоторая топология. Порядок < будем предполагать непрерывным относительно данной топологии. Это означает следующее. Пусть $Q \subseteq M.$ Элемент $x \in M$ называется *нижней (верхней) границей* множества Q, если для всякого элемента $y \in Q$ либо $x < y$, либо $x = y$ (либо $y < x$, либо $y = x$). Через $\overline{Q}$ будем обозначать замыкание множества Q. Порядок < называется *непрерывным* относительно данной топологии, если любая нижняя и любая верхняя граница произвольного $Q \subseteq M$ являются нижней и соответственно верхней границей и для его замыкания $\overline{Q}.$

Быть может, более естественно было бы определить непрерывность отношения порядка условием, что график отношения при объединении с диаюиалью будет замкнут на $M \times M$. Легко показать, что из такого определения наше вытекает.

Например, естесвенный порядок на числовой прямой непрерывен относительно естественной топологии этой прямой.

Лемма 7.2. *Если порядок непрерывен, то множество R_x всех элементов $y \in M,$ для которых либо $y < x$, либо $x = и$, является замкнутым.*

Действительно, по определению x является верхней границей для R_x; в силу непрерывности x является верхней границей и для $\overline{R}_x.$ Возьмем произвольный $y \in \overline{R}_x.$ Тогда либо $y < x$, либо $y = x$.

В обоих случаях $y \in R_x.$ Значит, $\overline{R}_x \subseteq R_x.$ Но всегда $\overline{R}_x \supseteq R_x.$ Следовательно, $\overline{R}_x = R_x.$

Лемма 7.3. *Если порядок непрерывен, то любое максимальное совершенное множество Q замкнуто.*

Доказательство. В силу того, что на Q порядок совершенный, для любого $x \in Q$ множество Q может быть разбито на две части $Q = Q_x^+ \cup Q_x^-.$ Здесь Q_x^+ — множество тех элементов $y \in Q,$ для которых $y < x$ либо $y = x$, а Q_x^- —множество тех элементов y, для которых $x < y$ либо $x = y$ (пересечение $Q_x^+ \cap Q_x^-$ состоит из одного элемента x). Так как замыкание объединения равно объединению замыканий, то

$$\overline{Q} = \overline{Q}_x^+ \cup \overline{Q}_x^-.$$

С другой стороны, $Q_x^+ \subseteq R_x$ и $\overline{Q}_x^+ \subseteq \overline{R}_x.$ Таким образом, по лемме 7.2 $\overline{Q}_x^+ \subseteq R_x.$ Таким образом, любой элемент $y \in \overline{Q}_x^+$ либо совпадает с x, либо удовлетворяет соотношению $y < x$. Аналогично показывается, что любой эчемент $z \in \overline{Q}_x^-$ либо совпадает с x, либо удовлетворяет соотношению $x < z$. Итак, любой элемент из замыкания $\overline{Q}$ сравним с x. Это утверждение справедливо для любого $x \in Q$. Итак, любой элемент $w \in \overline{Q}$ сравним с любым элементом $x \in Q$. Следовательно, если бы существовал элемент, принадлежащий множеству $\overline{Q} \setminus Q$, то этот элемент можно бпо бы добавить к множеству Q, сохранив совершенный порядок. Но это сделать невозможно в силу максимальности Q. Стало быть, $\overline{Q} \subseteq Q$, а значит $\overline{Q} = Q$. Лемма доказана.

Из этих лемм следует, что пересечения $F_x = R_x \cap Q$—замкнутые множества. Если $x_1, x_2, \ldots$ — элементы из Q, то пересечение любой конечной группы этих множеств $F_{x_1} \cap F_{x_2} \cap \ldots \cap F_{x_n}$ не пусто. Действительно, поскольку $\{x_1, x_2, \ldots, x_n\} \subseteq Q$, порядок $<$ на $\{x_1, x_2, \ldots, x_n\}$ — совершенный. Так как $\{x_1, x_2, \ldots, x_n\}$— конечное множество, в нем есть наименьший элемент. Пусть этот наименьший элемент есть x_1. Тогда ясно, что $F_{x_1} \cap F_{x_2} \cap \ldots \cap F_{x_n} = F_{x_1}$; следовательно, интересующее нас пересечение непусто, так как содержит элемент x_1. Итак, система множеств $\{F_x\}$ $(x \in Q)$ является центрированной системой замкнутых множеств.

Теорема 7.5. *Пусть М — компактное топологическое пространство, а < — непрерывный порядок на нем. Тогда для любого элемента $y \in M$ существует минимальный элемент x_0 такой, что $x_0 < y$ либо $x_0 = y$.*

Доказательство. По теореме 7.3 существует максимальное совершенное множество $Q \subseteq M$, содержащее y. По одному из определений компактного пространства пересечение системы множеств $\{F_x\}$ $(x \in Q)$ непусто. Пусть x_0 — элемент этого пересечения. Так как $x_0 \in$ Q, то x_0 сравним с y. Покажем, что x_0 — минимальный элемент множества Q. Действительно, если существует $z \in Q$, для которого $z < x_0$, то R_z не содержит элемента x_0 и, следовательно, F_z не содержит x_0, т. е. x_0 не входит в пересечение всех F_x. Итак, x_0 является минимальным элементом множества Q. Значит, $x_0 < y$ или $x_0 = y$. Но если бы x_0 не был минимальным элементом множества M, то нашелся бы $w \in M$ такой, что $w < x_0$. Этот элемент w можно было бы присоединить к Q, не нарушая совершенства порядка. В силу максимальности Q это невозможно. Итак, x_0 есть минимальный элемент в M, причем $x_0 < y$ или $x_0 = y$. Теорема доказана.

Легко получить следующее обобщение этой теоремы:

Теорема 7.5'. *Пусть M — топологическое пространство, а $<$ — непрерывный порядок на нем. Тогда, если любое множество R_x* всех элементов $y \in M$, для которых $y < x$ или $y = x$, *компактно, то для любого $y \in M$ существует минимальный элемент x_0 такой, что $x_0 < y$ или $x_0 = y$.*

Теперь перейдем к изучению нестрогих порядков. Введем следующее

Определение 7.5. Отношение A на множестве M называется *отношением нестрогого порядка* (или *нестрогим порядком),* если оно может быть представлено в виде

$$A = A_1 \cup E, \tag{7.1}$$

где Ai — строгий порядок на M, а E — диагональное отношение.

Отсюда следует, что отношение нестрогого порядка рефлексивно. Легко проверить, что оно и транзитивно. Однако, в отличие от строгого порядка, оно не асимметрично, а только антисимметрично. Более того, $A \cap A^{-1} = E$. В самом деле, из (7.1)

$$A \cap A^{-1} = (A_1 \cup E) \cap (A_1^{-1} \cup E) = \\ = (A_1 \cap A_1^{-1}) \cup (A_1 \cap E) \cup (A_1^{-1} \cap E) \cup E.$$

В силу свойств строгого порядка (в частности, в силу того, что A_1^{-1} также будет строгим порядком) все члены в скобках суть пустые множества.

Любое отношение нестрогого порядка рефлексивно, антисимметрично и транзитивно. Легко видеть, что если A рефлексивно, антисимметрично и транзитивно, то A — нестрогий порядок, так как $A = (A \setminus E) \cup E$, а $A \setminus E = A_1$— строгий порядок. Таким образом, нестрогий порядок можно было бы ввести аксиоматически как рефлексивное, транзитивное и антисимметричное отношение. Ни одно из этих свойств не следует из других, что легко проверить соответствующими примерами.

Нестрогий порядок A мы назовем *совершенным,* если для любой пары x, y верно либо xAy, либо yAx. Из антисимметричности нестрогого порядка следует, что одновременное выполнение xAy и yAx означает совпадение $x = y$. Легко проверить, что справедлива

Лемма 7. 4. *Если A — совершенный нестрогий порядок, то $A_1 = A \setminus E$ есть совершенный строгий порядок. Обратно, если A_1 — совершенный строгий порядок, то $A = A_1 \cup E$ есть совершенный нестрогий порядок.*

Введем такое

Определение 7.6. Отношение A на множестве M называется *отношением квазипорядка* (или *квазипорядком),* если оно рефлексивно и транзитивно.

Очевидно, отношения квазипорядка являются обобщением отношений эквивалентности и одновременно обобщением отношений нестрогого порядка. Пусть теперь квазипорядок *А* является одновременно эквивалентностью и нестрогим порядком. Предположим, что выполнено xAy и $x \neq y$. Тогда по симметричности эквивалентности верно *yAx*. С другой стороны, в силу антисимметричности нестрогого порядка *yAx* не выполняется. Отсюда вытекает

Лемма 7.5. *Если отношение А есть одновременно эквивалентность и нестрогий порядок, то оно есть отношение равенства.*

Пример. Пусть имеется отображение

$$f\colon M \to \mathbf{R},$$

где **R** — множество всех вещественных чисел (числовая ось). Введем на *М* отношение *А* условием:

$$xAy, \quad \text{если} \quad f(x) \leqslant f(y).$$

Ясно, что *А* рефлексивно, так как $f(x) \leqslant f(x)$. Транзитивность отношения *А* видна из следующего рассуждения: если xAy и yAz, то $f(x) \leqslant f(y)$ и $f(y) \leqslant f(z)$, а значит, и $f(x) \leqslant f(z)$, т. е. *xAz*. Если $x \neq y$ и $f(x) = f(y)$, то *xAy* и *y Ax*. Таким образом, если отображение *f* не инъективно, то *А* не антисимметрично. Очевидно, для любой пары *x* и *y* выполнено либо $f(x) \leq f(y)$, либо $f(y) \leqslant f(x)$, т. е. либо *xAy*, либо *yAx*.

Теперь покажем, что каждый квазипорядок порождает некоторый порядок. Для этого нам нужна

Теорема 7.6. *Если А — квазипорядок, то отношение* $B=A\cap A^{-1}$ *есть эквивалентность.*

Доказательство. Рефлексивность и транзитивность отношения *В* вытекает из известных лемм. Докажем симметричность отношения *В*. Пусть выполнено *xBy*. Это значит, что одновременно выполнены соотношения *xAy* и *yAx*. Но это равносильно выполнению соотношений yAx и $yA^{-1}x$, т. е. $y\, A\cap A^{-1}\, x = yBx$. Значит, *В* симметрично. Лемма доказана.

Пусть *А* — квазипорядок на множестве *М*. Обозначим через $\mathfrak{M}$ совокупность классов эквивалентности по отношению $B = A \cap A^{-1}$. Будем говорить, что два класса *X* и *Y* из $\mathfrak{M}$ находятся в отношении *А**, если в этих классах можно выбрать по представителю $x \in X$ и $y \in Y$, так что выполнено *xAy*. Мы будем говорить, что отношение *А** *индуцируется* квазипорядком *А*.

Теорема 7.7. *Отношение А* на множестве классов эквивалентности* $\mathfrak{M}$, *индуцированное квазипорядком А, является нестрогим порядком.*

Доказательство. Рефлексивность отношения *А** следует из того, что для любого класса Х и любого представителя $x \in X$ верно *xAx*, а следовательно, справедливо XA^*X. Транзитивность проверяется немного сложнее. Пусть для классов верны соотношения XA^*Y и $YA^*Z.$ Это значит, во-первых, что для некоторых представителей $x \in X,\ y_1 \in Y$ выполнено соотношение xAy_1 и, во-вторых, что для некоторых представителей $y_2 \in Y,\ z \in Z$ выполнено соотношение $y_2Az.$ Поскольку $y_1 \in Y$ и $y_2 \in Y,$ имеем y_1By_2 и, значит, $y_1Ay_2.$ Из $xAy_1,$ y_1Ay_2 и y_2Az по транзитивности квазипорядка A получаем $xAz.$ Значит, для классов $XA^*Z.$ Наиболее нетривиальным является доказательство антисимметричности отношения A^*. Пусть выполнено $XA^*Y.$ Это значит, что для некоторых представителей $x \in X$ и $y \in Y$ верно

$$xAy. \tag{7.2}$$

Предположим, что одновременно верно $YA^*X,$ т. е. существуют такие представители $x' \in X$ и $y' \in Y,$ что

$$y'Ax'. \tag{7.3}$$

По определению класса эквивалентности $y\, A \cap A^{-1}\, y'.$ Тогда по транзитивности из (7.2) и yAy' следует

$$xAy'. \tag{7.4}$$

С другой стороны, из (7.3) и $x'Ax$ вытекает

$$y'Ax = xA^{-1}y'. \tag{7.5}$$

Сравнивая (7.4) и (7.5), получаем

$$x\left(A \cap A^{-1}\right)y',$$

т. е. x и y принадлежат общему классу по $A \cap A^{-1}.$ Значит, $X \cap Y \neq \varnothing$ и, следовательно, $X = Y$, что доказывает антисимметричность отношения A^*. Тем самым наша теорема доказана.

Итак, по квазипорядку на множестве M можно сконструировать нестрогий порядок, «склеив» некоторые объекты из M.

Для предыдущего примера квазипорядка A, задаваемого числовой функцией f на M, элементами множества $\mathfrak{M}$ служат такие множества, где функция f принимает фиксированное значение. Такие множества обычно называются *областями уровня.* Порядок A^*, индуцированный квазипорядком A на множестве $\mathfrak{M},$ определяется условием: $E \leqslant E'$ $(E \leqslant E'$ означает, конечно, $EA^*E'),$ если для любого $x \in E$ и для любого $x' \in E'$ имеем $f(x) \leqslant f(x').$

Пусть, в частности, множество M есть множество точек на топографической карте, а квазипорядок задается условием: $x \leqslant y,$ если высота $f(x)$ точки x над уровнем моря не превосходит высоту $f\,(y)$ точки y над уровнем моря. Тогда элементами множества $\mathfrak{M}$ являются горизонтали, а порядок A^* совпадает с порядком «отметок высоты» на этих горизонталях. Если квазипорядок A был *совершенным* (то есть для любых x и y либо xAy либо yAx), то легко убедиться, что порядок A^* на классах также будет совершенным. В самом деле, возьмем два произвольных класса: X и Y — и в них два произвольных представителя: $x \in X$ и $y \in Y$. Поскольку A — совершенный квазипорядок, выполнено по крайней мере одно из соотношений: xAy или yAx. Значит, верно либо XA^*Y, либо YA^*X.

В заключение параграфа рассмотрим пример. Пусть M — множество ситуаций, между которыми требуется произвести выбор. Например, множество мест возможной работы. (Разумеется, можно подставить сюда десятки других примеров разной степени серьезности.) В теории исследования операций существует следующая рекомендация, как выполнить обоснованный выбор. Сопоставим каждому месту работы набор признаков. Например, (1) расстояние от места жительства, (2) творческая удовлетворенность, (3) зарплата, (4) перспективы роста, (5) наличие интересных коллег. Каждому из этих факторов сопоставим вес, отражающий наше представление о значении данного фактора. Скажем, веса 30, 10, 40, 10, 10 означают, что мы ищем близкую прибыльную работу, а веса 20, 30, 10, 10. 30 выражают наше стремление найти работу, дающую максимальное удовлетворение, не забывая при этом о минимуме жизненных удобств. Затем опишем каждое из предполагаемых мест, давая ему оценки по всем показателям так, чтобы максимальная оценка не превосходила веса, который мы уже приписали данному фактору. В следующей таблице мы приводим возможною расстановку весов для пяти предполагаемых мест работы.

Признак	Максимальный вес	I	II	III	IV	V
(1) Расстояние от дома	10	5	10	0	10	5
(2) Творческая удовлетворенность	30	20	10	30	15	20
(3) Зарплата	10	10	10	0	5	0
(4) Перспективы роста	20	20	15	20	10	15
(5) Интересные коллеги	30	10	15	15	5	10
Суммарная оценка . . .	100	65	60	65	45	50

Итак, мы на множестве возможных ситуаций M задали оценочную функцию f, которая определяет совершенный квазипорядок на M. По теореме 7.7, склеив равноценные ситуации (в нашем случае I и III), мы

получим совершенный нестрогий порядок. Стало быть, мы можем найти оптимальный класс ситуаций. В этом классе мы можем выбирать случайно — скажем, бросая монетку. Это все очень хорошо, так как дает уверенность в обоснованности выбора. Но, с другой стороны, этот метод навязывает совершенный порядок там, где его по существу нет. Скажем, в уашем примере довольно ясно, что I и III места работы, набравшие одинаковые веса и формально равноценные, совсем не равноценны с точки зрения нашего выбора. Эти места работы существенно различны (одно — лучше по одним факторам, другое—но другим), и нам надо опять задать себе вопрос, чего мы хотим в действительности. Здесь математическая модель явления создала иллюзию простоты ситуации там, где ее в действительности ует. Таким образом, с числовыми оценками реальных явлений следует обращаться осторожно. Это не компрометирует сам метод весовых оценок — из него видно, что IV место работы заведомо не стоит рассматривать.
Но применять подобные оценки можно, лишь понимая их ограниченность и грубость. В реальных ситуациях выбора обычно нет совершенного порядка. Вводя этот порядок в модели, нужно отдавать себе отчет в степени допускаемого произвола.

7.2. Операции над отношениями порядка

Начнем опять с простейшей операции A^{-1}. Из известных лемм вытекает

Теорема 7.8. *Если отношение A является строгим порядком (нестрогим порядком, квазипорядком), то и отношение A^{-1} является строгим порядком (соответственно нестрогим порядком, квазипорядком).*

Легко проверить также, что если отношение A является совершенным строгим порядком (совершенным нестрогим порядком, совершенным квазипорядком), то и отношение A^{-1} является совершенным строгим порядком (соответственно совершенным нестрогим порядком, совершенным квазипорядком).

Из известных лемм вытекает

Теорема 7.9. *Если A и B — строгие порядки (нестрогие порядки, квазипорядки), то пересечение* $A \cap B$ *также является строгим порядком (соответственно нестрогим порядком, квазипорядком).*

Замечание. Пусть A — строгий порядок, а B — нестрогий порядок. Тогда $B = B_1 \cup E$, где B_1 — строгий порядок. Поскольку

$$A \cap B = A \cap (B_1 \cup E) = (A \cap B_1) \cup (A \cap E) = A \cap B_1,$$

пересечение строгого и нестрогого порядка есть строгий порядок.

Свойство «быть совершенным порядком» не обязано сохраняться при пересечении. Это проще всего увидеть из следующих соображений. Пусть *А* — совершенный порядок (строгий или нестрогий), тогда $A \cap A^{-1} = \varnothing$ (или =*Е*). Значит, $A \cap A^{-1}$ на множестве более чем из одного элемента не является совершенным порядком.

Объединение порядков в общем случае не является порядком. Это хорошо видно на таком примере. Пусть *А* — совершенный нестрогий порядок, тогда A^{-1} — есть отношение того же типа. Однако объединение $A \cup A^{-1}$ есть полное отношение, и, следовательно, не является порядком. Условие, когда объединение порядков есть снова порядок, дает

Теорема 7.10. *Если А и В — строгие порядки, то объединение* $A \cup B$ *является строгим порядком в том и только том случае, когда*

$$BA \cup AB \subseteq A \cup B. \tag{7.6}$$

Доказательство. Антирефлексивность объединения вытекает из известной леммы. Достаточно убедиться, что условие (7.6) равносильно транзитивности объединения. В самом деле, транзитивность отношения $A \cup B$ означает, что $(A \cup B)(A \cup B) \subseteq A \cup B,$ или что $A^2 \cup B^2 \cup BA \cup AB \subseteq A \cup B.$ Если выполнено это последнее условие, то $BA \cup AB \subseteq A^2 \cup B^2 \cup BA \cup AB \subseteq A \cup B.$ Если же выполнено (7.6), то, учитывая $A^2 \subseteq A,\ B^2 \subseteq B,$ получаем

$$A^2 \cup B^2 \cup BA \cup AB \subseteq A \cup B \cup BA \cup AB \subseteq$$
$$\subseteq A \cup B \cup A \cup B = A \cup B.$$

Теорема доказана.

Для нестрогих порядков это условие выглядит несколько иначе:

Теорема 7.11. *Для того чтобы объединение* $A \cup B$ *нестрогих порядков А и В было нестрогим порядком, необходимо и достаточно выполнение условий*

$$\begin{cases} BA \cup AB \subseteq A \cup B, \\ A \cap B^{-1} \subseteq E. \end{cases} \tag{7.7}$$

Доказательство. Пусть сначала выполнены условия (7.7). Рефлексивность объединения $A \cup B$ вытекает из рефлексивности операндов. Имеем, далее, $(A \cup B)^{-1} = A^{-1} \cup B^{-1}$. Отсюда

$$(A \cup B) \cap (A \cup B)^{-1} = (A \cup B) \cap (A^{-1} \cup B^{-1}) =$$
$$= (A \cap A^{-1}) \cup (B \cap B^{-1}) \cup (B \cap A^{-1}) \cup (A \cap B^{-1}) =$$
$$= E \cup E \cup (A \cap B^{-1})^{-1} \cup (A \cap B^{-1}) = E.$$

Значит, объединение $A \cup B$ антисимметрично. Далее,

$$(A \cup B)(A \cup B) = A^2 \cup AB \cup BA \cup B^2 \subseteq A \cup B \quad (7.8)$$

и $A \cup B$ тем самым транзитивно. Пусть, наоборот, $A \cup B$ — нестрогий порядок. Тогда, в силу транзитивности, имеем условие (7.8). Из него следует, что $BA \cup AB \subseteq A \cup B.$ Условие антисимметричности $(A \cup B) \cap (A \cup B)^{-1} \subseteq E$ мы можем записать в виде

$$(A \cap A^{-1}) \cup (B \cap B^{-1}) \cup (B \cap A^{-1}) \cup (A \cap B^{-1}) \subseteq E.$$

Отсюда уже вытекает, что $A \cap B^{-1} \subseteq E.$ Теорема доказана.

Замечание. При помощи известнной леммы легко проверяется, что если *A* и *B* — рефлексивные отношения, то условие (7.6) равносильно условию

$$AB = BA = A \cup B.$$

Произведение порядков *AB* также не обязано быть порядком. Это видно из того хотя бы, что для совершенного нестрогого порядка *A* произведение $AA^{-1} \supseteq A \cup A^{-1}$ есть полное отношение. Было бы любопытно отыскать простое необходимое и достаточное условие, при выполнении которого произведение *AB* порядков *A* и *B* было бы порядком. Достаточным условием является, например, такое: если *A* и *B* — строгие порядки и выполнены соотношения

$$\begin{cases} AB = BA, \\ A \cap B^{-1} = \varnothing, \end{cases}$$

то *AB* — строгий порядок (отсюда видно, что надо быть осторожным в попытках построить иерархическую классификацию путем комбинирования разных отношений порядка: род — вид, часть — целое и т. п.).

Доказательство этого утверждения предоставляем читателю.

По поводу транзитивного замыкания $\hat{A}$ заметим, что оно всегда совпадает с исходным порядком *A* в силу его транзитивности.

В заключение данного параграфа мы рассмотрим еще одну операцию, которая для порядков, в некотором смысле, обратна к транзитивному замыканию. Идея явного определения этой операции и ее последовательного применения принадлежит С. Я. Фитиалову.

Определение 7.7. *Редукцией* отношения *A* называется отношение *A"*, определяемое условием:

$$A^r = A \setminus A^2. \quad (7.9)$$

Это означает, что xA^ry выполняется в тех и только тех случаях, когда выполнено само отношение *xAy,* но не существует «промежуточного» *z* такого, что *xAz* и *zAy*. Отношение xA^ry означает «непосредственное подчинение» элемента *x* элементу *y.* Отметим, что

$$A^r \subseteq A. \tag{7.10}$$

Легко проверить также, что для любого отношения *A*

$$(\hat{A})^r \subseteq A. \tag{7.11}$$

На рис. 7.1 мы фактически изображали графы для отношения A^r, а не *A*. Дело в том, что отношение A^r (для случая отношений порядка на конечных множествах) содержит всю нужную информацию об отношении A (см. теорему 7.12), но изображается существенно более простым графом. Сравните, например, граф отношения *A* и граф его редукции A^r на рис. 7.3.

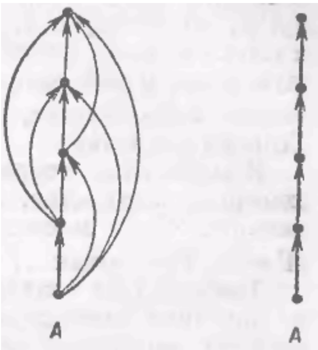


Рис. 7.3.

Обычно вместо графа отношения порядка *A* принято изображать граф отношения A^r, хотя это далеко не всегда оговаривается. Основанием для этого как раз и служит теорема 7.12 (см. ниже). Чтобы перейти в этом случае от отношения A^r к *A,* надо выделить все пути на графе отношения A^r и замкнуть их стрелками.

Сам факт, что отношение *A* восстанавливается по его редукции, не столь тривиален. Так, из (7.9) видно, что, для рефлексивного отношения *A,* $A_r = 0$ и, стало быть, редукция A^r не позволяет восстановить исходное отношение *A*.

Теорема 7.12. *Если A — строгий порядок на конечном множестве M, то транзитивное замыкание редукции совпадает с исходным порядком:*

$$\widehat{A^r} = A\ . \tag{7.12}$$

Доказательство. Из (7.10) и известной теоремы $\widehat{A^r} \subseteq \hat{A} = A.$ Докажем обратное включение. Пусть *xAy.* Отметим, что если

$$xAz_1, \quad z_1Az_2,\ \ldots,\ z_{k-1}Az_k, \quad z_kAy, \tag{7.13}$$

то ввиду транзитивности и антирефлексивности отношения *A* в цепочке элементов $x,\ z_1,\ z_2,\ \ldots,\ z_k,\ y$ любые два элемента различны. Рассмотрим всевозможные цепочки элементов

$z_1, z_2, \ldots, z_k \ (k \geqslant 0)$ такие, что выполняется (7.13). Поскольку *M* — конечное множество и ввиду сделанного только что замечания, таких цепочек конечное число. Значит, среди них существует цепочка максимальной длины. Возьмем ее. (Если цепочек максимальной длины несколько, возьмем любую из них.) Из (7.13) и того, что цепочка $z_1, z_2, \ldots, z_k$ имеет максимальную длину, вытекает

$$xA^rz_1, \quad z_1A^rz_2, \ldots, z_{k-1}A^rz_k, \quad z_kA^ry. \tag{7.14}$$

В самом деле. Если, к примеру, не выполняется $z_1A^rz_2$, то $z_1A^2z_2$, т. е. существует такое *u,* что z_1Au и uAz_2. Но тогда цепочка $z_1, u, z_2, \ldots, z_k$ имеет большую длину и обладает свойством (7.13). Из (7.14) вытекает $x\widehat{A^r}y$. Значит, $A \subseteq \widehat{A^r}$. Мы получили (7.12). Теорема доказана.

К сожалению, теорема 7.12 не переносится на бесконечные множества. Например, если *A* — обычный порядок < на множестве действительных чисел, то $A^r = \varnothing$. Тем самым $\widehat{A^r} = \varnothing$ и $\widehat{A^r} \neq A$.

Теорема 7.12 означает, что для строгих порядков на конечных множествах по отношению A^r можно однозначно восстановить исходное отношение *A*. Более того, редукция A^r есть минимальное отношение, позволяющее восстановить *A*. Точный смысл этого утверждения раскрывает

Теорема 7.13. *Если отношение B таково, что* $\widehat{B} = A$, то $A^r \subseteq B$.

Доказательство. Предположим, что выполнено xA^ry. Из (7.10) xAy; по условию теоремы тогда существует такое *n,* что xB^ny. Однако, в силу $B \subseteq \widehat{B}$, справедливы включения $B \subseteq A$ и $B^n \subseteq A^n$. Значит, верно соотношение xA^ny. Поскольку xA^ry, $n = 1$. Значит, xBy. Теорема доказана.

Из теоремы 7.12 вытекает, что если *A* — строгий порядок на конечном множестве и выполнено xAy, то существует минимальное число *n,* при котором $x(A^r)^ny$. Это *n* характеризует длину минимального пути в графе отношения A^r, который надо пройти, чтобы из вершины *x* попасть в *y*.

Установим некоторые свойства редукций строгих порядков.

Определение 7.8. Отношение *B* называется *антитранзитивным,* если при всех $n > 2$

$$B \cap B^n = \varnothing. \tag{7.15}$$

Иначе говоря, если выполнена цепочка соотношений $xBx_1, x_1Bx_2, \ldots, x_nBy$, то невозможно xBy. В сущности, это значит, что непосредственная связь в графе отношения *B* между вершинами *x* и *y* исключает обходный путь (отметим, что любое антитранзитивное отношение асимметрично н, следовательно, антирефлексивно).

Теорема 7.14. *Если A — строгий порядок, то отношение A^r антитранзитивно.*

Доказательство. Предположим, что существует цепочка $x_1, x_2, \dots, x_n$ такая, что

$$xA^r x_1, \quad x_1 A^r x_2, \ \dots, \ x_n A^r y.$$

Но тогда

$$xAx_1, \quad x_1 A x_2, \ \dots, \ x_n A y.$$

Ввиду транзитивности отношения A x_1Ay. Из xAx_1 и x_1Ay вытекает xA^2y и, следовательно, xA^ry не верно. Теорема доказана.

Полезно рассмотреть отношение, изображенное на рис. 7.4: это циклический граф, его транзитивным замыканием служит полный граф, поскольку, двигаясь по циклу, можно из любой точки попасть в любую, в том числе — в самое себя.

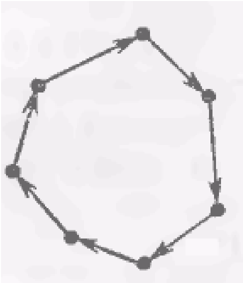


Рис. 7.4

Это отношение не является антнтранзитивным, поскольку $A^{n+1} = A$, где n — число вершин.

Лемма 7.6. *Если отношение B антитранзитивно, то*

$$(\widehat{B})^r = B. \qquad (7.16)$$

Доказательство. Ввиду (7.11) достаточно доказать включение $B \subseteq (\widehat{B})^r$. Пусть для некоторой пары x, y выполнено соотношение xBy и не выполнено $x(\widehat{B})^r y$. Так как $B \subseteq \widehat{B}$, то $x\widehat{B}y$ тоже выполнено. Стало быть, $x(\widehat{B})^2 y$. Но тогда существует $n \geq 2$, для которого $xB^n y$, а это по (7.15) не совместимо с xBy. Полученное противоречие доказывает (7.16).

Равенство (7.16) естественно сопоставить с (7.12).

Если в графе отношения B имеется контур

$$x_1, \ x_2, \ \dots, \ x_n, \ x_1,$$

то $(\widehat{B})^r \neq B$, так как $x_1 B x_2$, но не выполнено $x_1 (\widehat{B})^r x_2$.

(Поскольку x_1Bx_2, значит, $x_1\widehat{B}x_2$. Поскольку $x_1, x_2, ..., x_n, x_1$ — контур, $x_2\widehat{B}x_2$. Из $x_1\widehat{B}x_2$ и $x_2\widehat{B}x_2$ вытекает $x_1(\widehat{B})^2x_2$. Из $x_1\widehat{B}x_2$ и $x_1(\widehat{B})^2x_2$ следует, что не верно $x_1(\widehat{B})^r x_2$.) Однако из отсутствия в графе отношения *В* контуров не вытекает $(\widehat{B})^r = B$. Например, если *В* — обычный строгий порядок на множестве действительных чисел, то в графе отношения *В* нет контуров, но $\widehat{B} = B,\ (\widehat{B})^r = B^r = \varnothing \neq B$.

Легко видеть, что, каково бы ни было отношение *В,* транзитивное замыкание $\widehat{B}$ тогда и только тогда не является антирефлексивным, когда в графе отношения *В* есть контуры. Отсюда вытекает

Лемма 7.7. *Каково бы ни было отношение В, транзитивное замыкание В тогда и только тогда является строгим порядком, когда в графе отношения В нет контуров.*

Теперь может быть легко получена обратная к теореме 7.14

Теорема 7.15. *Если отношение В антитранзитивно, то В есть редукция некоторого строгого порядка.*

Доказательство. Согласно лемме 7.6 $B = (\widehat{B})^r$. По лемме 7.7 достаточно убедиться, что в графе отношения *В* нет контуров. Предположим, что в этом графе есть контур

$x_1,\ x_2,\ \ldots,\ x_n,\ x_1.$

Тогда имеем $x^1B^{n+1}x_2$, т. е. $B \cap B^{n+1} \neq \varnothing$, что противоречит антитранзитивности отношения *В.* Теорема доказана.

7.3. Древесные порядки

В этом параграфе мы будем изучать важный специальный класс отношений порядка — так называемые *древесные порядки.*

Пусть имеется множество *М* с отношением строгого порядка <. Элемент x_0 мы будем называть *наибольшим,* если для всякого элемента $y \in M$, отличного от x_0, выполнено соотношение $y<x_0$. Легко видеть, что наибольший элемент (если он существует) единствен. Полезно отметить также, что для любого строгого порядка, в котором существует наибольший элемент, этот элемент является единственным максимальным (если строгий порядок на конечном множестве имеет единственный максимальный элемент, то этот элемент является наибольшим.)

Определение 7.9. Отношение строгого порядка < на множестве *М* называется *отношением древесного порядка* (или *древесным порядком),* если

1) из того, что $x < y$ и $x < z$ следует, что y и z сравнимы;
2) во множестве $\langle M, <\rangle$ существует наибольший элемент.

Множество M с заданным на нем древесным порядком, т. е. пару $\langle M, <\rangle$, мы будем называть *деревом,* а наибольший элемент — *корнем* дерева.

Условие 1) означает, что для любого элемента $x \in M$ на множестве элементов, больших чем x, исходный древесный порядок превращается в совершенный порядок.

Нетрудно видеть, что совершенный порядок, в котором существует наибольший элемент, есть частный случай древесного.

Установим несколько свойств древесного порядка.

Л е м м а 7.8. *Если A — древесный порядок на M, то на множестве $M(x)$,* состоящем из самого x и всех элементов $y \in M$ таких, что *yAx, отношение A также задает древесный порядок.* (Множество $M(x)$ с порядком A естественно назвать *поддеревом* дерева $\langle M, A\rangle$.)

Доказательство. Первое условие выполняется, очевидно, для любого подмножества множества M. Очевидно также, что наибольшим элементом в $M(x)$ является сам x.

Л е м м а 7.9. *Если A — древесный порядок на конечном множестве M, то для всякого x, отличного от корня x_0, существует ровно один y, для которого выполнено xA^ry.*

Доказательство. Предположим сначала, что существуют такие y и z *($y \neq z$),* что xA^ry и xA^rz. По определению древесного порядка, поскольку xAy, xAz и $y \neq z$, имеем yAz или zAy. Для определенности положим yAz. Тем самым получается, что выполнены два соотношения xAy и yAz. Следовательно, невозможно xA^rz. Итак, мы доказали, что не может быть двух разных элементов, «непосредственно старших», чем данный. Предположим теперь, что для элемента x не существует такого y, что xA^ry. Тогда, легко видеть, не существует такого y, что $x\bar{A}^ry$. Поскольку M — конечное множество, не существует такого y, что xAy (теорема 7.12). Значит, x — максимальный элемент, т. е. $x = x_0$. Лемма доказана.

Если M — множество неположительных действительных чисел с отношением <, то этот древесный порядок не удовлетворяет заключению леммы 7.9.

Лемма 7.10. *Пусть < — древесный порядок на конечном множестве M. Тогда для любых несравнимых элементов $x \in M$ и $y \in M$ существует единственный элемент $z \in M$, для которого* 1) $x < z$; 2) $y < z$; 3) *если* $x < w$ *и* $y < w$, *то* $z \le w$.

Доказательство. Поскольку x и y несравнимы, ни один из них не является корнем дерева. Обозначим через M_x множество всех элементов z, для которых $x < z$, а через M_y— аналогичное множество для y. В силу условия 1) определения 7.9 отношение < на M_x (и на M_y) является совершенным строгим порядком. Так как M_x и M_y содержат корень, то $M_x \cap M_y \neq \varnothing$ и отношение < на $M_x \cap M_y$ является совершенным строгим порядком. Ясно, что множество $M_x \cap M_y$ состоит из всех элементов w, для которых одновременно $x < w$ и $y < w$. Поскольку это множество конечно, то в нем есть наименьший элемент z (лемма 7.1); для любого $w \in M_x \cap M_y$ имеем $z \leqslant w$.

Следующий пример показывает, что в этой лемме условие конечности существенно. Пусть M является объединением полупрямой $(-\infty, 0]$ и двух элементов x и y. Порядок на полупрямой — обычное числовое отношение <, а любая точка на полупрямой больше x и y. Сами элементы x и y не сравнимы. Для этих двух элементов утверждение леммы не верно, хотя определенный нами порядок — древесный.

С помощью доказанных лемм можно убедиться, что граф, изображающий редукцию A^r древесного порядка A на конечном множестве M, действительно имеет древовидную структуру. Назовем *окрестностью* элемента y совокупность элементов z, для которых выполнено zA^ry. Будем изображать A^r по ярусам (рис. 7.5).

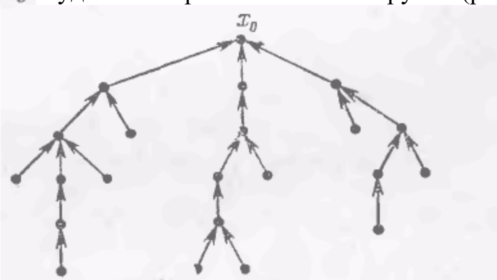


Рис. 7.5. Древесный порядок.

В первом ярусе поместим корень дерева — наибольший элемент x_0. Во втором ярусе поместим элементы, входящие в окрестность элемента x_0. В третьем ярусе поместим элементы, входящие в окрестности элементов второго яруса, и т. д. Ясно, что стрелки в графе могут идти только от яруса к ярусу. При этом от каждого элемента к верхнему ярусу идет ровно одно ребро, а к нижнему может идти сколько угодно ребер. Итак, мы видим, что граф имеет структуру дерева. Общее число ярусов называется *высотой* дерева. Максимальное число элементов в одной окрестности (максимальное число ростков, выходящих из одной вершины) называется *шириной* дерева.

Высота дерева *h*, ширина *d* и общее число вершин *n* связаны очевидным неравенством

$$n \leqslant 1 + d + d^2 + \ldots + d^{h-1} = \frac{d^h - 1}{d - 1}.$$

Это неравенство обращается в равенство в том и только том случае, когда окрестность каждого элемента (кроме, конечно, элементов самого нижнего яруса) состоит из *d* элементов.

М. В. Арапов предложил следующим образом характеризовать *сложность конечного дерева.* Обозначим через *d(x)* число элементов в окрестности элемента *x.* Определим *сложность* σ*(x) вершины x* следующим рекуррентным правилом:

$$\sigma(x) = d(x) + \sigma(y), \tag{7.17}$$

где *у* — тот единственный элемент, для которого $xA^{\tau}y.$ Иначе говоря, сложность вершины *х* складывается из количества ростков, выходящих вниз из этой вершины, и сложности вершины предыдущего яруса, соединенной с *х.* При $x = x_0$ принимаем $\sigma(y) = 0.$ *Сложность дерева* σ*(D)* определяется как суммарная сложность его вершин:

$$\sigma(D) = \sum_{x \in M} \sigma(x). \tag{7.18}$$

Из равенства (7.17) легко вывести, что

$$\sigma(x) = d(x) + \sum_{x < y} d(y).$$

(Знак *х < у* под знаком суммы показывает, что сумма берется по всем таким *у,* для которых *х < у.)* Подставляя это выражение для σ*(x)* в (7.18), получаем

$$\sigma(D) = \sum_{y \in M} d(y)\, k(y), \tag{7.19}$$

где через *k(y)* обозначено, сколько раз величина *d(y)* участвует в выражении для σ*(D).* Ясно, что *k(y)* есть число тех *x,* для которых $x \leqslant y.$ Иначе говоря, *k(y)* равно числу вершин в поддереве, для которого *у* является корнем. Для деревьев D_1, D_2 и D_3, изображенных на рис. 7.8, сложности равны соответственно:

$$\sigma(D_1) = 2 \cdot 7 + 2 \cdot 2 \cdot 3 = 26,$$
$$\sigma(D_2) = 3 \cdot 7 + 3 \cdot 4 = 33,$$
$$\sigma(D_3) = 2 \cdot 7 + 2 \cdot 5 + 2 \cdot 3 = 30.$$

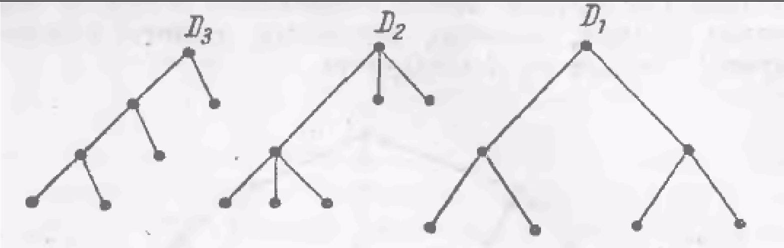


Рис. 7.6. Деревья различной сложности.

Мы здесь специально выбирали деревья с одинаковым количеством вершин, чтобы была заметной зависимость сложности от структуры дерева.

Обозначим через σ_n минимальную сложность дерева с n вершинами. Для вычисления σ_n можно получить рекуррентную формулу. Пусть D_n — дерево минимальной сложности с n вершинами. Пусть $m=d(x_0$ — число отростков, выходящих из его корня. Пусть, наконец, $D^1, D^2, \ldots, D^m$ — поддеревья дерева D_n, начинающиеся со второго яруса. Тогда на основании (7.9)

$$\sigma(D_n) = d(x_0)\cdot n + \sum_{x\in D^1} d(x)\,k(x) + \ldots + \sum_{x\in D^m} d(x)\,k(x).$$

Или, что то же самое,

$$\sigma(D_n) = mn + \sigma(D^1) + \ \ldots\ + \sigma(D^m). \qquad (7.20)$$

Но для дерева минимальной сложности поддеревья также должны иметь минимальную сложность. Иначе мы могли бы уменьшить сумму в (7.20). Обозначим через k_i число вершин в D^i. Сумма чисел k_i равна числу всех вершин в D_n, за исключением корня. Итак,

$$\sigma_n = mn + \sum_{i=1}^{m} \sigma_{k_i},$$

где $k_1 + k_2 + \ \ldots\ + k_m = n - 1$. В силу минимальности величины $\sigma(D_n)$ строение поддеревьев должно быть таково, чтобы эта сумма обратилась в минимум. Значит, окончательно, для отыскания σ_n можно написать рекуррентное уравнение

$$\sigma_n = \min\left(mn + \sum_{i=1}^{m} \sigma_{k_i}\right).$$

В этом уравнении минимум берется по всевозможным m и наборам $\langle k_1, k_2, \ldots, k_m\rangle$, для которых $k_1 + k_2 + \ \ldots\ + k_m = n - 1$. Заметим, что идея получения этого уравнения фактически взята из динамического программирования.

Е. Н. Ефимова сумела получить для величины a_n асимптотику вида

$$\sigma_n \sim n \ln n.$$

* * *

Древесные порядки можно рассматривать не только для конечных множеств. Только леммы 7.9, 7.10 и возможность пользоваться редукцией (теорема 7.12) зависели от конечности множества *М.*

Хороший пример бесконечного дерева можно получить следующим образом. Пусть *М* — множество всех кортежей $x = \langle \varepsilon_0, \varepsilon_1, \ldots, \varepsilon_n \rangle$, где $\varepsilon_0 = 0$, а $\varepsilon_1, \varepsilon_2, \ldots$ принимают значения 0 или 1. Порядок *А* зададим следующим условием. Пусть $x = \langle \varepsilon_0, \varepsilon_1, \ldots \ldots, \varepsilon_n \rangle$ и $y = \langle \eta_0, \eta_1, \ldots, \eta_m \rangle$. Мы будем считать выполненным соотношение *xAy,* если *m* < *n* и, для всех $i \leqslant m$, $\varepsilon_i = \eta_i$. Таким образом, *xAy* означает, что кортеж *у* «вложен» в кортеж *х.* Нетрудно видеть, что если *у* и *z* оба «вложены» в одни и тот же кортеж *х,* то кто-то из них «вложен» в другой. Кортеж x_0 = (0) является, очевидно, наибольшим: поскольку 0 начинает любой кортеж из *М,* то для всех $x \neq x_0$ имеет место xAx_0. В данном случае утверждение леммы 7.9 остается справедливым. Этот порядок можно представлять себе как дерево бесконечной высоты, в котором каждая вершина пускает два ростка.

В предыдущем примере полностью сохранилось понятие яруса. Именно, *n*-й ярус состоял из всех кортежей из *М* длины *n.*

Следующий пример является по существу обобщением предыдущего на «кортежи континуальной длины».

Рассмотрим множество *М,* состоящее из функций *f,* определенных при $0 \leqslant t < +\infty$, принимающих значения 0 и 1 и таких, что *f* (0) = 0. Определим на *М* отношение строгого порядка < условием: *f* < *g,* если *f* и *g* не совпадают и существует такое $a \geqslant 0$, что $g(t) = f(t)$ при $0 \leqslant t \leqslant a$ и $g(t) = 0$ при *t* > *a*.

Проверим, что этот порядок является древесным. В самом деле, пусть *f*<*g* и *f* < g_1. Тогда существуют *a* и a_1 такие, что при $t > a$ $g(t) = 0$, при $t > a_1$ $g_1(t) = 0$, при $0 \leqslant t \leqslant a$ $f(t) = g(t)$ и при $0 \leqslant t \leqslant a_1$ $f(t) = g_1(t)$. Примем, что $a_1 \geqslant a$. Тогда из вышенаписанного вытекает, что, при $0 \leqslant t \leqslant a$, $g(t) = g_1(t)$ и, при $t > a$, $g(t) = 0$. Это означает, что либо *g* и g_1 совпадают, либо g_1 < *g*. Обозначим через f_0 функцию, тождественно равную нулю. Ясно, что, какова бы ни была функция $f \in M$, отличная от f_0, $f < f_0$. Таким образом, f_0 — наибольший элемент (корень).

В этом примере мы имеем ситуацию, которую можно было бы интерпретировать как непрерывное ветвление. Для любого t_0 > 0

совокупность всех функций f таких, что $f(t_0) = 1$ и $f(t) = 0$ при $t > t_0$, как бы образует ярус ранга f_0.
Отметим одно важное обстоятельство. Если в определении древесного порядка отказаться от существования наибольшего элемента, то для конечного множества M мы получим вместо дерева объединение некоторого множества попарно непересекающихся деревьев. Поэтому в случае конечного M вместо условия 2) в определении 7.9 можно взять любое условие, гарантирующее связность соответствующего графа. Например, возможны следующие условия:
2') максимальный элемент единствен;
2") для любых несравнимых элементов x и y существует такой элемент z, что $x < z$ и $y < z$.
В случае бесконечного множества M ситуация оказывается другой. Так, для множества M действительных чисел с обычным порядком выполнено условие 1) определения 7.9 и не выполнено условие 2). Максимальных элементов в этом упорядоченном множестве нет, т. е. условие 2') выполнено. Условие 2") здесь также выполнено.
Приведем в заключение пример «почти-древесного» порядка. Множество M состоит из всех пар вида $\langle m, n\rangle$, где m и n — целые числа и $m \geqslant 0$ (M образует целочисленную решетку на полуплоскости). Определим на множестве M отношение A. Фиксируем произвольное целое число n. Положим по определению:

$$\begin{array}{ll} \langle 0, n+1\rangle A \langle 0, n\rangle, & \langle 1, n+1\rangle A \langle 0, n\rangle, \\ \langle 2, n+1\rangle A \langle 1, n\rangle, & \langle 3, n+1\rangle A \langle 1, n\rangle, \\ \langle 4, n+1\rangle A \langle 2, n\rangle, & \langle 5, n+1\rangle A \langle 2, n\rangle, \\ \langle 6, n+1\rangle A \langle 3, n\rangle, & \langle 7, n+1\rangle A \langle 3, n\rangle \end{array}$$

и т. д. В общем виде:

$$\langle 2k, n+1\rangle A \langle k, n\rangle, \quad \langle 2k+1, n+1\rangle A \langle k, n\rangle.$$

(Советуем читателю попробовать нарисовать граф отношения A.) Легко видеть, что граф отношения A не имеет контуров. Поэтому отношение B — A есть строгий порядок (лемма 7.7).
Этот строгий порядок удовлетворяет условию 1) из определения 4.9 и не удовлетворяет условию 2) того же определения. Однако для этого порядка выполняются условия 2') и 2"). Полезно заметить, что для него выполнены заключения лемм 7.9 и 7.10. (В предыдущем примере заключение леммы 7.9 не выполнено.) В отличие от предыдущего примера здесь мы имеем $\widehat{B^r} = B$, поскольку $B^r = A$ (см. лемму 7.6).
Для любой пары $\langle m, n\rangle$ множество всех пар $\langle l, p\rangle$, для которых $\langle l, p\rangle B \langle m, n\rangle$, образует дерево.

7.4. Множества с несколькими порядками

Мы будем в этом параграфе рассматривать только *конечные* множества (причем конечность будет подразумеваться, а не указываться явно) с несколькими отношениями порядка, связанными определенными условиями «согласования». Содержательные примеры подобных ситуаций, играющих важную роль в математической лингвистике, будут рассмотрены позже. Поэтому ниже мы будем вести изложение на формальном уровне.

Пусть дано множество М, отношение совершенного строгого порядка < на нем и отношение строгого порядка $\Rightarrow$. Редукцию последнего отношения мы будем обозначать через →. Соотношение $x \Rightarrow y$ мы советуем читателю воспринимать как *«x больше y».* Поэтому, в графе отношения $\Rightarrow$ стрелка редукции → будет вести от *большего* элемента к *меньшему.* Множество *M* с двумя такими отношениями $\langle M, <, \Rightarrow \rangle$ мы будем называть *дважды упорядоченным множеством.*

Если выполнено либо $x < z < y$, либо $y < z < x$, то мы будем говорить, что *z* находится *между x* и *y.* Мы будем говорить, что дважды упорядоченное множество *M* удовлетворяет *условию* Π_1, если из $x \rightarrow y$ и того, что *z* находится между *x* и *y,* следует соотношение $x \Rightarrow z$.

Для элемента $x \in M$ мы обозначим через *M(x)* множество, состоящее из самого *x* и всех элементов *y,* для которых $x \Rightarrow y$.

Лемма 7.11. *Пусть* $\langle M, < \rangle$*—отношение совершенного строгого порядка. Если* $x_1, x_2, \ldots, x_n$*—различные элементы из M и w находится между* x_1 *и* x_n*, то либо w совпадает с некоторыми* x_i $(2 \leqslant i \leqslant n-1)$, *либо существует такое* $i (1 \leqslant i \leqslant n-1)$, *что w находится между* x_i *и* x_{i+1}.

Доказательство. Пусть для определенности $x_1 < w < x_n$. Допустим, что *w* отлично от $x_2, x_3, \ldots \ldots, x_{n-1}$. Тогда либо $x_2 > w$, либо $x_2 < w$. Если $x_2 > w$, то $x_1 < w < x_2$. Если $x_2 < w$, рассмотрим x_3. Либо $w < x_3$, либо $x_3 < w$. Если $w < x_3$, $x_2 < w < x_3$. Если $x_3 < w$. рассмотрим x_4 и т. д. Поскольку $w < x_n$ и $\{x_2, x_3, \ldots, x_{n-1}\}$ — конечное множество, через конечное число шагов мы найдем искомое x_i.

Теорема 7.16. *Дважды упорядоченное множество M тогда и только тогда удовлетворяет условию* Π_1 *когда из* $y \in M(x)$, $z \in M(x)$ *и* $y < w < z$ *вытекает*

$$w \in M(x).$$

Доказательство. Пусть сначала *M* удовлетворяет условию Π_1. Если *w* совпадает с *x,* то $w \in M(x)$. Рассмотрим случай *w* < *x.* Поскольку *y* < *w*, *x* ≠ *y.* Вследствие $x \Rightarrow y$ существует последовательность $x = x_1, x_2, \ldots, x_{n-1}, x_n = y$ такая, что при всех *i* $x_i \rightarrow x_{i+1}$. Поскольку $x_n = y < w < x = x_1$, по лемме 7.11 либо, для некоторого i, $w = x_i$, либо существует такое *i*, что *w* находится между x_i и x_{i+1}. В первом случае $x \Rightarrow w$ и $w \in M(x)$. Во втором случае, в силу условия Π_1, $x_i \Rightarrow w$. Поскольку $x \Rightarrow x_i$, имеем $x \Rightarrow w$ и $w \in M(x)$. В случае *x* < *w* рассуждения проходят таким же образом, если заметить, что *w* находится теперь между *x* и *z*. Доказательство в обратную сторону предоставляем читателю.

В силу теоремы 7.2 дважды упорядоченное множество *M* можно изображать отрезком натурального ряда $\{1, 2, \ldots, m\}$, а отношение < понимать как соответствующее числовое неравенство. *Интервалом* мы будем называть любое множество $[i, j]$, состоящее из всех натуральных чисел *l*, удовлетворяющих неравенствам: $i \leqslant l \leqslant j$. Предыдущая теорема означает: условие Π_1 равносильно тому, что все множества *M(x)* суть интервалы.

П р и м е р. Пусть *M* = {1, 2, 3, 4}, а отношение → определено условиями $1 \rightarrow 2$, $1 \rightarrow 3$, $4 \rightarrow 2$, $4 \rightarrow 3$. Тогда

$$M(1) = \{1, 2, 3\},\ M(2) = \{2\},\ M(3) = \{3\}$$

и $M(4) = \{2, 3, 4\}$. Нетрудно проверить, что это дважды упорядоченное множество удовлетворяет условию Π_1 (рис. 7.7).

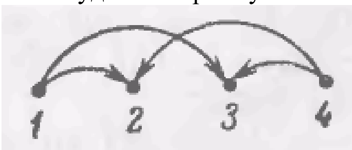


Рис. 7.7.

Отметим, что $M_1(1) \cap M(4) \neq \varnothing$, но ни одно из этих множеств не содержится в другом.

Полезную информацию о взаимном расположении множеств *M(x)* дает

Теорема 7.17. *Если отношение* $\langle M, \Rightarrow \rangle$ *есть древесный порядок, то для любых несовпадающих x и y либо* $M(x) \cap M(y) = \varnothing$, *либо* $M(x) \subset M(y)$, *либо* $M(y) \subset M(x)$.

Доказательство. Пусть $M(x) \cap M(y) \neq \varnothing$ и $w \in M(x) \cap M(y)$. Если $w \neq x$ и $w \neq y$, то имеем $x \Rightarrow w$ и $y \Rightarrow w$. В силу древесностн порядка и несовпадения *x* и *y* имеем либо $x \Rightarrow y$, либо $y \Rightarrow x$. Если же $w = x$, то $y \Rightarrow x$, а при *w* = *y* имеем $x \Rightarrow y$. Если

$x \Rightarrow y$, то по транзитивности имеем $x \Rightarrow z$ для любого $z \in M(y)$, т. е. $M(x) \supset M(y)$. Если же $y \Rightarrow x$, то получаем обратное включение.

Условимся теперь изображать множество *M* на горизонтальной оси, а стрелки, выражающие отношение →, проводить только сверху этой оси. Мы будем говорить, что дважды упорядоченное множество *M* удовлетворяет *условию* Π_1, если можно все стрелки → провести так, чтобы они не пересекались и не накрывали максимальные элементы (то есть элементы *x*, для которых соотношение $y \to x$ не выполнено ни при каком *y*.

Теорема 7.18. *Условие* Π_2 *влечет условие* Π_1.

Доказательство. Пусть дважды упорядоченное множество *M* удовлетворяет условию Π_2. Проведем стрелки, выражающие отношение →, соответствующим «хорошим» образом. Пусть $x \to y$, а *z* находится между *x* и *y*. Элемент *z* не может быть максимальным элементом, поскольку он накрывается стрелкой, ведущей из *x* в *y*. Тогда существует z_1 для которого $z_1 \to z$. Элемент z_1 находится между *x* и *y*, либо совпадает с *x* или *y*, ибо иначе стрелка $z_1 \to z$ пересекалась бы со стрелкой $x \to y$. Аналогичным рассуждением мы построим цепочку $z_n \to z_{n-1} \to \ldots \to \to z_1 \to z$, где либо все z_i лежат между *x* и *y*, либо z_n совпадает с *x* или *y*. Так как все z_i различны и множество *M* конечно, то для какого-то *n* элемент z_n совпадает с *x* или *y*. Но тогда будем иметь $x \Rightarrow z$. Таким образом, дважды упорядоченное множество *M* удовлетворяет условию Π_1.

Пример на рис. 7.7 показывает, что условие Π_1 может выполняться, когда условие Π_2 не выполнено. Однако когда $\Rightarrow$ является древесным порядком, условия Π_1 и Π_2 равносильны. А именно, имеет место

Теорема 7.19. *Если отношение* $\Rightarrow$ *является древесным порядком и дважды упорядоченное множество удовлетворяет условию* Π_1, *то оно удовлетворяет условию* Π_2.

Доказательство. Пусть x_0 — корень дерева, а $x_1 < x_2 < \ldots < x_n$ — все те элементы, для которых выполнено $x_0 \to x_i$. Ясно, что все стрелки, исходящие из x_0, можно провести без пересечений. Множества $M(x_1), M(x_2), \ldots, M(x_n)$ являются по теореме 7.16 интервалами. Включение $M(x_i) \supset M(x_j)$ невозможно, так как влекло бы условие $x_i \Rightarrow x_j$; следовательно, по теореме 7.17 эти интервалы не могут пересекаться. Между разными множествами $M(x_i)$ и $M(x_j)$ не могут проходить стрелки; иначе было бы выполнено $x_i \Rightarrow w$, где $w \in M(x_j)$. Если все элементы из $M(x_i)$, кроме самого x_i лежат между x_{i-1} и x_i, то стрелки внутри $M(x_i)$ можно провести так, чтобы они не пересекались со стрелками, идущими от корня. Аналогично будет и в случае, когда $M(x_i)$ лежит между x_i и x_{i+1}. Пусть теперь

$M(x_i) = M^1(x_i) \cup \{x_i\} \cup M^2(x_i)$, где $M^1(x_i)$ лежит между x_{i-1} и x_i, а $M^2(x_i)$ лежит между x_i и x_{i+1}. Покажем, что нет ни одной стрелки, ведущей из $M^1(x_i)$ в $M^2(x_i)$ (и соответственно в обратном направлении). Действительно, существование такой стрелки означало бы, что $y \to z$, где $y \in M^1(x_i)$ и $z \in M^2(x_i)$. Но так как $y < x_i < z$, то, согласно условию $П_1$, было бы $y \Rightarrow x_i$, что невозможно. Итак, все остальные стрелки, не идущие из x_0, можно провести без пересечений со стрелками, идущими из x_0. Но каждое из $M(x_i)$ является дважды упорядоченным множеством, удовлетворяющим условию $П_1$ и сужение $\Rightarrow$ на $M(x_i)$ является древесным порядком. Поэтому и стрелки внутри $M(x_i)$ можно провести без пересечений со стрелками, выходящими из x_i. Продолжая это рассуждение, мы легко убедимся, что все стрелки можно провести без пересечений. Так как корень x_0 не входит ни в одно из множеств $M(x_i)$ и никакое $M(x_i)$ не расположено по обе стороны x_0, то, очевидно, все стрелки можно провести так, чтобы x_0 не накрывалось. Теорема доказана.

Заметим, что решающую роль в доказательстве играл тот факт, что все $M(x_i)$ либо не пересекаются, либо вложены одно в другое. Поэтому формулировку последней теоремы можно было бы несколько уточнить.

Существует еще одна полезная формулировка связи между обоими отношениями порядка на дважды упорядоченном множестве. Она имеет смысл только для случая, когда отношение $\Rightarrow$ является древесным порядком. (Правда, ее можно расширить на те «недревесные» ситуации, для которых удается ввести понятие яруса.)

Сначала изобразим элементы множества M целочисленными точками от 1 до n на оси абсцисс в координатной плоскости. Затем каждой точке $x \in M$ сопоставим такую точку x' на восстановленном из x перпендикуляре к оси абсцисс, чтобы расстояние от x' до x равнялось уменьшенному на единицу номеру яруса, на котором находится x.

Если $x \to y$, то точки x' и y' соединим отрезком. На рис. 7.8, 7.9 и 7.10 проведены соответствующие построения.

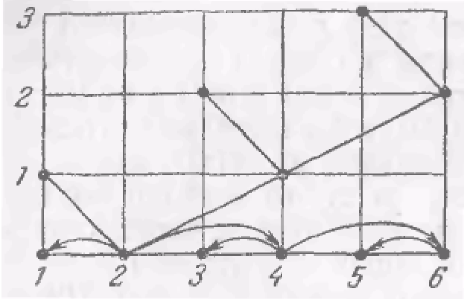


Рис. 7.8.

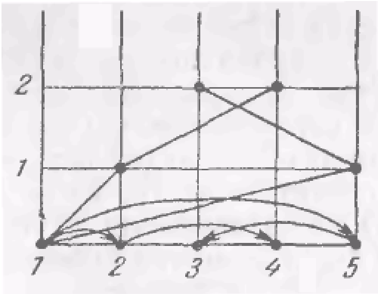


Рис. 7.9.

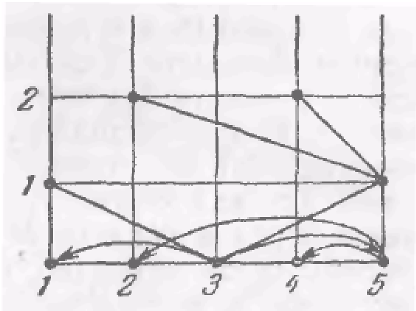


Рис. 7.10.

Условие Π_3 состоит в том, что а) проведенные отрезки не пересекаются и б) продолжение перпендикуляра выше любой точки x' не пересекает ни одного отрезка.

На рис. 7.8 условие Π_3 выполнено, а на рис. 7.9 и 7.10 не выполнено.

Теорема 7.20. *Если отношение* $\Rightarrow$ *является древесным порядком, то условие* Π_3 *равносильно условию* Π_1.

Доказательство. Покажем сначала, что из $П_3$ следует $П_1$. Предположим, что существуют *x, y* и *z,* для которых $x \to y$, *z* находится между *x* и *y* и не выполнено $x \Rightarrow z$. Пусть z_0 — максимальный элемент среди всех таких *z* (при фиксированных *x* и *y).* Из условия $П_3$ следует, что точка z'_0 не может лежать под отрезком, соединяющим *x', y'*. Отсюда, в частности, вытекает, что z_0 не может быть корнем дерева. Но в силу древесности порядка тогда существует *w*, для которого $w \to z_0$. По принятому предположению *w* не находится между *x* и *y* и не совпадает с *x* или *y.* Возможны следующие варианты расположения:

$$x < z_0 < y < w;$$
$$w < x < z_0 < y;$$
$$y < z_0 < x < w;$$
$$w < y < z_0 < z.$$

Рассмотрим первый случай. Так как z'_0 лежит выше отрезка *x'y',* то либо отрезок z'_0w' пересекает отрезок *x'y',* либо отрезок z'_0w' лежит выше точки *y'.* В обоих случаях условие $П_3$ нарушается. Остальные три варианта исследуются совершенно аналогично. Итак, мы доказали, что из $П_3$ следует $П_1$.

Покажем теперь, что из $П_1$ следует $П_3$. Предположим, что $П_3$ не выполнено. Рассмотрим сначала случай, когда продолжение перпендикуляра, проходящего через точку *x',* пересекает выше точки *x'* отрезок *y'z'*. Это означает, что точка *x* расположена между *y* и *z,* $y \to z$ (или $z \to y$), и хотя бы одна из точек *y*' или *z'* лежит выше, чем *x'.* Тогда другая лежит не ниже, чем *x'.* Следовательно, невозможно $y \Rightarrow x$, что противоречит условию $П_1$. Второй случай состоит в том, что некоторые отрезки $x'y'$ и $z'u'$ пересекаются. Это возможно только при условии, что нижние и верхние концы этих отрезков лежат соответственно на одном уровне. Но тогда невозможно ни одно из соотношений

$x \Rightarrow z,\ x \Rightarrow u,\ y \Rightarrow z,\ y \Rightarrow u.$ С другой стороны, либо *z,* либо *u* находится между *x* и *y,* т. е. согласно $П_1$ одно из перечисленных соотношений должно было бы выполняться. Итак, теорема доказана.

Следствие. *Если* $\Rightarrow$ *есть древесный порядок, то* $П_3$ *равносильно* $П_2$.

* * *

Перейдем теперь к изучению другого тина множеств с двумя отношениями порядка.

Упорядоченным деревом мы будем называть множество *М* с заданными на нем отношением древесного порядка $\subset$ и отношением строгого порядка < (т. е. тройку $\langle M, \subset, <\rangle)$, если выполнены следующие условия:

1) если $x \subset y$, $z \subset u$, $y < u$, то $x < z$;

2) если *х* и *у* несравнимы по отношению $\subset$, то они сравнимы по отношению <.

В частности, на подмножестве концевых вершин дерева $\langle M, \subset\rangle$ отношение < образует совершенный порядок. Обозначим через *Q(х)* окрестность элемента *х*. Тогда на множестве *Q(х)* отношение < также задает совершенный порядок.

Поскольку *М* конечно, любое множество *Q (х)* можно перенумеровать так, что максимальный (по отношению <) элемент получит номер 0, следующий за ним номер 1 и т. д. Так как каждый элемент *у* входит ровно в одно множество *Q{х)* (благодаря древесности порядка $\subset$),то тем самым любой вершине *у* (кроме корня дерева) оказался приписан некоторый целочисленный вес *m(у).* На рис. 7.11 изображено дерево с весами вершин, на котором порядок < в каждом *Q(х)* выражается размещением вершин слева направо.

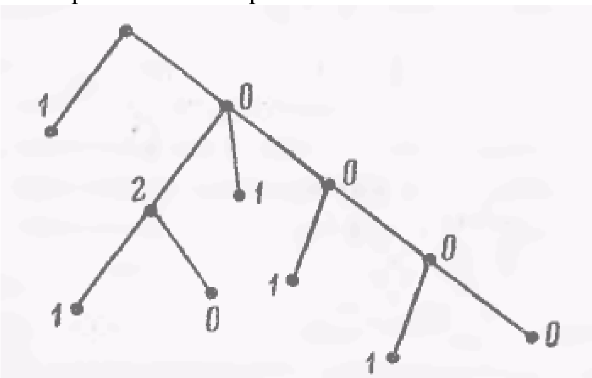


Рис. 7.11.

Для концевой вершины *у* определим величину

$$\gamma(y) = \Sigma m(x),$$

равную сумме весов вершин, лежащих на пути из корня дерева в вершину *у,* включая вес самой вершины *у.* Так, для семи концевых вершин дерева на рис. 7.11 мы получаем, идя слева направо, следующие значения для $\gamma(y)$:

$$1, 3, 2, 1, 1, 1, 0.$$

Глубиной дерева мы будем, следуя В. Ингве (V. H. Yngve), называть величину

$$\gamma = \max_{y} \gamma(y).$$

Так, дерево на рис. 7.11 имеет глубину $\gamma = 3$. Полезно заметить, что глубина имеет смысл только для упорядоченного дерева, а просто для дерева она не определена. Малая величина глубины означает геометрически, что ветвления идут, в основном, вправо, т. е. что дерево устроено асимметрично.

* * *

В заключение параграфа изучим некоторые свойства множеств с тремя отношениями порядка. Пусть на множестве M заданы три отношения строгого порядка: $<$, $\subset$ и $\Rightarrow$, для которых выполняются следующие условия:

1) Множество M с отношениями $\subset$ и $<$ является упорядоченным деревом;

2) отношение $\Rightarrow$ определено на множестве концевых вершин $M_k \subset M$ предыдущего дерева и образует на M_k древесный порядок;

3) обозначим через $S(x) \subset M_k$ множество концевых вершин y, для которых $y \subset x$. Тогда S (x) должно быть деревом по отношению $\Rightarrow$.

4) если $S(x) \cap S(x_1) = \varnothing$, $y \in S(x)$, $z \in S(x_1)$ и $y \to z$, то y — корень дерева $S(x)$. (Ясно, что z является в этом случае корнем дерева $S(x_1)$.)

5) если существует такое u, что $y < u < z$, то существуют x и x_1 такие, что $y \in S(x)$, $z \in S(x_1)$, и не существует w, для которого $x < w < x_1$. (Легко проверить, что в этом случае $S(x) \cap S(x_1) = \varnothing$.)

Теорема 7.21. *При сформулированных выше условиях дважды упорядоченное множество* $\langle M_k, <, \Rightarrow \rangle$ *удовлетворяет условию* Π_1.

Доказательство. Пусть $y \to z$ и u находится между y и z. Покажем, что $y \Rightarrow u$. Рассмотрим случай $y < u < z$. (Противоположный случай исследуется аналогичным способом.) Согласно свойству 5) выберем $S(x)$ и $S(x_1)$, для которых $y \in S(x)$ и $z \in S(x_1)$. В силу $y \to z$ и свойства 4) y и z суть корни деревьев $S(x)$ и $S(x_1)$. Покажем, что u входит либо в $S(x)$, либо в $S(x_1)$.

Действительно, пусть $u \notin S(x)$ и $u \notin S(x_1)$. Тогда элемент u заведомо несравним с x и x_1, поскольку $u \subset x$ и $u \subset x_1$ отрицается предположением, а соотношения $u \supset x$ или $u \supset x_1$ невозможны, поскольку $u \in M_k$. Тогда элемент u должен быть сравнимым с x и x_1 по отношению $<$. В силу свойства 1) упорядоченного дерева и условия

$y < u < z$ должно быть $x < < u < x_1$. Но это противоречит выбору x и x_1 (условию 5). Значит, непременно $u \in S(x)$ или $u \in S(x_1)$. В первом случае $x \Rightarrow u$, а во втором случае $y \Rightarrow u$, т. е. опять-таки $x \Rightarrow u$. Теорема доказана.

7.5 Отношения между геометрическими объектами

Многие хорошо известные из школьной математики понятия в сущности являются названиями бинарных отношений, а основные связанные с ними теоремы выражают свойства этих отношений.

Пусть M — множество всех прямых на плоскости. Соотношение $X \parallel Y$ означает, что прямые X и Y параллельны, то есть не имеют общих точек. Установим некоторые свойства этого отношения.

1. Отношение $\parallel$ антирефлексивно. Действительно, никакая прямая не параллельна сама себе.

2. Отношение $\parallel$ симметрично. Это видно из того, что в определении параллельности обе прямые равноправны.

3. Отношение $\parallel$ почти транзитивно, а именно: если $X \parallel Y$ и $Y \parallel Z$, то либо $X \parallel Z$, либо прямые X и Z совпадают. Действительно, если бы это было не так, то прямые X и Z пересекались бы, то есть имели бы ровно одну общую точку. Но, как известно из геометрии, если прямая Z пересекается с одной из параллельных X, то она пересекается и с другой из параллельных Y, т. е. было бы невозможно соотношение $Z \parallel Y$.

Таким образом, отношение параллельности между прямыми не обладает еще хорошими свойствами. Но сказанное выше позволяет легко сообразить, какое отношение, родственное параллельности, будет отношением эквивалентности. А именно, определим отношение,

$$\parallel\!| = \parallel \cup E,$$

которое выполняется, когда прямые либо параллельны, либо совпадают. По определению $X \parallel\!| X$для любой прямой X. Симметричность отношения $\parallel\!|$ также очевидна. Наконец, если $X \parallel\!| Y$ и $Y \parallel\!| Z$, то $X \parallel\!| Z$. В самом деле, если $X \parallel Y$ и $Y = Z$, то $X \parallel Z$; если $X = Y$ и $Y \parallel Z$, то $X \parallel Z$. Наконец, если $X \parallel Y$ и $Y \parallel Z$, то по сказанному ранее либо $X \parallel Z$, либо $X = Z$. Во всех случаях имеем $X \parallel\!| Z$.

Отношение $\parallel\!|$ на множестве прямых очень естественно выглядит в алгебраической форме. Если на плоскости ввести декартовы координаты x и y, то всякая прямая, не перпендикулярная к оси Ox (не вертикальная), задается уравнением $y=kx + b$. Иначе говоря, любая (за указанным исключением) прямая X определяется парой чисел (k, b). Пусть прямая X задается уравнением $y=kx + b$, а прямая Y — уравнением $y=k'x + b'$. Тогда соотношение $X \parallel\!\mid Y$ выполняется в том и только том случае, когда $k = k'$. Соотношение $X \parallel Y$ означает, что $k = k'$ и одновременно $b \neq b'$, т. е. прямые различны. Это видно из того, что $k = \operatorname{tg} \alpha$, где α — угол наклона прямой к оси Ox. Для вертикальных прямых можно положить $k = \infty \quad (\alpha = 90^\circ)$, и условие $k = k'$ будет по-прежнему означать $X \parallel\!\mid Y$. Однако это соглашение не очень красиво, так как при $k = \infty$ у нас не определен второй параметр, различающий параллельные прямые.

В аналитической геометрии дается более универсальная (так называемая *нормальная)* форма уравнения прямой:

$$x \cos \alpha + y \sin \alpha - p = 0,$$

которая описывает прямую любого вида. Здесь p — длина перпендикуляра, опущенного из начала координат на прямую (рис. 7.12), α — угол наклона этого перпендикуляра к оси абсцисс.

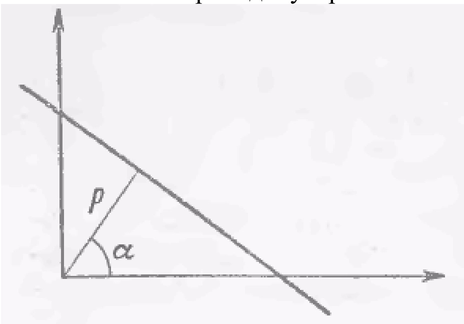


Рис. 7.12.

Тем самым каждой прямой взаимно-однозначно сопоставлена пара чисел $\langle \alpha, p \rangle$, где $0 \leqslant \alpha < 2\pi$ и $0 \leqslant p < +\infty$. Соотношение $X \parallel\!\mid Y$ означает, что для соответствующих прямых α=α′ или $\alpha = \alpha' + \pi$. Каждой прямой соответствует точка на плоскости параметров α и p, лежащая в области, указанной на рис. 7.13.

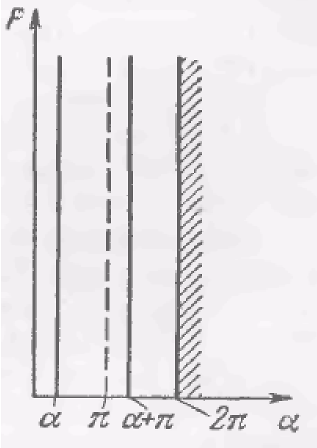


Рис. 7.13.

Пары вертикальных прямых $\alpha = \text{const}$ и $\alpha + \pi = \text{const} \quad (0 \leqslant \alpha < \pi)$ суть классы эквивалентности отношения $|||$.

На множестве прямых на плоскости существует еще одно важное отношение: $X \perp Y$ (*X* перпендикулярна к *Y*). Отношение перпендикулярности обладает следующими важными свойствами:

1. Антирефлексивность. Невозможно $X \perp X.$
2. Симметричность. Если $X \perp Y,$ то $Y \perp X.$
3. Если $X \perp Y$ и $Y \perp Z,$ то невозможно $X \perp Z.$ Из $X \perp Y$ и $Y \perp Z$ следует, очевидно, $X \parallel\mkern-3mu| Z.$ Обратно, если $X \parallel\mkern-3mu| Z,$ то существует общий перпендикуляр *Y* к прямым *X* и *Z,* т. е. такое *Y*, что $X \perp Y$ и $Y \perp Z.$

Оба эти утверждения означают, что квадрат отношения перпендикулярности есть отношение $|||$ «усиленной параллельности»:

$$\perp\perp = \perp^2 = |||\,.$$

Введем на *M* еще одно отношение *X* Пер *Y*, означающее, что прямые *X* и *Y* имеют хотя бы одну общую точку, т. е. пересекаются или совпадают. Ясно, что отношение Пер рефлексивно и симметрично (но не транзитивно) и является, стало быть, отношением толерантности.

Выберем на плоскости некоторую точку *P* и рассмотрим множество K_P всех прямых, проходящих через эту точку. Легко видеть, что K_P есть класс толерантности. Действительно, любые прямые из K_P имеют общую точку, а именно, саму точку *P*. С другой стороны, любая прямая *X,* не входящая в K_P, не пересекается с некоторой прямой из K_P, а именно, с прямой, проходящей через точку P и параллельной прямой X. Предоставляем читателю проверить, что классы K_P образуют базис.

Существуют и другие классы толерантности. Например, множество всех прямых, касающихся некоторой полуокружности с выкинутой концевой точкой, образуют класс толерантности. Действительно, среди этих прямых нет параллельных друг другу. А для любой прямой, не входящей в данное множество, можно построить параллельную к ней прямую, касающуюся данной полуокружности.

Пусть теперь M — множество всех треугольников на плоскости. Читатель легко убедится, что равенство и подобие треугольников суть отношения эквивалентности (заметим, что равенство треугольников в геометрии вовсе не означает их тождества (совпадения).).

Обозначим через M_K множество окружностей на плоскости и определим отношение $X \models Y$ условием, что окружность X находится внутри окружности Y. Ясно, что это отношение антирефлексивно и трапзитивно, т. е. является строгим порядком. Этот порядок не является совершенным, так как существуют пары окружностей, из которых ни одна не лежит внутри другой.

Множеству всех прямых присвоим обозначение $M_П$. Тогда мы можем рассмотреть отношения между прямыми и окружностями. Примером такого отношения является X Кас Y — прямая X касается окружности Y.

Произведение Кас $(Кас)^{-1}$ есть отношение на множестве прямых и ХКас$(Кас)^{-1}$ Y равносильно тому,что существует окружность V такая, что X Кас V и Y Кас V. Итак, X Кас $(Кас)^{-1}Y$ означает, что у прямых X и Y существует общая касательная окружность V. Но такая окружность существует для любых двух прямых. Стало быть, отношение Кас$(Кас)^{-1}$ выполнено для любых двух прямых, а, значит, является полным отношением на $M_П$.

Отношение $(Кас)^{-1}$Кас определено на множестве окружностей M_K и $X(Кас)^{-1}$КасY означает, что существует прямая W, для которой WКасX и WКасY, т. е. что к окружностям X и Y можно провести общую касательную.

7.6. Отношения между уравнениями

Пусть теперь множество M состоит из уравнений вида

$$f(x) = g(x). \qquad (\xi)$$

Рассматриваемое уравнение мы будем обозначать греческой буквой, которую мы для этого поставили в одной строчке с уравнением.

Корнем уравнения называется такое действительное число *a*, что при подстановке в обе части уравнения числа *a* вместо *x* мы получаем равные числа. Множество всех корней уравнения ξ мы будем обозначать R_ξ.

Например, для уравнения

$$x^2 = x^3 \qquad (\xi_1)$$

множество R_{ξ_1} состоит из чисел 0 и 1. Для уравнения

$$\cos x = \sin x \qquad (\xi_2)$$

множество R_{ξ_2} состоит из всех чисел вида

$$x = \frac{\pi}{4} + \pi n \qquad (n = 0,\ \pm 1,\ \pm 2,\ \ldots)$$

и является бесконечным. Для уравнения

$$1 + x^2 = -1 \qquad (\xi_3)$$

множество корней R_{ξ_3} пусто, поскольку при любом действительном значении *x* левая часть положительна, а правая отрицательна. Зато для уравнения

$$(x-1)^2 = x^2 - 2x + 1 \qquad (\xi_4)$$

множество корней R_{ξ_4} есть множество всех действительных чисел.

Введем теперь отношения между уравнениями:

Определение 7.6.1. Уравнения ξ и η называются *равносильными:*

$$\xi \approx \eta$$

если их множества корней совпадают: $R_\xi = R_\eta$.

Из того, что равенство двух множеств есть отношение эквивалентности, легко получается, что отношение равносильности $\approx$ есть отношение эквивалентности. В школьной алгебре изучаются преобразования уравнений, которые переводят уравнение ξ в равносильное ему уравнение η.

Определение 7.6.2. Уравнение ξ *не сильнее* уравнения η: $\xi \Rightarrow \eta$, если $R_\xi \subseteq R_\eta$. В этом случае естественно говорить также, что уравнение η *не слабее* уравнения ξ (в этом случае уравнение η называют также *выводным* из уравнения ξ (или *следствие»* уравнения ξ).)

Легко проверить, что отношение $\Rightarrow$ рефлексивно и транзитивно, т. е. является квазипорядком. Ясно также, что из $\xi \Rightarrow \eta$ и $\eta \Rightarrow \xi$ вытекает равносильность $\xi \approx \eta$. Обратно, из равносильности $\xi \approx \eta$ следует $\xi \Rightarrow \eta$ и $\eta \Rightarrow \xi$. Таким образом, $\approx\ = \ \Rightarrow \cup (\Rightarrow)^{-1}$.

На множестве уравнений, имеющих хотя бы один корень, легко определить естественное отношение толерантности — наличие общих корней: $R_\xi \cap R_\eta \neq \varnothing$.

Можно ввести еще отношение $\approxeq$ эффективной равносильности уравнений. Уравнения ξ и η мы будем называть *эффективно*

равносильными, если каждое из них можно преобразовать в другое с помощью конечного числа применений разрешенных приемов из фиксированного списка (предполагается, конечно, что преобразования, входящие в этот фиксированный список, сохраняют равносильность).
В силу транзитивности отношения $\approx$ любое число применения таких приемов не нарушает равносильности уравнений. Поэтому эффективно равносильные уравнения являются равносильными, что можно записать как включение одного из этих отношений в другое:

$$\approxeq \subseteq \approx .$$

Так, уравнения

$$\frac{x-1}{x^2+1}=2 \qquad (\xi_5)$$

и

$$2x^2-x+3=0 \qquad (\xi_6)$$

эффективно равносильны:

$$\xi_5 \approxeq \xi_6,$$

так как уравнение ξ_6 можно получить из уравнения ξ_5 с помощью цепочки известных из школьной алгебры преобразований, и обратно.
С другой стороны, уравнения

$$x-1=0 \qquad (\xi_7)$$

и

$$2^x-1-x=0 \qquad (\xi_8)$$

равносильны (оба имеют один корень: $x= 1$), но не эффективно равносильны, так как ξ_8 нельзя получить из ξ_7 алгебраическими преобразованиями. Таким образом, справедливо строгое включение:

$$\approxeq \subset \approx .$$

8. Отображения отношений

8.1. Гомоморфизмы и корреспонденции

Нам уже неоднократно приходилось сопоставлять разные множества и отношения на них. Например, произвольное пространство толерантности и множество S_H непустых подмножеств множества H. Или множество с заданным на нем квазипорядком и множество классов эквивалентности и индуцированным порядком. В этом разделе мы введем важные общие понятия, позволяющие говорить о сопоставлении различных множеств с заданными на них отношениями.

Пусть имеются два отношения: $\langle A, M\rangle$ и $\langle B, L\rangle$. Сопоставить эти отношения — это значит соотнести элементам множества L какие-то элементы множества M и указать, какую информацию о выполнении отношения B несет тот факт, что между некоторыми элементами множества M выполнено отношение A. Запись $\alpha\colon \langle A, M\rangle \to \langle B, L\rangle$ будет обозначать далее, что α — отображение множества M в множество L, а $\langle A, M\rangle$ и $\langle B, L\rangle$ — отношения. Советуем читателю вспомнить определения сюръективного, инъективного и биективного отображения.

Определение 8.1. Отображение $\alpha\colon M \to L$ называется *гомоморфным отображением* (или *гомоморфизмом)* отношения $\langle A, M\rangle$ в отношение $\langle B, L\rangle$, если из xAx' следует $\alpha(x)B\alpha(x')$.

Иначе говоря, из того, что выполняется отношение A для прообразов, следует, что выполняется отношение B для образов. На рис. 8.1 показаны два примера гомоморфизмов отношений.

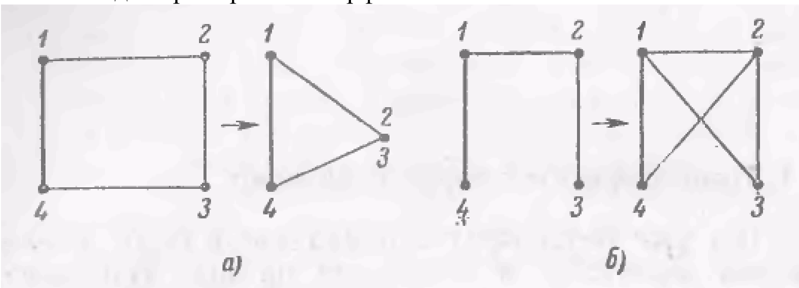


Рис. 8.1. Гомоморфизмы отношений.

Чтобы не проводить лишних стрелок, соответствие вершин указано их нумерацией. В частности, на рис. 8.1, *а* указано, что вершины *2* и *3* переходят при отображении в одну вершину. На рис. 8.1 (а также на рис. 8.2 и 8.3) изображены симметричные отношения и поэтому в графах не проведены стрелки.

Определение 8.2. Отображение $\alpha\colon M \to L$ называется *корреспонденцией* отношения $\langle A, M\rangle$ в отношение $\langle B, L\rangle$, если из $\alpha(x)B\alpha(x')$ следует xAx'.

Иначе говоря, выполнение отношения B для пары образов влечет выполнение отношения A для любой пары их прообразов. (Термин «корреспонденция» для таких отображений впервые использовал С. К. Шаумян.)

На рис. 8.2 приведены два примера корреспонденции между отношениями.

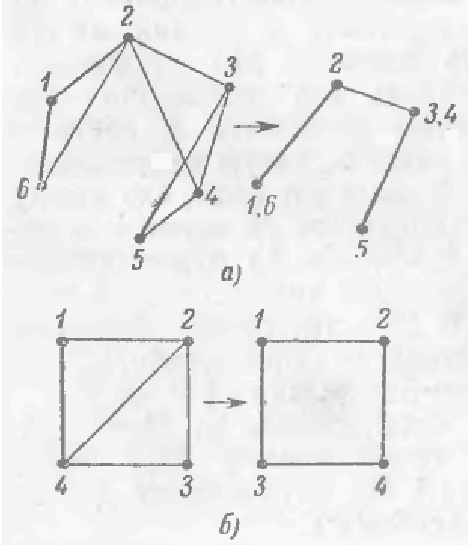


Рис. 8.2. Корреспонденции отношений

Поучительно проанализировать, какие ребра на отображаемом графе необходимы, чтобы отображение было корреспонденцией. На рис. 8.2, *а* не обязательны (их изъятие не нарушит корреспонденцию) только ребра (1,6) и (3,4).

Если отображение $\alpha\colon M \to L$ биективно, то *а* тогда и только тогда является корреспонденцией, когда обратное отображение $\alpha^{-1}\colon L \to M$ является гомоморфизмом. Если обратного отображения не существует, топонятие корреспонденции не сводится к понятию гомоморфизма.

Понятие корреспонденции оказалось полезным в задачах математической лингвистики.

Если отображение *а* сюръективно, то гомоморфизм *а* мы будем называть *эпиморфизмом;* если *а* инъективно. то гомоморфизм *а* называется *мономорфизмом;* если, наконец, *а* биективно, гомоморфизм *а* называется *изоморфизмом.* Таким образом, для любого множества *М* тождественное отображение ε множества *М* на себя является изоморфизмом пустого отношения $\langle \varnothing, M \rangle$ в полное отношение $\langle M^2, M \rangle$, что, конечно, мало согласуется с тем, что обычно в математике связывается со словом «изоморфизм». Более разумным является вводимое в следующей фразе понятие *k*-изоморфнзма. Гомоморфизм (эпиморфизм, мономорфизм, изоморфизм) ее называется *k-го-моморфизмом* (соответственно *k-эпиморфизмом,* *k-мономорфизмом, k-изоморфизмом)*, если *а* одновременно является корреспонденцией.

Хороший пример *k*-гомоморфизма можно извлечь из предыдущей главы. Пусть *M* — множество уравнений, а *L* — множество, состоящее из множеств действительных чисел. Рассмотрим отображение $\varphi\colon M \to L,$ сопоставляющее каждому уравнению $\xi \in M$ множество его корней $R_\xi \in L.$ Ясно, что разным уравнениям может соответствовать одно и то же множество корней. Но, согласно определению 7.6.1, отображение φ является *k*-гомоморфизмом отношения $\langle \approx, M \rangle$ в отношение $\langle =, L \rangle,$ поскольку равносильным уравнениям соответствуют совпадающие множества корней и, обратно, если множества корней для двух уравнений совпадают, то сами уравнения равносильны. Аналогично, согласно определению 7.6.2, то же отображение φр будет также *k* -гомоморфизмом отношения $\langle \Rightarrow, M \rangle$ в отношение $\langle \subseteq, L \rangle.$

Теорема 6.3 означает, что для любого пространства толерантности $\langle M, \tau \rangle$ существует *k*-гомоморфизм пространства $\langle M, \tau \rangle$ в $\langle S_H, \hat{\tau} \rangle,$ где S_H—множество непустых подмножеств множества *H;* при этом $F_1 \hat{\tau} F_2$ означает, что $F_1 \cap F_2 \neq \varnothing.$ В частном случае, когда пространство $\langle M, \tau \rangle$ - безъядерное, существует даже *k* -мономорфизм пространства $\langle M, \tau \rangle$ в пространство $\langle S_H, \hat{\tau} \rangle$ (теорема 6.3'').

Рассмотренное в разделе 7 взаимоотношение между квазипорядками и порядками допускает следующую интерпретацию в новых терминах. Пусть $\langle A, M \rangle$ квазипорядок. Тогда существует *k* -эпиморфизм

$$\alpha\colon \langle A,\ M \rangle \to \langle B,\ L \rangle,$$

где *B* —нестрогий порядок, а $L = M/A \cap A^{-1}.$

Покажем теперь, что всякий гомоморфизм может быть расширен до *k* -гомоморфизма:

Лемма 8.1. *Пусть* $\alpha\colon M \to L$ — *отображение и* $\langle C, L \rangle$ — *отношение. Тогда* 1) *существует единственное отношение* $\langle D, M \rangle$ *такое, что* α *является k-гомоморфизмом отношения* $\langle D, M \rangle$ *в отношение* $\langle C, L \rangle;$ *2) для любого отношения* $\langle B, M \rangle,$ *для которого* α *является гомоморфизмом отношения* $\langle B, M \rangle$ *в отношение* $\langle C, L \rangle,\ B \subseteq D.$

Доказательство. 1) Определим отношение *D* на множестве *M* условием:

$$xDx', \quad \text{если} \quad \alpha(x)\, C\alpha(x'). \tag{8.1}$$

Очевидно, α является k-гомоморфизмом отношения $\langle D, M\rangle$ в отношение $\langle C, L\rangle$. Докажем теперь единственность. Пусть α является также k -гомоморфизмом отношения $\langle A, M\rangle$ в отношение $\langle C, L\rangle$. Пусть сначала xAx'. Поскольку α — гомоморфизм отношения $\langle A, M\rangle$ в отношение $\langle C, L\rangle$, $\alpha(x) C \alpha(x')$. Из (8.1) вытекает xDx'. Значит, $A \subseteq D$. Пусть теперь xDx'. Из (8.1) имеем $\alpha(x) C \alpha(x')$. Поскольку α является корреспонденцией отношения $\langle A, M\rangle$ в отношение $\langle C, L\rangle$, xAx'. Значит, $D \subseteq A$. Итак, $A = D$.

2) Включение $B \subseteq D$ доказывается так же, как мы в первой части доказывали включение $A \subseteq D$.

Пусть ε — тождественное отображение множества M на себя, а B и C — отношения на M. Легко видеть, что отображение

$$\varepsilon\colon \langle B, M\rangle \to \langle C, M\rangle$$

тогда и только тогда является гомоморфизмом, когда $B \subseteq C$ (см. рис. 8.1, *б*). С другой стороны, чтобы отображение ε было корреспонденцией, необходимо и достаточно обратное включение $B \supseteq C$ (см., в частности, рис. 8.2, *б*).

Рассмотрим теперь, какие свойства отношений сохраняются при различных типах отображений.

Лемма 8.2. *Пусть* $\alpha\colon \langle A, M\rangle \to \langle B, L\rangle$ *— гомоморфизм. Тогда* 1) *если* α *сюръективно и A рефлексивно, то B рефлексивно;* 2) *если B антирефлексивно, то A антирефлексивно.*

Доказательство. 1) В силу сюръективности отображения α для всякого $y \in L$ существует прообраз x, для которого $\alpha(x) = y$. Так как выполнено xAx, по определению эпиморфизма выполнено yBy. Таким образом, рефлексивность отношения A влечет рефлексивность отношения B. 2) Пусть теперь B антирефлексивно. Предположим, что существует такой $x \in M$, что xAx. Тогда для образа $y = \alpha(x)$ справедливо yBy, что противоречит антирефлексивности отношения B. Лемма доказана.

Аналогично этому имеет место

Лемма 8.3. *Пусть* $\alpha\colon \langle A, M\rangle \to \langle B, L\rangle$ *— корреспонденция. Тогда* 1) *из рефлексивности отношения B следует рефлексивность отношения A,* 2) *если* α *сюръективно, то из антирефлексивности отношения A следует антирефлексивность отношения B.*

Доказательство. Действительно, если всегда $\alpha(x) B \alpha(x)$, то по определению корреспонденции справедливо и xAx, т. е. A рефлексивно. Если бы хотя бы для одного элемента из L выполнялось $\alpha(x) B \alpha(x)$, то A не могло бы быть антирефлексивным.

Для сохранения остальных свойств отношений необходимо, чтобы отображение α было одновременно и эпиморфизмом и корреспонденцией.

Лемма 8.4. *Если* $\alpha\colon \langle A, M\rangle \to \langle B, L\rangle$ *есть k-эпиморфизм, то B симметрично тогда и только тогда, когда симметрично A.*

Доказательство. Предположим, что *A* симметрично. Тогда если $yBy',\ y=\alpha(x),\ y'=\alpha(x'),$ то (по определению корреспонденции) *xAx'*. Отсюда *x'Ax* и (по определению гомоморфизма) *y'By*. Пусть теперь *B* симметрично, и пусть *xAx',* тогда $\alpha(x)B\alpha(x')$ (по определению гомоморфизма). Значит, $\alpha(x')B\alpha(x)$ и (по определению корреспонденции) *x'Ax*. Лемма доказана.

Из лемм 8.2, 8.3 и 8.4 непосредственно вытекает

Теорема 8.1. *Если* $\alpha\colon \langle A, M\rangle \to \langle B, L\rangle$*—k-эпиморфизм, то отношение B тогда и только тогда является толерантностью, когда A — толерантность.*

Имеет место также

Лемма 8.5. *Если* $\alpha\colon \langle A, M\rangle \to \langle B, L\rangle$*— k-изоморфизм, то B антисимметрично тогда и только тогда, когда антисимметрично A.*

Лемма 8.6. *Если* $\alpha\colon \langle A, M\rangle \to \langle B, L\rangle$*— k-эпиморфизм, то B транзитивно тогда и только тогда, когда транзитивно A.*

Доказательство. Предположим сначала, что *B* транзитивно. Пусть *xAx'* и *x'Ax''*. Тогда $\alpha(x)B\alpha(x')$ и $\alpha(x')B\alpha(x'')$; из-за транзитивности отношения *B* верно $\alpha(x)B\alpha(x'')$. Но тогда и *xAx''*. Предположим теперь, что *A* транзитивно. Пусть *yBy'* и *y'By''*. Тогда для любых прообразов *xAx'* и *x'Ax''*. Стало быть, ввиду предположенной транзитивности отношения *A, xAx''*. Значит, *yBy''*.

Из доказанных лемм вытекает

Теорема 8.2. *Если отображение* $\alpha\colon \langle A, M\rangle \to \langle B, L\rangle$ *является k-эпиморфизмом, то отношение B тогда и только тогда является эквивалентностью (квазипорядком), когда A — эквивалентность (соответственно квазипорядок).*

Теорема 8.3. *Пусть отображение* α *множества M на множество L является гомоморфизмом (корреспонденцией) отношения* $\langle A, M\rangle$ *в отношение* $\langle B, L\rangle$ *и гомоморфизмом (корреспонденцией) отношения* $\langle A_1, M\rangle$ *в отношение* $\langle B_1, L\rangle$. *Тогда отображение* α *является также гомоморфизмом (корреспонденцией) отношений* $\langle A\cup A_1, M\rangle,\ \langle A\cap A_1, M\rangle,\ \langle AA_1, M\rangle$ *в отношения* $\langle B\cup B_1, M\rangle,\ \langle B\cap B_1, M\rangle,\ \langle BB_1, M\rangle$ *соответственно.*

Доказательство предоставляем читателю.

8.2. Минимальный образ и каноническое пополнение отношения

Начнем с доказательства необходимых лемм.

Лемма 8.7. *Пусть* $\alpha\colon M \to L$ — *сюръективное отображение,* $\langle B, M\rangle$ — *произвольное отношение,* $\langle A, M\rangle$ — *такое отношение, что*

$$x_1 A x_2 \text{ равносильно } \alpha(x_1) = \alpha(x_2).$$

Тогда, 1) *если существует такое отношение* $\langle C, L\rangle$, *что* α *является* *k-эпиморфизмом отношения* $\langle B, M\rangle$ *в отношение* $\langle C, L\rangle$, *то*

$$ABA = B, \tag{8.2}$$

2) *если выполняется* (8.2), *то существует такое отношение* $\langle C, L\rangle$, *что* α *является* *k-эпиморфизмом отношения* $\langle B, M\rangle$ *в отношение* $\langle C, L\rangle$.

Доказательство. 1) Поскольку *A* рефлексивно, $B \subseteq ABA$. Докажем, что $ABA \subseteq B$. Пусть $xABAx'$. Тогда существуют такие x_1, x_2, что xAx_1, x_1Bx_2, x_2Ax'.

Из x_1Bx_2 и того, что α— гомоморфизм, вытекает

$$\alpha(x_1)\,C\alpha(x_2).$$

Из xAx_1 получаем $\alpha(x) = \alpha(x_1)$, из x_2Ax' вытекает $\alpha(x_2) = \alpha(x')$. Значит,

$$\alpha(x)\,C\alpha(x').$$

Поскольку α — корреспонденция, xBx'. Следовательно, $ABA \subseteq B$.

2) Пусть $y, y' \in L$. Поскольку α — сюръекция, существуют такие $x, x' \in M$, что $\alpha(x) = y,\ \alpha(x') = y'$.

Определим на *L* отношение *C:*

$$yCy', \text{ если } xBx'.$$

Наше определение не зависит от выбора прообразов *x, x'.* В самом деле. Пусть x_1, x'_1 — другие прообразы элементов *y, y'* соответственно. Таким образом, $\alpha(x_1) = y,\ \alpha(x'_1) = y'$. Покажем, что xBx' тогда и только тогда, когда $x_1Bx'_1$. Допустим, что xBx'. Докажем $x_1Bx'_1$. Поскольку $\alpha(x_1) = \alpha(x)$, x_1Ax. По аналогичной причине $x'Ax'_1$. Из $x_1Ax,\ xBx'$ и $x'Ax'_1$ вытекает $x_1ABAx'_1$. Из (8.2) получаем $x_1Bx'_1$. Аналогично из $x_1Bx'_1$ выводится xBx'. Из определения отношения *C*, очевидно, вытекает, что xBx' тогда и только тогда, когда $\alpha(x)\,C\alpha(x')$. Значит, α — *k*-эпиморфизм. Лемма доказана.

Легко видеть, что для произвольных эквивалентностей *A* и *B* равенство (8.2) равносильно включению $A \subseteq B$. Ясно, что (8.2) всегда выполняется при $A = E$, т.е. тогда, когда α — биекция.

Интересно выяснить, когда для отношения $\langle B, M\rangle$ существуют неинъективные k -эпиморфизмы (т. е. k -эпиморфизмы, не являющиеся k -изоморфизмами). Из леммы 8.7 следует, что отношение $\langle B, M\rangle$ допускает неинъективные k -эпиморфизмы только в томслучае, когда существует отношение эквивалентности A на множестве M (отличное от отношения равенства) такое, что $ABA = B$. В частности, при $B = E$ такого A не существует.

Определение 8.3. Пусть B — отношение. Определим отношение B^+ условием: xB^+y, если 1) xBz тогда и только тогда, когда yBz; 2) zBx тогда и только тогда, когда zBy. Таким образом, соотношение xB^+y означает, что в графе отношения B из x и y выходят стрелки к одним и тем же вершинам и в x и y входят стрелки из одних и тех же вершин. Назовем B-отношением, *ассоциированным* с B.

Тривиальным образом можно убедиться, что B^+ будет эквивалентностью для любого исходного отношения B. Отношение B^+ склеивает в классы все элементы, имеющие одинаковые связи в графе отношения B.

Отметим, что в разделе 6 мы уже рассматривали фактически переход от отношения B к отношению B^+: если τ — отношение толерантности, а θ — отношение из (6.3), то $\tau^+ = \theta$.

Лемма 8.8. *Имеет место тождество*

$$B^+BB^+ = B. \tag{8.3}$$

Доказательство. Ясно, что $B \subseteq B^+BB^+$. Докажем обратное включение: $B \supseteq B^+BB^+$. Пусть выполнено xB^+BB^+y. Тогда существуют z, w такие, что xB^+z, zBw и wB^+y. Из xB^+z и zBw вытекает xBw. Аналогично, из wB^+y и xBw вытекает xBy. Значит, $B^+BB^+ \subseteq B$, и тем самым (8.3) доказано.

Лемма 8.9. *Для того чтобы для отношения эквивалентности A и для произвольного отношения B выполнялось равенство $ABA=B$, необходимо и достаточно, чтобы было*

$$A \subseteq B^+. \tag{8.4}$$

Доказательство. Пусть выполнено (8.4). Тогда имеем, учитывая лемму 8.8, $B \subseteq ABA \subseteq B^+BB^+ = B$, т. е. $ABA= B$. Пусть теперь выполнено $ABA = B$. Допустим, что xAy. Докажем, что xB^+y. По определению отношения B^+ надо доказать, что xBz тогда и только тогда, когда yBz, и что zBx тогда и только тогда, когда zBy. Докажем, что xBz влечет yBz. Итак, пусть xBz. Поскольку A — эквивалентность, из xAy вытекает $y\,Ax$. Кроме того, zAz. Из $y\,Ax$, xBz и zAz следует $yABAz$, а значит, и yBz. Остальное доказательство предоставляем читателю. Лемма доказана.

Из леммы 8.9 и леммы 8.7 вытекает

Теорема 8.4. *Для того чтобы существовал неинъективиый k-эпиморфизм отношения $\langle B, M\rangle$, необходимо и достаточно, чтобы отношение B^+ не было отношением равенства.*

(Доказывая достаточность, надо, чтобы использовать лемму 8.7, положить $L = M/B^+$ и в качестве α взять отображение, которое каждый элемент переводит в свой класс эквивалентности.)

Легко проверить, что если B — эквивалентность, то $B^+ = B$. Если же B — толерантность, то xB^+y означает, что x и y принадлежат общему ядру. Тем самым выясняется, что пространство толерантности тогда и только тогда не допускает нетривиальных k-эпиморфизмов, когда оно безъядерное. Вообще говоря, отношение B несравнимо с ассоциированным с ним отношением B^+. Однако имеет место

Л е м м а 8.10. *Если B рефлексивно, то* $B^+ \subseteq B$.

Доказательство. Ввиду (8.3) и рефлексивности отношений B, B^+ имеем $B^+ \subseteq B^+BB^+ = B$, т. е. лемма доказана.

Рассмотрим теперь отображение $\alpha\colon\ M \to L$ и отношение B на M. Пусть $\mathfrak{M}_B(\alpha)$ - множество всех отношений C на L таких, что $\alpha\colon\ \langle B, M\rangle \to \langle C, L\rangle$ — гомоморфизм. Множество $\mathfrak{M}_B(\alpha)$ непусто, так как в него всегда входит полное отношение.

Обозначим через B^α пересечение всех отношений из $\mathfrak{M}_B(\alpha)$. По теореме 8.3 (точнее — некоторому ее обобщению) отображение

$$\alpha\colon\ \langle B,\ M\rangle \to \langle B^\alpha,\ L\rangle \tag{8.5}$$

есть гомоморфизм. Отношение B^α обладает свойством минимальности: если C таково, что $\alpha\colon\ \langle B, M\rangle \to \langle C, L\rangle$ — гомоморфизм, то $B^\alpha \subseteq C$. Легко видеть, что верно и обратное: если $B^\alpha \subseteq C$, то $\alpha\colon\ \langle B, M\rangle \to \langle C, L\rangle$ — гомоморфизм.

Определение 8.4. Отношение B^α называется α-*образом* отношения B.

Пример на рис. 8.3 показывает, что α-образ отношения B не обязан быть эквивалентностью даже, если B — эквивалентность. (На рисунке мы не изображаем петли, которые имеются в каждой точке обоих графов.)

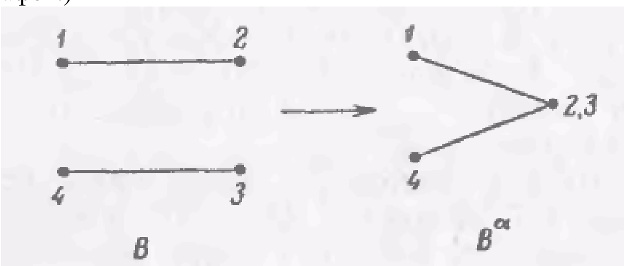


Рис. 8.3. Эквивалентность переходит в толерантность.

Лемма 8.11. *Если* α *— сюръективное отображение и B — толерантность, то* B^α *— толерантность.*

Доказательство. Рефлексивность отношения B^α следует из леммы 8.2. Симметричность отношения B^α получается так. Докажем сначала, что если $yB^\alpha y'$, то хотя бы для одной пары прообразов верно *xBx'.* В противном случае мы могли бы взять отношение *C* на *L* такое, что $y\tilde{C}y'$ не выполнено, а для всех остальных пар $y_1\tilde{C}y_2$ тогда и только тогда, когда $y_1B^\alpha y_2$.

Очевидно, что $\tilde{C} \subset B^\alpha$ и $\alpha\colon \langle B, M\rangle \to \langle \tilde{C}, L\rangle$ — гомоморфизм. Но это невозможно в силу минимальности отношения B^α. Теперь из *xBx′* следует $x'Bx$ и $y'B^\alpha y$.

Из примера на рис. 8.3 видно, что если *B* — эквивалентность, то B^α может оказаться только толерантностью.

Ввиду леммы 8.1 мы можем сформулировать

Определение 8.5. Пусть $\alpha\colon M \to L$—отображение и *B* — отношение на *M*. α-*пополнением* отношения *B* называется такое отношение B_α, для которого отображение

$$\alpha\colon \langle B_\alpha, M\rangle \to \langle B^\alpha, L\rangle \qquad (8.6)$$

есть *k*-гомоморфизм.

Из леммы 8.1 вытекает

$$B \subseteq B_\alpha. \qquad (8.7)$$

Теорема 8.5. *Пусть A — произвольная эквивалентность на M,* $\alpha\colon M \to M/A$ *— отображение, переводящее каждый элемент* $x \in M$ *в его класс эквивалентности по A, и B— отношение на M. Тогда*

$$B_\alpha = ABA. \qquad (8.8)$$

Доказательство. Мы используем еще раз свойство α-образа B^α, установленное в доказательстве леммы 8.11: соотношение $yB^\alpha y'$ выполнено тогда и только тогда, когда существует пара прообразов *x* и x' $(\alpha(x) = y,\ \alpha(x') = y')$ таких, что *xBx'.* Пусть $xB_\alpha x'$. Поскольку α является гомоморфизмом отношения $\langle B_\alpha, M\rangle$ в отношение $\langle B^\alpha, L\rangle$, $\alpha(x)B^\alpha\alpha(x')$. Тогда ввиду вышеуказанного свойства α-образа B^α существуют такие $z, z' \in M$, что $\alpha(z) = \alpha(x),\ \alpha(z') = \alpha(x')$ и zBz'.

Из $\alpha(z) = \alpha(x)$ вытекает *xAz.* По аналогичной причине *z'Ax'.* Из xAz, zBz' и $z'Ax'$ вытекает $xABAx'$.

Аналогично доказывается, что $xABAx'$ влечет $xB_\alpha x'$.

Тем самым (8.8) доказано.

Из (8.8) легко следует, что α-пополнение B_a удовлетворяет условию (8.2). Действительно,

$$AB_\alpha A = A(ABA)A = A^2BA^2 = ABA = B_\alpha.$$

Итак,

$$B_\alpha = AB_\alpha A. \tag{8.9}$$

Теперь мы исследуем, когда α-пополнение B_α принадлежит тому же типу, что и исходное отношение *В.* Благодаря (8.8) вопрос сводится к чисто алгебраической задаче: когда произведение *ABA*, где *А*—эквивалентность, принадлежит к тому же типу, что и *В?*

Сначала исследуем случай, когда *В* — отношение эквивалентности. Простой алгебраический критерий дает

Теорема 8.6. *Если А и В — эквивалентности на множестве М, то для того, чтобы произведение ABA было эквивалентностью, необходимо и достаточно выполнение равенства*

$$ABA = A \hat{\bigcup} B. \tag{8.10}$$

Доказательство. Так как $A \hat{\bigcup} B$—эквивалентность, то условие (8.10) достаточно. Пусть теперь *ABA* — эквивалентность. Из очевидных включений $A \subseteq ABA$ и $B \subseteq ABA$ следует $A \bigcup B \subseteq ABA.$ Беря транзитивное замыкание от обеих частей и учитывая известную теорему, получаем

$$A \hat{\bigcup} B \subseteq ABA. \tag{8.11}$$

С другой стороны, в силу

$$ABA \subseteq BAB \bigcup ABAB \bigcup ABA \bigcup BABA = (AB \bigcup BA)^2$$

имеем

$$ABA \subseteq A \hat{\circ} B. \tag{8.12}$$

Но по известной теореме $A \hat{\bigcup} B = A \hat{\circ} B;$ сравнивая включения (8.11) и (8.12), получаем (8.10). Теорема доказана.

Более простое достаточное условие состоит в том, что *AB* = *BA*. Тогда $ABA = AAB = AB,$ но по теореме Шика *AB* является эквивалентностью.

В. Б. Борщев построил пример двух отношений эквивалентности *А* и *В,* для которых $AB \neq BA$ и $ABA \neq BAB,$ но *ABA* есть эквивалентность.

Этот пример состоит в следующем. Классы по отношению *В* суть {1}, {2, 3} и {4}, а по отношению *А* — {1, 2} и {3, 4}. Легко подсчитать, что отношение *ABA* — полное. Произведения *AB, BA* и *BAB* показаны на рис. 8.4.

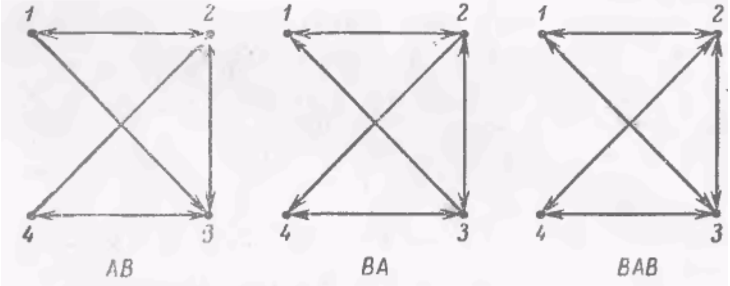


Рис. 8.4.

Можно сформулировать необходимые и достаточные условия в терминах «почти-коммутативности» отношений *A* и *B*. Соответствующий результат выглядит таким образом:

Теорема 8.7. *Если A и B — отношения эквивалентности, то произведение ABA будет эквивалентностью в том и только в том случае, когда*

$$BAB \subseteq ABA. \qquad (8.13)$$

Доказательство. Докажем сначала, что если имеет место (8.13), то *ABA* — эквивалентность. Поскольку $(ABA)^{-1} = A^{-1}B^{-1}A^{-1} = ABA$, ABA симметрично. Докажем еще транзитивность отношения *ABA*. Из (8.13) следует

$$(ABA)(ABA) = A(BAAB)A = \\ = A(BAB)A \subseteq A(ABA)A = ABA.$$

Значит, *ABA* транзнтивно. Пусть теперь *ABA* — эквивалентность. Поскольку *ABA* транзитивно. имеем $(ABA)(ABA) \subseteq ABA$. Но отсюда следует

$$BAB \subseteq A(BAB)A = (AB)A(BA) = \\ = (AB)(AA)(BA) = (ABA)(ABA) \subseteq ABA.$$

Теорема доказана.

Сходный результат получается и для порядков. Он выглядит следующим образом:

Теорема 8.8. *Пусть A — эквивалентность, B — строгий порядок. Для того чтобы произведение ABA являлось строгим порядком, необходимо и достаточно, чтобы выполнялись условия*

$$\begin{cases} BAB \subseteq ABA, \\ A \cap B = \varnothing. \end{cases} \qquad (8.14)$$

Доказательство. 1) Достаточность. Пусть выполнено (8.14). Сначала докажем антирефлексивность отношения *ABA*. Предположим, что для некоторого *x* выполнено *xABAx*. Тогда существуют *y* и *z* такие, что одновременно *xAy, yBz* и *zAx*. Из *zAx* и *xAy* следует *zAy*, а затем и *yAz*. Итак, выполнено *yBz* и *yAz*. Но это невозможно в силу $A \cap B = \varnothing$. Ввиду полученного противоречия антирефлексивность отношения *ABA* доказана. Транзитивность произведения *ABA* доказывается в точности, как в предыдущей теореме.

2) Необходимость. Пусть *ABA* является строгим порядком. Предположим, что $A \cap B \neq \varnothing$, т. е. что существует пара элементов *x* и *y* такая, что выполнены одновременно соотношения *xAy* и *xBy*. Тогда выполнена тройка соотношений: *xAx, xBy* и *yAx,* т. е. выполнено *xABAx* и *ABA* не является тем самым строгим порядком. Отсюда следует, что $A \cap B = \varnothing$. Включение $BAB \subseteq ABA$ доказывается в точности, как в предыдущей теореме.

Теорема доказана.

Рассуждая аналогично, читатель легко сумеет доказать, что если *B* — квазипорядок и *A*—эквивалентность, то *ABA* будет квазипорядком в том и только том случае, когда $BAB \subseteq ABA$.

Подведем некоторый итог проделанной работе. Пусть имеется отношение $\langle B, M\rangle$ и отображение α множества *M* в некоторое множество *L:*

$$\alpha\colon\ M \to L.$$

Тогда на множестве L однозначно определяется *минимальный образ* B^α отношения *B*. Иными словами, по отношению *B* и отображению α строится отношение B^α на L, так что отображение

$$\langle B,\ M\rangle \xrightarrow{\alpha} \langle B^\alpha,\ L\rangle$$

оказывается гомоморфизмом, обладающим следующим свойством: если *D* — некоторое отношение на *L,* то отображение $\alpha\colon\ M \to L$ будет гомоморфизмом $\alpha\colon\ \langle B, M\rangle \to \langle D, L\rangle$ в том и только том случае, когда

$$D \supseteq B^\alpha.$$

Гомоморфизм α отношения $\langle B, M\rangle$ в его минимальный образ $\langle B^\alpha, L\rangle$, вообще говоря, не является корреспонденцией. Однако существует единственное *каноническое пополнение* B_α отношения *B,* $B_\alpha \supseteq B$, для которого отображение

$$\alpha\colon\ \langle B_\alpha,\ M\rangle \to \langle B^\alpha,\ L\rangle$$

является *k*-гомоморфизмом.

Иными словами, отображение $\alpha\colon\ M \to L$ для каждого отношения *B* на *M* «вкладывается» в диаграмму:

$$\begin{tikzcd} \langle B, M\rangle \arrow[r, "\alpha"] \arrow[d, "\varepsilon"'] & \langle B^{\alpha}, L\rangle \\ \langle B_{\alpha}, M\rangle \arrow[ur, "\alpha"] & \end{tikzcd}$$

Здесь $\varepsilon\colon\ M \to M$— тождественное отображение, задающее гомоморфизм $\langle B,\ M\rangle \xrightarrow{\varepsilon} \langle B_{\alpha},\ M\rangle,$ горизонтальная стрелка является гомоморфизмом, а диагональная стрелка является k-гомоморфизмом. Отношения B^{α} и $B_{\alpha} \supseteq B$ определены однозначно по отношению B и отображению α. Подчеркнем важность формулы (8.8), выражающей явным образом каноническое пополнение B_{α} отношения B через отношение B.

Полученные в этом параграфе результаты имеют некоторое значение для математической теории классификационных систем.

Всякая классификация элементов некоторого множества M основана на выборе системы разбиений этого множества на классы. Тем самым на M возникает некоторая система отношений эквивалентности $S = \{A_1, A_2, \ldots\}$. Любое из отношений эквивалентности, принадлежащих S, например, A_1, порождает сюръективное отображение

$$\alpha\colon\ M \to L,$$

где L — множество классов эквивалентности по A_1 и отображение α сопоставляет каждому элементу из M соответствующий класс эквивалентности по A_1. Таким образом, xA_1y равносильно тому, что $\alpha(x) = \alpha(y)$. Тогда отношение $A_2 \in S$ индуцирует на множестве L отношение A_2^{α}. Из леммы 8.11 и примера на рис. 8.3 видно, что отношение A_2^{α}, индуцируемое эквивалентностью A_2 на классах эквивалентности по другому отношению, может оказаться только толерантностью. Стало быть, при описании достаточно сложной классификационной системы нельзя ограничиться отношениями эквивалентности, а приходится рассматривать более общие отношения толерантности. Это связано с тем, что в классификационных системах всегда изучаются не только отношения между самими объектами, но и отношения между классификационными рубриками, являющимися по существу классами эквивалентности по одному из отношений, характеризующих классификационную систему.

9. Примеры решения типовых задач

Отношение эквивалентности

Пример 1. 1) Отношение равенства E на любом множестве является отношением эквивалентности. Равенство — это минимальное отношение эквивалентности в том смысле, что при удалении любой пары из E (т.е. любой единицы из диагонали матрицы E) оно перестает быть рефлексивным и, следовательно, уже не является эквивалентностью.

2) Утверждения вида $(a + b)(a — b) = a^2 — b^2$, или $\sin^2 x + \cos^2 x = 1$, состоящие из формул, соединенных знаком равенства, задают бинарное отношение на множестве формул, описывающих суперпозиции элементарных функций. Это отношение обычно называется отношением равносильности и определяется следующим образом: формулы равносильны, если они задают одну и ту же функцию. Равносильность, хотя и обозначается знаком =, отличается от отношения равенства E, так как оно может выполняться для различных формул (можно считать, что знак равенства в таких соотношениях относится не к формулам, а к функциям, которые ими описываются). Отношение E для формул — это совпадение формул по написанию. Оно называется *графическим равенством.*

3) Рассмотрим множество треугольников на плоскости, считая, что треугольник задан, если заданы координаты его вершин. Два треугольника называются конгруэнтными (иногда их называют просто равными), если они при наложении совпадают, т.е. могут быть переведены друг в друга путем некоторого перемещения. Конгруэнтность является отношением эквивалентности на множестве треугольников.

4) Отношение «иметь один и тот же остаток от деления на 7» является эквивалентностью на N. Это отношение выполняется, например, для пар (11, 46), (14, 70) и не выполняется для пар (12, 13), (14, 71).

Пусть на множестве M задано отношение эквивалентности R. Осуществим следующее построение. Выберем элемент $a_1 \in M$ и образуем класс (подмножество M) C_1, состоящий из a_1 и всех элементов, эквивалентных a_1; потом выберем элемент $a_2 \notin C_1$ и образуем класс C_2, состоящий из a_2 и всех элементов, эквивалентных a_2, и т.д. Получится система классов C_1, C_2 (возможно, бесконечная) такая, что любой элемент из M входит хотя бы в один класс, т.е. $\bigcup_{i} C_i = M$. Эта система классов имеет следующие свойства: 1) она образует разбиение, т.е. классы, которые попарно не пересекаются; 2) любые два элемента из одного класса эквивалентны; 3) любые два элемента из разных классов неэквивалентны. Все эти свойства следуют из рефлексивности, симметричности и транзитивности R. Действительно, если бы классы, например C_1 и C_2, пересекались, то они имели бы общий элемент b, эквивалентный a_1 и a_2, но тогда через транзитивность R было бы a_1Ra_2, что противоречит построению C_2. Аналогично доказываются другие две свойства.

Построенное разбиение, т.е. система классов называется системой *классов эквивалентности* по отношению R. Мощность этой системы называется *индексом* разбиения. С другой стороны, любое разбиение M на классы определяет некоторое отношение эквивалентности, а именно, отношение «входить в тот же класс данного разбиения».

Пример 2. 1) Все классы эквивалентности по отношению равенства E состоят из одного элемента.

2) Формулы, которые описывают одну и ту же элементарную функцию, находятся в одном классе эквивалентности по отношению равносильности. В этом примере счетны само множество формул, множество классов эквивалентности (т.е. индекс разбиения) и каждый класс эквивалентности.

3) Разбиение множества треугольников по отношению конгруэнтности имеет континуальный индекс, причем каждый класс также имеет мощность континуум.

4) Разбиение N по отношению «иметь общий остаток от деления на 7» имеет конечный индекс 7 и состоит из 7 счетных классов 0, 7, 14 ...; 1, 8, 15 ...; 2, 9, 16 ...; ...; 6, 13, 20 ...

Отношение порядка.

Напомним, что отношение называется *отношением нестрогого порядка,* если оно рефлексивно, антисимметрично и транзитивно. Отношение называется *отношением строгого порядка,* если оно антирефлексивно, антисимметрично и транзитивно. Оба типа отношений называются отношениями порядка. Элементы a, b называются сравнимыми по отношению порядка R, если выполняется aRb или bRa. Множество M, на которой заданы отношения порядка, называется полностью упорядоченным, если любые два элемента M сравнимы, и частично упорядоченным в противном случае.

Пример 3. 1) Отношения $\leq$ и $\geq$ для чисел являются отношениями нестрогого порядка, отношения $<$ и $>$ — отношениями строгого порядка. Оба отношения полностью упорядочивают множества *N* и *R*.

2) Определим отношение $\leq$ и $<$ на R^n следующим образом:

$(a_1 \ \ldots, \ a_n) \leq (b_1, \ \ldots, \ b_n)$, если $a_1 \leq b_1, \ldots, a_n \leq b_n$;

$(a_1 \ \ldots, \ a_n \ ) < (b_1, \ \ldots, \ b_n)$, если $(a_1, \ \ldots, \ a_n) \leq (b_1, \ \ldots, \ b_n)$ и хотя бы в одной координате i выполнено $a_i < b_i$. Эти отношения определяют частичный порядок на R^n: (5, 1/2, —3) < (5, 2/3, —3); (5, 1/2, -3) и (5, 0, 0) не сравнимы.

3) На системе подмножеств множества *М* отношение включения $\subseteq$ задает нестрогий частичный порядок, а отношение строгого включения $\subset$ задает строгий частичный порядок. Например, $\{1, 2\} \subset \{1, 2, 3\}$, а $\{1, 2\}$ и $\{1, 3, 4\}$ не сравнимы.

4) Отношение подчиненности на предприятии задает строгий частичный порядок. В нем несравнимыми являются сотрудники разных отделов.

5) Пусть в списке букв конечного алфавита *А* порядок букв зафиксирован, т.е. всегда один и тот же, как, например, в русском или латинском алфавите или в алфавите цифр. Тогда этот список определяет полное упорядочение букв, которые назовем отношением предшествования и обозначим $\sqsubset$ ($a_i \prec a_j$, если a_i предшествует a_j в списке букв). На основе отношения предшествования букв строится отношение предшествования слов, определяемое следующим образом. Пусть даны слова $\alpha_1 = a_{11} \ldots a_{1m}$ и $\alpha_2 = a_{21} \ldots \ldots a_{2m}$. Тогда $\alpha_1 \prec \alpha_2$, если и только если либо

l) $\alpha_1 = \beta a_i \gamma$, $\alpha_2 = \beta a_j \delta$ и $a_i \prec a_j$ (β, γ, δ — некоторые слова, возможно, пустые, a_i и a_j — буквы), либо

2) $\alpha_2 = \alpha_1 \beta$, где β — непустое слово.

Это отношение задает полное упорядочение множества всех конечных слов в алфавите *А,* которе называется *лексикографическим упорядочением* слов.

Наиболее известным примером лексикографического упорядочения является упорядочение слов в словарях. Например, лес $\prec$ лето (случай 1 определения: β=ле,с, $\prec$ т, γ пусто, δ =о), поэтому слово «лес» расположено в словаре раньше слова «лето»; лес $\prec$ лесть (случай 2 определения: β=ть).

Если рассматривать числа в позиционных системах счисления (например, в двоичной или десятичной) как слова в алфавите цифр, то их лексикографическое упорядочение совпадает с обычным, если все сравнимые числа имеют одинаковое число разрядов. В общем же случае эти два вида упорядочения могут не совпадать: например, 10 < 1073 и 20 < 1073, но 10 ⊏ 1073, а 20 f 1073. Для того чтобы они совпадали, нужно выравнять число разрядов у всех сравниваемых чисел, приписывая слева нули. В данном примере при этом получим 0020 p 1073. Такое выравнивание автоматически происходит при записи целых чисел в ЭВМ.

Пример 4. Пусть задано произвольное множество A. Тогда отношение $\subseteq$ на $\mathcal{P}(A)$ есть тривиальное отношение порядка. ($X \subseteq X$ для всех X; если $X \subseteq Y$ и $Y \subseteq X$, то $X = Y$; $X \subseteq Y$ и $Y \subseteq Z \Rightarrow X \subseteq Z$.).

Пример 5. На основе порядка, определенного на N, мы можем формально получить обычные отношения порядка на множествах чисел Z, Q и R.
Сначала рассмотрим Z. Чтобы облегчить рассуждения, разобьем Z таким образом:

$$Z = N \cup \{0\} \cup A.$$

Поэтому $A=\{-x:\ x \in N\}$. Определим отношение (которое будем называть полным отношением порядка) на Z рассмотрением всевозможных элементов x и y из разбиения $\{N \cup \{0\} \cup A\}$. .

Если $x = y$, то $x \le y$ и $y \le x$. Пусть $x \ne y$. Тогда:

а) если $x,y \in$ N, то порядок в Z тот же самый, что и в N;

б) если $x,y \in A$, то $x \le y$ тогда и только тогда, когда $-y \le -x$ в N (т.е. —5≤ -4, так как 4≤5);

в) если $x = 0$ и $y \in$ N, то $x \le y$;

г) если $x \in A$ и $y = 0$, то $x \le y$;

д) если $x \in A$ и $y \in$ N, то $x \le y$ или в противном случае $y \le x$.

На основе порядка на Z и обычных арифметических операций с целыми числами мы можем определить порядок на Q:

$a/b \le c/d$ тогда и только тогда, когда $a * d \le b * c$.

Наконец, определим отношение порядка на множестве действительных чисел R. Рассмотрим десятичные представления двух действительных положительных чисел:

$$D = \ldots 0d_n \ldots d_2 d_1 d_0 \delta_1 \delta_2 \ldots,$$

$$C = \ldots 0c_m \ldots c_2 c_1 c_0 \gamma_2 \gamma_1 \ldots\ldots$$

Если $d_i=c_i$ и $\delta_i=\gamma_i$ для всех i, то $D=C$ и, следовательно, $D \le C$ и $C \le D$. В противном случае:

а) если $d_n \ne 0$, $c_m \ne 0$ и $n \ne m$, то $D \le C$, если $n<m$, и $C \le D$, если $m < n$;

б) если $n=m$ и $d_i \ne c_i$, но $d_j=c_j$ для всех j таких, что $i<j \le n$, то из $d_i < c_i$ следует, что $D \le C$, и, обратно, если $c_i < d_i$, то $C \le D$;

в) если $m= n$ и $d_i=c_i$ для всех i, но $\delta_k \neq \gamma_k$ для некоторых k и $\delta_j=\gamma_j$ для всех j таких, что $0 < j < k$, тогда $C \leq D$ если $\delta_k<\gamma_k$ и $D \leq C$, если $\delta_k<\gamma_k$. Отрицательные числа могут быть исследованы так же, как в Z.

Отношение толерантности.

Отношение толерантности τ на множестве M удовлетворяет свойствам рефлексивности и симметричности. Упорядоченная пара (x, y) принадлежит множеству $\tau \subset M \times M$, если: 1) $x\tau x$ и 2) из $x\tau y$ следует $y\tau x$. Для этого отношения, в отличие от эквивалентности, транзитивность не обязательна, и значит эквивалентность есть частный случай толерантности.

Отношение толерантности представляет собой экспликацию интуитивных представлений о сходстве и неразличимости. Каждый объект неразличим сам с собой (рефлексивность), а сходство двух объектов не зависит от того, в каком порядке они сравниваются (симметричность). В то же время, если один объект сходен с другим, а другой сходен с третьим, то это вовсе не означает, что все они обязательно сходны между собой, т.е. свойство транзитивности может не выполняться.

Например, толерантность на множестве точек плоскости может быть задана свойством: расстояние между любой парой точек не превышает величины a. С этим свойством обычно связывается моделирование зрительного органа. Очевидно, толерантностью может быть острота зрения, т.е. условие того, что любые пары точек неразличимы для глаза в его поле зрения.

Развлекательным примером толерантности является популярная задача «превращение мухи в слона» (муха-мура-тура-тара-кара-каре-кафе-кафр-каюр-каюк-крюк-крок-срок-срок-сток-стон-слон). Здесь отношение толерантности определяется сходством между четырехбуквенными словами, если они отличаютсся тольки одной буквой. Если определить отношение между словами как наличие хотя бы одной общей буквы, то толерантными будут пересекающиеся слова кросворда.

Толерантность кортежей. На множестве кортежей (векторов) $x=(x_1, x_2, \ldots, x_n)$ толерантность можно задать различными способами, например, обусловить наличие в паре кортежей хотя бы одной общей компоненты x_i.

Компонентами кортежа могут быть любые объекты. Если они принимают целочисленное значение от 0 до $m - 1$, то кортеж можно рассматривать как n-разрядное число, записанное в позиционной системе счисления с основой m. Например, кортеж $x = (7, 0, 4, 9, 2)$ отвечает десятичному числу 70492. Количество всех таких кортежей, очевидно, равняется m^n.

При $m = 2$ имеем двоичный кортеж, его компоненты принимают значение 0 или 1. Для каждого $x = (x_1, x_2, ..., x_n)$ существует только один не толерантный к нему кортеж $x' = (1 — x_1, 1 — x_2, \ldots 1\text{-}x_n)$.....

Двоичный кортеж можно трактовать также как содержимое n-разрядного регистра вычислительной машины. Состояние машины определяется содержимым всех его регистров, т.е. множеством двоичных кортежей. Если два состояния машины различаются содержимым некоторого ограниченного числа регистров, то говорят, что эти состояния толерантны, а машину называют *толерантным автоматом.*

Толерантность числовых функций. Каждый кортеж $x=(x_1, x_2, ..., x_n)$ компоненты которого — некоторые действительные числа x_i, можно считать числовой функцией, заданной на множестве $(1, 2, ..., n)$. Каждому числу i $(1 \le i \le n)$ эта функция сопоставляет число x_i. Толерантность двух функций означает, что хотя бы в одной точке они принимают одинаковые значения (точка A на рис. 1).

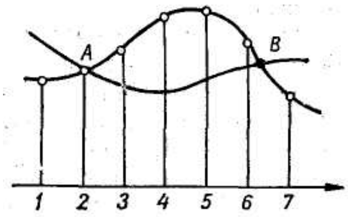


Рис. 1. Толерантность функций.

Если функции определены на некотором отрезке действительных чисел, то толерантность на множестве таких функций означает совпадения хотя бы одного из значений двух функций, соответствующих одному и тому же аргументу. Другими словами, толерантными являются функции, графики которых пересекаются (точки пересечения A и B на рис. 1).

Многомерный симплекс. Рассмотрим совокупность S_q всех непустых подмножеств множества $q + 1$ натуральных чисел

$H = \{1, 2, ..., q+1\}$. Определим на этой совокупности отношения толерантности: два подмножества толерантны если они содержат хотя бы один общий элемент. При q=0,1,2,3 множество S_q можно представить соответственно точкой, отрезком, треугольником и тетраэдром (рис. 2), отображая одноэлементные подмножества вершинами, двухэлементные — ребрами, трехэлементные — гранями и четырехэлементные - геометрическим телом (тетраэдром).

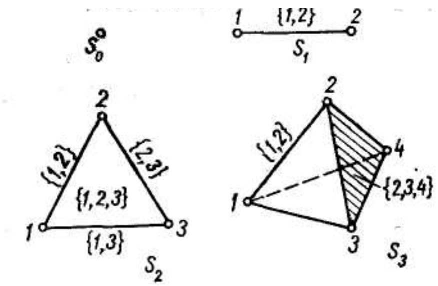


Рис. 2. Многомерные симплексы (q = 0, 1, 2, 3).

Если q>3, то геометрическое представление множества S_q в обычном трехмерном пространстве теряет наглядность, но может быть формально продолженно в абстрактном пространстве, которое имеет q измерений.

Множество S_q называют q-мерным *симплексом.* Симплекс обобщает понятие отрезка, треугольника и тетраэдра на многомерный случай. Подмножества, которые содержат k+1 элемент, рассматриваются как *k-мерные грани.* Толерантность граней симплекса (наличие общих вершин) означает их *геометрическую инцидентность.*

Толерантность в множестве подмножеств. Пусть H— произвольное конечное множество, элементами которого могут быть объекты любой природы (предметы, числа, фигуры, свойства и т.п.), и $S_н$ — множество всех ее непустых подмножеств. Если H содержит q элементов, количество элементов в $S_н$ равно 2^q — 1 (вычитаемая единица соответствует пустому подмножеству универсума H).

Толерантность в множестве $S_н$ можно задать условием: два подмножества $X, Y \in S_н$ $(X \subset H$ и $Y \subset H)$ толерантны, если они содержат хотя бы один общий элемент. Это значит, что $X \tau Y$ при условии

$$X \cap Y \neq \varnothing .$$

Пусть, например, $H=\{\alpha_1,\alpha_2,\alpha_3,\alpha_4\}$ и заданы подмножества $X=\{\alpha_1,\alpha_2\}$;$Y=\{\alpha_1,\alpha_3,\alpha_4\}$; $Z=\{\alpha_4\}$; согласно определению $X\tau Y$ и $X\tau Z$, но X и Z не толерантны, так как ни один из элементов из X не содержится в Z.

Сходство как толерантность. Сходство между различными объектами имеет точный смысл только тогда, когда указанная совокупность признаков, относительно которого это сходство устанавливается. Два объекты считаются *сходными (толерантными)*, если они обладают хотя бы одним общим признаком.

Пусть $M=\{x_1,x_2,...,x_m\}$ множество объектов и $H=\{\alpha_1,\alpha_2,...,\alpha_г\}$ – множество признаков. Каждому элементу x_i из M отвечает некоторое подмножество H_i признаков ($H_i \subset H$). Следовательно, сходство между объектами x_i и x_j определяется толерантностью соответствующих им подмножеств H_i и H_j из H, т.е. условием $H_i \cap H_j \neq \varnothing$.

Если рассматривать соответствие между объектами и признаками как бинарное отношение A между множеством объектов M и множеством признаков H, то элементами A будут упорядоченные пары $(x,\alpha)\in A$, в каждой из которых первая координата является объектом, а вторая — признаком. Очевидно, множество H_i признаков объекта x_i является сечением $A(x_i)$ этого отношения, т.е. $H_i=A(x_i)$. Как было показано ранее, все такие сечения полностью определяют бинарное отношение A.

Итак, толерантность τ на множестве объектов M можно задать с помощью некоторого всюду определенного бинарного отношения A от M к H следующим образом: для любой пары объектов x_i и x_j из M имеет место $x_i\tau\, x_j$, если и только если $A(x_i)\cap A(x_j) \neq \varnothing$.

Действительно, отношение τ симметрично, так как из $x_i\tau x_j$ следует $A(x_i)\cap A(x_j)\neq\varnothing$, но $A(x_i)\cap A(x_j) = A(x_j)\cap A(x_i)$, следовательно, $x_j\tau x_i$. Оно также рефлексивно, поскольку отношение A определено на всем M. В этом и только в этом случае множество $A(x_i)\cap A(x_i)=A(x_i)$ не пусто для любого $x_i\in M$. Следовательно, τ — отношение толерантности.

Классы толерантности. Множество $L\subset M$, любые два элемента которого толерантны, называют *предклассом толерантности.*

Если толерантность связана с отношением A, то обратное отношение A^{-1} устанавливает соответствие между признаками и объектами. Каждому признаку $\alpha_i\in H$ соответствует некоторая совокупность объектов из M, которые обладают этим признаком. Такая совокупность определяется сечением $A^{-1}(\alpha_i)$ и является предклассом. Следовательно, множество всех различных сечений $A^{-1}(\alpha_i)=L_i$ образует некоторое множество предклассов, причем каждый объект из M входит хотя бы в один предкласс.

Некоторые из предклассов могут быть связаны отношениям включения $L_i \subset L_j$. Если некоторый предкласс не является подмножеством никакого другого предкласса, то он является максимальным предклассом и называется *классом толерантности.*

Чтобы выделить из множества предклассов L_i классы толерантности, необходимо попарно сравнить все предклассы относительно включения. При этом, если $L_i \subset L_j$, то L_i отбрасывается и продолжаеться сравнение L_j с другими предклассами до тех пор, пока не останутся только предклассы, которые не связаны отношениям включения. Они и образуют совокупность классов толерантности $\{ K_1, K_2, \ldots, K_P\}$

Различные классы толерантности могут содержать одинаковые элементы и, следовательно, являются пересекающимися множествами. В то же время их объединение равно множеству M, т.е. K_1 □ $K_2 \cup \ldots \cup K_p = M$. Говорят, что классы толерантности образуют *покрытие* множества M.

Пусть, например, $M=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7)$ и $H=\{\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5\}$, причем объект x_1 наделен признаками α_3 и α_5, т.е. $H_1 = \{\alpha_3, \alpha_5\}$ и аналогично $H_2 = \{\alpha_2, \alpha_4\}$; $H_3 = \{\alpha_3, \alpha_5\}$; $H_4 = \{\alpha_3, \alpha_4\}$; $H_5 = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$; $H_6 = \{\alpha_3, \alpha_5\}$; $H_7 = \{\alpha_4\}$.

Соответствующее отношение выражается матрицей

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
α_1					1		
α_2		1		1	1		
α_3	1		1		1	1	
α_4		1		1	1		1
α_5	1		1			1	

Отсюда определяем предклассы как сечения обратного отношения:
$L_1 = \{x_5\}$, $L_2 = \{x_2, x_4, x_5\}$, $L_3 = \{ x_1, x_3, x_5, x_6\}$; $L_4 = \{x_2, x_4, x_5, x_7\}$; $L_5 = \{x_1, x_3, x_6\}$.

Так как $L_1 \subset L_2 \subset L_3 \subset L_4$ и $L_5 \subset L_3$, то классами толерантности являются L_3 и L_4, т.е. $K_1=L_3=\{x_1, x_3, x_5, x_6\}$ и $K_2=L_4=\{x_2,x_4,x_5, x_7\}$

Толерантность и эквивалентность. Изложенный способ образования классов толерантности подсказывает связь между толерантностью и эквивалентностью.

В частном случае, когда каждый объект характеризуется только одним признаком (отношение A от M к H есть отображение или функция), классы толерантности объединяют только те объекты, которые обладают данным признакрм. При этом толерантность переходит в эквивалентность, а классы толерантности — в классы эквивалентности. Другая формулировка условия совпадения толерантности с эквивалентностью требуют, чтобы классы толерантности не пересекались один из одним.

Более глубокая связь между отношениями толерантности и эквивалентности устанавливается при рассмотрении подмножеств $H_i = A(x_i)$ признаков, соответствующих объекту $x_i \in M$.

Разобьем M на непересекающиеся классы, поместив в каждый класс все объекты x_i из M, для которых H_i совпадают. Тогда это разбиение $\{M_1, M_2, ...\}$ определит на множестве M отношение эквивалентности. Так, для рассмотренного ранее примера имеем

$$M_1 =\{x_1\ x_3, x_6\};\ M_2 =\{x_2, x_4\};\ M_3 =\{x_5\}, M_4 =\{x_7\}.$$

Так как эквивалентность является частным случаем толерантности, то каждый класс эквивалентности является подмножеством какого-либо класса толерантности либо совпадает с ним. Значит классы эквивалентности являются предкласами или классами толерантности (в нашем примере

$$M_1 \subset K_1, M_2 \subset K_2, M_3 \subset K_1, M_3 \subset K_2, M_4 \subset K_2).$$

Отсюда также следует, что класс эквивалентности связан отношениям включение с пересечением классов толерантности, подмножествами которых он является ($M_3 \subset K_1 \square K_2$). Если классы толерантности не пересекаются, то толерантность переходит в эквивалентность. Ясно также, что класс толерантности выражается через объединение входящих в него классов эквивалентности

$$(K_1 = M_1 \cup M_3;\ K_2 = M_2 \cup M_3 \cup M_4).$$

Матрица толерантности. Пусть толерантность определена как совокупность классов толерантности $\{K_1, K_2, ...\}$. Внутри каждого такого класса любые два элемента толерантны, следовательно, кроме рефлексивности и симметричности имеет место и транзитивность. Поэтому подобно эквивалентности, классу толерантности соответствует заполненный единичными элементами квадрат, диагональ которого располагается по главной диагонали матрицы. Но в отличие от эквивалентности эти квадраты пересекаются, так что в целом толерантность не транзитивна.

Если при записи матрицы расположить элементы множества совокупностями, соответствующими классам эквивалентности и толерантности, то для рассмотренного ранее примера имеем:

	x_1	x_3	x_6	x_5	x_2	x_4	x_7		
x_1	1	1	1	1					
x_3	1	1	1	1				M_1	
x_6	1	1	1	1					K_1
x_5	1	1	1	1	1	1	1	M_3	
x_2				1	1	1	1	M_2	K_2
x_4				1	1	1	1		
x_7				1	1	1	1	M_4	

Отношение толерантности полностью определяется его *картой*, на которой изображаются классы эквивалентности (или любое другое покрытие) и классы толерантности (рис. 3).

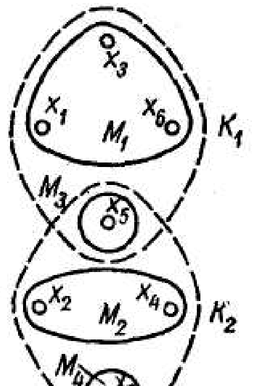


Рис. 3. Карта отношения толерантности.

Индивидуальные тестовые задачи

Отношение эквивалентности

1. Покажите, что каждое из следующих отношений является эквивалентностью:

а) подобие в множестве всех треугольников на плоскости;

б) принадлежность к одной группе в множестве студентов факультета;

в) равенство веса в множестве разновесов;

г) равномощность в произвольной системе множеств;

2. Опишите характерные свойства графика отношения в множестве действительных чисел, если это отношение:

а) рефлексивное,

б) симметричное,

в) транзитивное.

Что можно сказать о графике отношения эквивалентности?

3. Отношение эквивалентности на множестве $M=\{1, 2, 3, 4, 5, 6,7\}$ задано разбиением на классы: $M_1=\{1, 4\}$; $M_2=\{2, 3, 7\}$; $M_3=\{5, 6\}$. Представьте это отношение множеством упорядоченных пар и матрицей.

4. Укажите свойства, которыми обладают приведенные ниже отношение и объясните, почему они не являются эквивалентностями:

а) «x кратно y» в множестве целых чисел;

б) «x имеет общие точки с y» в множестве прямых на плоскости;

в) «x знаком с y» в множестве людей.

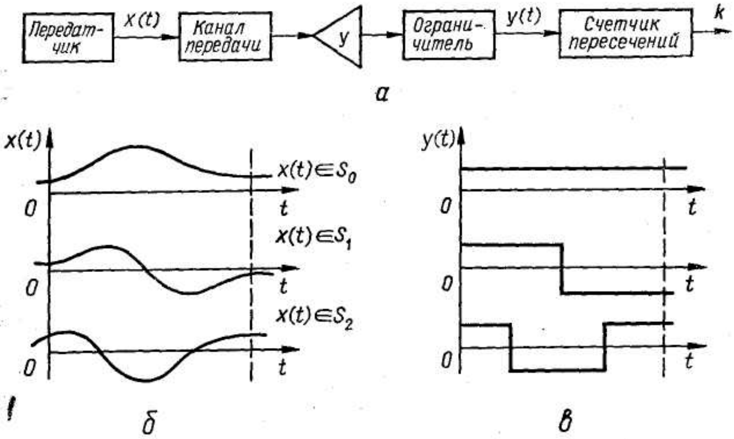
5. Пусть $S=\{x(t)|P(x)\}$ - множество сигналов $x(t)$, которые характеризуются свойством $P(x)$. Один со способов различения принимаемых сигналов состоит в подсчете числа пересечений с нулевым уровнем за определенный промежуток времени (см. рис.), что соответствует разбиению множества S на непересекающиеся подмножества:

$S_k=\{x(t)|x(f)$ имеет k несовпадающих пересечений на заданном интервале времени$\}$ ($k = 0, 1, 2, ...$).

а) Охарактеризуйте отношение, задаваемое этим разбиением, как отношение эквивалентности.

б) Укажите систему представителей данного разбиения.

в) Как следует определить классы эквивалентности, чтобы их число не превышало величины n?



Различение сигналов на основе числа пересечений нулевого уровня: *а* — система передачи сигналов; *б* — переданные сигналы; *в* — принятые сигналы (после ограничителя).

6. Охарактеризуйте отношения эквивалентности, порождаемые следующими отображениями:

а) *f(x)* - отец *x*;

б) *f(x)* — рост *x;*

в) *f(x)*— начальная буква имени *x;*

г) *f(x)* — взаимно-однозначное отображение;

д) *f(x) = a (a* — постоянная величина).

7. Покажите, что если *A* и *B* — отношение эквивалентности, то их пересечение $A \cap B$ также является эквивалентностью.

8. Покажите, что отношение A^{-1}, симметричное отношению эквивалентности *A,* также является эквивалентностью.

9. В общем случае объединение *A*☐ *У* отношений эквивалентности *A* и *B* не является эквивалентностью. Приведите пример, который подтверждает это положение.

10. Для того чтобы композиция *AB* отношений эквивалентности *A* и *B* была эквивалентностью, необходимо и достаточно, чтобы *A* и *B* коммутировали, т.е. *AB* = *BA.* Докажите это утверждение.

Отношение порядка

1. Показать, что приведенные ниже отношения являются отношениями порядка и определить тип упорядоченности:

а) « x тяжелее y» в множестве деталей;

б) «x подчинен y» в множестве должностей;

в) «x длиннее y» в множестве отрезков на плоскости;

г) «x старше y» в множестве людей;

д) «x не превосходит y» в множестве номеров домов на улице;

е) «из x следует y» в множестве высказываний;

ж) «x находится внутри y» в множестве кругов на плоскости.

2. В множестве комплексных чисел задано отношение A такое, что для любых $x, y \in$ Z имеет место соотношения xAy, если действительная часть числа x меньше действительной части числа y. Покажите, что A — квазипорядок, и опишите множество классов эквивалентности, в котором этот квазипорядок индуцирует строгий порядок.

3. Пять приборов оценивают по четырем показателям в десятибалльной системе с коэффициентом весомости α_i (i=1,2,3,4). Определите на основании приведенной таблицы комплексные показатели качества приборов и упорядочите согласно этим показателям множество приборов. К какому типу относится полученное упорядочение?

Показатели качества	Коэффициенты весомости	Оценка приборов (в баллах)				
		1	2	3	4	5
x_1	0,2	5	4	7	8	6
x_2	0,1	6	6	4	9	3
x_3	0,5	4	6	3	5	7
x_4	0,3	9	5	10	7	8

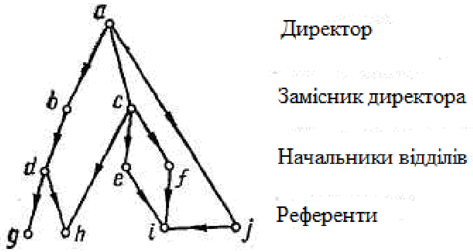
4. На рисунке показано графическое изображение организационной структуры, определяемой соотношением «x начальник y» со свойствами:

1) никто не является собственным начальником;

2) никто не может быть одновременно начальником и подчиненным другого лица;

3) начальник начальника некоторого лица является начальником этого лица.

Определите тип упорядоченности такой организации.



5. Вес (значимость) каждого лица в организации (задача 4) можно оценивать целым числом *m(x),* которое определяется по формуле

$$m(x) = \sum_k k n_k,$$

где n_i — число подчиненных уровня *k*. Уровень подчинения равен длине кратчайшего пути от вышестоящего лица *x* к подчиненному *y*. Если *x* не имеет подчиненных, то $m(x) = 0$.

а) Определите веса всех лиц в организации, заданной вышеприведенным рисунком.

б) Упорядочите множество всех лиц данной организации по их весам и укажите, к какому типу относится эта упорядоченность.

6. Может ли оказаться (при определении понятия уровня, как это сделано в задаче 5), что некоторое лицо в организации характеризуется меньшим весом, чем его подчиненный? Приведите пример.

7. Если уровень подчинения *k* определить как самый длинный путь от вышестоящего лица к подчиненному, то ситуация, которая описана в задаче 6 не может иметь места. Докажите это положение. Определите значимость всех лиц в организации (см. рис), приняв приведенное здесь определение уровня *k*.

8. Пусть *A* - подмножество действительных чисел, состоящие из числа —1 и из чисел $0 \le x < 1$. Покажите, что:

а) любое число, большее или равное 1, является мажорантой множества *A*, а любое число, которое меньшее или равное —1, — минорантой этого множества;

б) верхняя грань (супремум) равна 1, а нижняя грань (инфинум) равна—1.

9. Покажите, что если отношение *A* — строгий порядок (нестрогий порядок, квазипорядок), то симметричное ему A^{-1} также является строгим порядком (нестрогим порядком, квазипорядком).

10. Покажите, что если *A* и *B* - строгие порядки, то пересечение $A \cap B$ является строгим порядком. Распространите это положение на нестрогий порядок и квазипорядок.

11. Объединение порядков в общем случае не является порядком. Если A и B - строгие порядки, то объединение $A \cup B$ является строгим порядком, если и только если $BA \cup AB \subset A \cup B$. Для того чтобы объединение $A \cup B$ нестрогих порядков было также нестрогим порядком, необходимо и достаточно выполнение условий $BA \cup AB \subset A \cup B$ и $A \cup B^{-1} \subset$ Э. Приведите примеры.

Решетки

1. Доказать, что в каждой решетке:
   а) $(a \le b)\&(c \le d) \to a \square c \le b \cup d;$

   б) $(a \le b)\ \&\ (c \le b) \to a \cup c \le b;$

   в) $(a \le b) \to (\forall c)(a \cup c \le b \cup c);$

   г) $(a \le b)\&(c \le d) \to a \cap c \le b \cap d;$

   д) $(a \le b)\&(a \le c) \to a \le b \cap c;$

   е) $(a \le b) \to (\forall c)(a \cap c \le b \cap c).$

2. Докажите, что в каждой решетке для любых элементов a, b, c:

   а) $a \cup b \cap c \le (a \cup b) \cap (a \cup c);$

   б) $a \cap (b \cup c) \ge a \cap b \cup a \cap c;$

   в) $a \le b \to a \cup b \cap c \le (a \cup c) \cap b.$

3. Докажите, что решетка модулярна тогда и только тогда, когда для любых элементов a, b, c

$$a \cap b \cup c \cup (a \cap b) = (a \cap b \cup c) \cap (a \cup b).$$

4. Докажите, что решетка дистрибутивна тогда и только тогда, когда для любых трех элементов a, b, c

$$a \cap (b \cup c) \cup b \cap c = (c \cup a \cap b) \cap (a \cup b).$$

5. Найти множество «запрещенных фигур», выводящее частично упорядоченное множество из класса решеток.

Отношение толерантности

1. Какие из приведенных ниже отношений являются толерантностями?
   а) «x перпендикуляр на y» в множестве прямых;
   б) «x знаком с y» в множестве людей;

в) «x имеет общие точки с y» в множестве геометрических фигур;

г) «x рядом с y» в множестве книг на полке.

2. Пусть $M=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8\}$ - множество объектов и $H=\{\alpha_1,\alpha_2,\alpha_3,\alpha_4,\alpha_5,\alpha_6\}$ - множество признаков. Бинарное отношение задано матрицей

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
α_1		1	1	1	1			
α_2		1	1					
α_3	1					1		
α_4			1		1			
α_5		1	1		1		1	1
α_6	1	1				1		

а) определить предклассы и классы толерантности;

б) определить классы эквивалентности, которые связаны с отношением;

в) начертить карту толерантности.

3. Совокупность классов толерантности $K=\{K_1,K_2, ...\}$ в множестве M образует базис, если:

1) для всякой толерантной пары $x, y \in M$ существует класс толерантности из K, который содержит эти элементы;

2) удаление из K хотя бы одного класса приводит к потере этого свойства. Иначе говоря, для всякого $K_i \in K$ существует толерантная пара x, y, для которой K_i — единственный общий класс толерантности в K. Покажите, что для рассматриваемого примера существует только два базиса (найдите другой базис).

4. Пусть M — множество n-разрядных десятичных чисел и K_{ij} — подмножества этого множества $(K_{ij} \subset M)$, состоящие из всех тех чисел, которые в i-м разряде содержат цифру j (i =1, 2,..,n; j = 0, 1,..., 9).

а) Покажите, что K_{ij} — классы толерантности, образующие базис;

б) Сколько элементов содержит каждый из классов толерантности K_{ij} и сколько всего есть таких классов?

5. Покажите, что для любого рефлексивного отношения A отношение $A \cup A^{-1}$ и $A \cap A^{-1}$ являются толерантностями.

10. Элементы комбинаторного анализа

В дискретной математике мы по большей части имеем дело с конечными или конечно порожденными множествами и с системами конечных множеств — отношениями. Для того чтобы составить ясное представление о множестве (в том числе для построения математической модели множества), обычно надо выполнить анализ информации, составляющей описание этого множества. Типичная цель такого анализа — описание множества с помощью комбинаторных операций над некоторыми более простыми множествами, базисом, или даже в прямом перечислении его элементов. Особое место занимают задачи на подсчет или оценку мощности конечных множеств.

Соответствие между комбинаторными операциями над множествами и арифметическими операциями над их мощностями (эти мощности часто называют комбинаторными функциями) есть специальный вид соответствия между рекурсивными схемами задания множеств и функций. Можно сказать, что комбинаторный анализ является прикладпым разделом математической теории функциональных систем с операциями, в котором трактуются вопросы практической техники анализа дискретных множеств и отношений. Образно говоря, комбинаторный анализ относится к дедуктивным разделам математики так, примерно, как тактика в шахматной теории относится к правилам игры и основам стратегии: богаче конкретным содержанием, но беднее общностью и универсальностью понятий и формулировок результатов.

10.1. Комбинаторные операции и функции

1. Имеются два основных типа операций над множествами. К первому типу относятся так называемые *алгебраические* операции, такие как объединение $A_1 \cup A_2 \cup \ldots \cup A_k = \bigcup\limits_{i=1}^{n} A_i$ или пересечение

$A_1 \cap A_2 \cap \ldots \cap A_k = \bigcap\limits_{i=1}^{h} A_i$ нескольких множеств, а также теоретико-множественная разность $A \backslash B = \{x \backslash x \in A,\ x \notin B)$ и ее частный случай — дополнение $\overline{A} \Rightarrow B\backslash A$ множества A до множества B, когда $A \subseteq B$ и из контекста ясно, о каком B идет речь, симметрическая разность $A \otimes B \Rightarrow (A\backslash B) \cup (B\backslash A)$. Для операций этого типа характерно, что результирующее множество состоит из тех же элементов, из которых составлены множества, которые подвергаются операции, либо пусто. Операции другого типа называют *кардинальными,* при их применении возникают новые элементы. Таковы: прямое произведепие $A_1 \times A_2 \times ... \times A_k$, элементами которого являются всевозможные упорядоченные наборы вида $<a_1, a_2,...,a_k>$, где $a_1 \in A_1$, $a_2 \in A_2$,..., $a_k \in A_k$ (иногда эту операцию рассматривают как ассоциативную, иногда — нет, это всегда ясно из контекста; в частости, использование $A^k \Rightarrow A \times A \times ... \times A$ $_k$ раз подразумевает обычно ассоциативный вариант) (в ассоциативном случае элементы $A_1 \times A_2 \times ... \times A_k$ называют также словами и записывают в виде $a_1a_2...a_k$. Связную часть a_ia_{i+1} ... a_{i+j} слова $a_1a_2...a_k$ называют фрагментом этого слова; если i = 1, то префиксом, а если $i+j = k$, то суффиксом), булевская степень 2^A — множество всех подмножеств A:$2^A \Rightarrow \{X|X \subseteq A\}$ и кардинальная степень A^B—множество всех функций с областью определения B и областью значений A (т. е. отображений, сопоставляющих каждому $b \in B$ единственный элемент $f(b) \in A$).

2. Один из способов задания множеств при определенном универсальном множестве U (универсе) есть *характеристические функции.* По существу этот способ — одна из форм задания множества свойством его элементов. Характеристической функцией множества A в универсе U называется функция $s_A \in \{0, 1\}^U$, определенная правилом

$$s_A(x) = s_A^{(U)}(x) = \begin{cases} 1, & \text{если } x \in A, \\ 0, & \text{если } x \notin A. \end{cases}$$

Любую функцию, принимающую только два значения, скажем, f из $\{0, 1\}^U$ можно рассматривать как характеристическую функцию множества $N_f \Rightarrow \{x|f(x) = 1\}$, так как, очевидно, $s_{N_f}(x) \equiv f(x)$ (и равным образом $N_{s_A} \equiv A$) Следующее соотношение выражает фундаментальный принцип подсчета *мощности* конечного множества:

$$|A| = \sum_{x \in U} s_A(x) \tag{1}$$

3. Соотношения, перечисляемые в этом пункте, устанавливают связь алгебраических операций над множествами с арифметическими и булевскими операциями над их характеристическими функциями, после чего легко прослеживается и связь с операциями над соответствующими комбинаторными функциями:

$$s_U^{(U)}(x) \equiv 1, \; s_\varnothing^{(U)}(x) \equiv 0, \tag{2}$$

$$s_{\overline{A}}^{(U)}(x) = 1 - s_A^{(U)}(x) = \overline{s_A(x)}, \tag{3}$$

$$s_{A\cap B}(x) = s_A(x) \cdot s_B(x) = s_A(x) \,\&\, s_B(x), \tag{4}$$

$$s_{A\setminus B}(x) = s_A(x) - s_{A\cap B}(x) = s_A(x) \,\&\, \overline{s_B(x)}, \tag{5}$$

$$s_{A\otimes B}(x) = s_A(x) + s_B(x) - 2s_{A\cap B}(x) = s_A(x) \oplus s_B(x). \tag{6}$$

Проверка соотношений (2) — (6) вполне тривиальна, как и переход к комбинаторным функциям на основе (1). Проследим этот переход на одном примере:

$$s_{A\cup B}(x) = s_A(x) + s_B(x) - s_{A\cap B}(x)$$

влечет

$$|A \cup B| = \sum_{x\in U} s_{A\cup B}(x) = \sum_{x\in U} s_A(x) + \sum_{x\in U} s_B(x) - \\ - \sum_{x\in U} s_{A\cap B}(x) = |A| + |B| - |A \cap B|.$$

Используя индукцию по k, этот результат обобщается на случай объединения любого конечного числа k множеств:

$$|A_1 \cup A_2 \cup \ldots \cup A_k| = \\ = \sum_{\sigma=1}^{k} (-1)^{\sigma} \sum_{(j_1,\ldots,j_\sigma)} |A_{j_1} \cap \ldots \cap A_{j_\sigma}|, \tag{7}$$

где внутренняя сумма распространяется на всевозможные σ-элементпые множества индексов. Метод подсчета на основе (7) называют *методом включения и исключения.*

4. Своеобразие обозначений для кардинальных операций находит оправдание в связи с соответствующими операциями над комбинаторными функциями:

$$|A_1 \times A_2 \times \ldots \times A_k| = |A_1| \cdot |A_2| \ldots |A_k|, \tag{8}$$

$$|2^A| = 2^{|A|}, \quad |A^B| = |A|^{|B|}. \tag{9}$$

Правило произведения (8) доказывается легко — при k=2 индукцией по мощности одного из множеств и затем индукцией по k. Заметим, что (8) сохраняет силу независимо от того, рассматриваем ли мы операцию прямого произведения как ассоциативную или нет, т. е.

$$|A_1 \times (A_2 \times A_3)| = |(A_1 \times A_2) \times A_3| = |A_1 \times A_2 \times A_3|.$$

Соотношения (9) доказаны в п. 6.

5. Наиболее общее представлению о функциях связано с табличным способом задания функций. Для функции $f \in A^B$ часто рассматривают ее «естественное» продолжение на 2^B, полагая $f(\varnothing) = \varnothing$, и для непустых $C \subseteq B$ —

$$f(C) = \bigcup_{x\in C} \{f(x)\}. \tag{10}$$

Важным случаем функциональной связи между множествами являются *взаимно однозначные соответствия.* Между множествами A и B взаимно однозначное соответствие (если оно существует) устанавливается функцией $f \in A^B$ такой, что $f(B) = A$, и если $x, y \in B$ и $x \neq y$, то $f(x) \neq f(y)$. Для бесконечных множеств это понятие позволяет ввести понятие мощности множества, не обращаясь к понятию «количество», а для конечных позволяет получать оценки мощности, не прибегая к прямому подсчету. Если между изучаемым множеством A и множестпом B, мощность которого известна, установлено взаимно однозначное соответствие, то это сразу приводит к определению комбинаторной функции множества A.

6. **Примеры**. (а) Функции $\nu(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i \cdot 2^{n-i}$

устанавливает взаимно однотипное соответствие между множествами $A_n = \{0, 1\}^n$ и $B_n = \{0, 1, 2, ..., 2^{n\sim 1}\}$,

Свойство $\nu(A_n) = B_n$ докажем индукцией по n. При $n = 1$ проверяем, затем, полагая справедливым для n— 1, имеем

$$\nu(A_n) = \nu(\{0\} \times A_{n-1}) \cup \nu(\{1\} \times A_{n-1}) = \\ = B_{n-1} \cup \{2^{n-1} + x \mid x \in B_{n-1}\} = B_n.$$

Второе свойство: пусть $\langle x_1, \ldots, x_n \rangle \neq \langle y_1, \ldots, y_n \rangle$ и f— наименьшее, для которого $x_j \neq y_j$, скажем, $x_j = 1$, $y_j = 0$. Тогда

$$\nu(x_1, \ldots, x_n) - \nu(y_1, \ldots, y_n) = \\ = 2^{n-j} + \sum_{i=j+1}^{n} (x_i - y_i) \cdot 2^{n-i} \geqslant 2^{n-j} - \sum_{i=j+1}^{n} 2^{n-i} = 1,$$

т. е. $\nu(x_1, \ldots, x_n) \neq \nu(y_1, \ldots, y_n)$, ч. т. д.

(б) $|2^A| = |\{0, 1\}^A| = |\{0, 1\}^{|A|}| = 2^{|A|}$. Взаимно однозначное соответствие между 2^A и $\{0, 1\}^A$ было установлено в п. 2: $B \to s_B^{(A)}(x)$. Соответствие между функциями $\{0, 1\}^A$ и элементами $\{0,1\}^{|A|}$ получим, сопоставляя функции $f(x)$ упорядоченный набор ее значений

$$\langle f(a_1), f(a_2), \ldots, f(a_{|A|}) \rangle \in \{0, 1\}^{|A|}, \text{ где } f(a_i) \in \{0, 1\}.$$

Например, при $A = \{0,1\}$ эти соответствия таковы:

$$2^A \leftrightarrow \{0, 1\}^A \leftrightarrow \{0, 1\}^{|A|},$$
$$\varnothing \leftrightarrow f_1 = 0 \leftrightarrow \langle 0, 0 \rangle,$$
$$\{0\} \leftrightarrow f_2 = \bar{x} \leftrightarrow \langle 1, 0 \rangle,$$
$$\{1\} \leftrightarrow f_3 = x \leftrightarrow \langle 0, 1 \rangle,$$
$$\{0, 1\} \leftrightarrow f_4 = 1 \leftrightarrow \langle 1, 1 \rangle.$$

(в) Рассмотрим более подробно взаимно однозначное соответствие между A^B и $A^{|B|}$, доказывающее (9). Пусть $B = \{b_1, b_2, ..., b_n\}$, и пусть все функции A^B перечислены в левой части таблицы 1.

Таблица 1

x	b_1	$b_2 \dots b_n$
$f_1(x)$	$f_1(b_1)$	$f_1(b_2) \dots f_1(b_n) \to \langle f_1(b_1), f_1(b_2), \dots, f_1(b_n) \rangle$
$f_2(x)$	$f_2(b_1)$	$f_2(b_2) \dots f_2(b_n) \to \langle f_2(b_1), f_2(b_2), \dots, f_2(b_n) \rangle$
...	..	. . . .
$f_i(x)$	$f_i(b_1)$	$f_i(b_2) \dots f_i(b_n) \to \langle f_i(b_1), f_i(b_2), \dots, f_i(b_n) \rangle$
...	..	. . . .
$f_m(x)$	$f_m(b_1)$	$f_m(b_2) \dots f_m(b_n) \to \langle f_m(b_1), f_m(b_2), \dots, f_m(b_n) \rangle$

Справа от каждой функции показан соответствующий ей элемент A^n, это просто набор значений этой функции, упорядоченный в согласии с избранным упорядочением множества B. Очевидно, что каждый набор из A^n соответствует некоторой функции A^B, а различным функциям соответствуют различные наборы. Поэтому $|A^B| = |A|^{|B|}$.

(г) Число размещений элементов $B = \{b_1, b_2, ..., b_n\}$ по k ящикам $A = \{a_1, \dots, a_k\}$ равно k^n. Это следует из того, что каждому размещению можно сопоставить функцию $f \in A^B$ такую, что $f(x)$ есть номер ящика, в который помещен x. Любая функция из A^B задает некоторое размещение, а различным размещениям соответствуют различные функции.

(д) В примере (а) мы воспользовались рекурсией для определения множества A^n и его комбинаторной функции $|A^n|$:

$$A^1 = A, \quad A^n = A \times A^{n-1};$$

$$|A^1| = |A|, \quad |A^n| = |A| \cdot |A^{n-1}|.$$

Такой вид рекурсии, когда общий член параметрического семейства определяется схемой, включающей члены этого же семейства с меньшими значениями параметра, называют *рекуррентностью* (возвратом). Именно рекуррентность часто является итогом комбинаторного анализа множеств, так как последующее исследование — получение явного выражения комбинаторной функции или оценок для этой функции — перемещается в область функционального анализа. Рассмотрим, в качестве последней иллюстрации, множества $P_{A,k}$ всех k-элементных *перестановок* элементов из $A = \{1, 2, ..., n\}$, т. е. $P_{A,k}$ есть подмножество A^k, состоящее из всех наборов $<x_1, ..., x_k>$, у которых все элементы $x_1, ..., x_k$ различны. Пусть $A_i \Rightarrow A \setminus \{i\}$. Легко понять, что $P_{A,1} = A$ и

$$P_{A,k} = \bigcup_{i=1}^{n} \{i\} \times P_{A_i,k-1} \text{ для } k = 2, 3, \ldots, n.$$

Учитывая, что $|P_{A,k}| = |P_{B,k}|$ при $|A| = |B|$, имеем

$$P_{n,k} \Rightarrow |P_{A,k}|, \quad P_{n,1} = n, \quad P_{n,k} = n \cdot P_{n-1,k-1}.$$

Полученная простая рекуррентность дает

$$P_{n,n} = n! \Rightarrow n \cdot (n-1) \ldots 2 \cdot 1 \ (0! \Rightarrow 1)$$

— одну из основных комбинаторных функций наряду с k^n, и для произвольного $k = \overline{1, n}$

$$P_{n,k} = n(n-1) \ldots (n-k+1) = \frac{n!}{(n-k)!}.$$

10.2. Отношения порядка и нумерации

1. Пусть

$$A^+ \Rightarrow A \cup A^2 \cup \ldots = \bigcup_{i=1}^{\infty} A^i.$$

Отношениями на множестве A в широком смысле называются произвольные подмножества A^+, т. е. элементы $2^{(A+)}$ n-местными отношениями на множестве A называются произвольные подмножества A^n. Если представить элементы множества A *вершинами* (точками), а тот факт, что пара $<a_1, a_2>$ принадлежит отношению $\rho \subseteq A^2$, изобразить ориентированным отрезком (стрелкой), *ребром,* то такое представление ρ называется *графом отношения* ρ: $G_\rho = <A, \rho>$. При удачном расположении вершин и ребер графа G_ρ иногда может быть достигнуто хорошее понимание структуры отношения ρ и его свойств. Аналогичным образом для произвольного отношения ρ вводится понятие *гиперграфа* $G_\rho=<A, \rho>$, хотя с увеличением местности отношений ценность этого понятия становится все более проблематичной.

2. Способы задания отношений естественно имеют много общего со способами задания произвольных множеств, но есть и специфические моменты. Остановимся на двух приемах упрощения задания отношений и, в частности, возможности задания бесконечных отношений конечными средствами.

Пусть R — некоторый класс отношений, F — множество всех отиошний из R, включающих ρ, содержащее *наименьший элемент* (пересечение всех отношении из F), который обозначим $\rho' = (\rho)_n$. Тогда говорят, что ρ *семантически определяет* ρ' *в классе R.* Например, если *R* состоит из одного отношения ρ, то *(*$\varnothing$*)n* = ρ. Пусть {ρ_1, ρ_2, ρ_3, ρ_4, ρ_5} — отношения на *A* = {1, 2, 3}, графы которых изображены на рис. 1

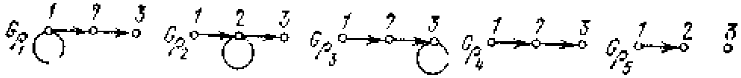


Рис. 1.

Если *R* = {ρ_1, ρ_2, ρ_3, ρ_4}, а $R_1 = R \cup \{\varnothing\}$, то имеем

$$(\varnothing)_R = \rho_4, \quad (\varnothing)_{R1} = \varnothing, \quad (\rho_5)_{R1} = \rho_4.$$

Пусть *s* — некоторое формальное правило построения по произвольному запасу наборов *X* некоторой совокупности наборов *s(X).* Через *s** обозначим итерированное правило

$$s: \quad s^0(X) \Rightarrow X, \quad s^{i+1}(X) \Rightarrow s(s^i(X))$$

и

$$s^*(X) \Rightarrow \bigcup_{i=0}^{\infty} s^i(X).$$

Тогда отношение ρ *синтаксически определяет* отношение ρ' *относительно правила s,* если ρ' = *s*(ρ), и относительно *s**, если *p*' = *s**(ρ).

3. Класс отношений, рассматриваемый в остальной части этого параграфа, — отношения порядка: *линейный порядок* и *частичный порядок.* Определения связаны с некоторыми свойствами бинарных отношений. Отношение ρ на *A* называется *рефлексивным,* если для любого $a \in A$ имеет место $<a, a> \in \rho$ и *антирефлексивным,* если для любого $a \in A$ имеет место $<a, a> \notin \rho$; *симметричным,* если для любых *a, b* ∈ *A* из $<a, b> \in \rho$ следует $<b, a> \in \rho$, и *антисимметричным,* если следует $<b, a> \notin \rho$; *транзитивным,* если для любых *a, b, c* ∈ *A* из $<a, b> \in \rho$ и $<b, c> \in \rho$ следует $<a, c> \in \rho$.

Термин «линейный порядок» отражает интуитивное представление о «полном» порядке в множестве. Полное упорядочение конечного множества *A* равносильно перенумерации его элементов, иными словами — для этого надо установить взаимно однозначное соответствие между *A* и начальным отрезком натурального ряда длины |*A*|, подобно тому, как это сделано в примере (а, п.6).

Мы будем рассматривать отношение строгого порядка «меньше», обозначаемое обычно символом «<» *(a < b).* Отношение нестрогого порядка определяется через строгий: «≤» ⇒ «<» ∪ « = ».

Очевидно, что отношение строгого порядка ρ на множестве *A* должно иметь следующие пять свойств:

антирефлексивность;

антисимметричность;

транзитивность;

сравнимость $\Rightarrow$ «для любых *a, b* $\in$ *A* либо $<a, b> \in \rho$, либо $<b, a> \in \rho$, либо $a = b$»;

ацикличность графа G_ρ.

Первые четыре из перечисленных условий принимаются за определение линейного порядка (пятое является их следствием).

4. Понятие частичного порядка вводится в связи с вопросом о возможности расширения отношения ρ до линейного порядка на множестве *A*.

Теорема 1. *Отношение ρ на A можно дополнить до линейного порядка на A в том и только в том случае, если граф* G_ρ *этого отношения ациклический.*

Доказательство. Ограничимся случаем конечного *A*.

Необходимость. Если $\rho \subseteq \rho_1$ и ρ_1 — линейный порядок, то G_{ρ_1} — ациклический, таков, следовательно, и G_ρ, который получается из G_{ρ_1} удалением некоторых ребер, в результате чего цикл возникнуть не может.

Достаточность. Скажем, что ρ' есть *элементарное расширение* ρ, если $\rho' = \rho \cup \{<a, b>\}$ (т. е. получается добавлением одной пары к ρ) и *элементарное транзитивное расширение,* если к тому же для некоторого $c \in A$ имеет место $<a, c> \in \rho$ и $<c, b> \in \rho$. Пусть $s(\rho)$ — транзитивное замыкание отношения ρ, т. е. результат применения элементарных транзитивных расширений до тех пор, пока это возможно.

Легко проверить следующие утверждения.

Лемма 1. *Элементарное транзитивное расширение ациклического отношения ациклично.*

Лемма 2. *Если ациклическое отношение ρ транзитивно, пара* $<a, b>$ *несравнима* (т. е. $a \neq b$, $<a, b> \notin \rho$ и $<b, a> \notin \rho$), *то* $\rho' = \rho \cup \{(a, b)\}$ *ациклично.*

Заметим, что при помощи элементарных расширений, удовлетворяющих леммам 1 и 2, невозможно расширение отношения, уже расширенного до линейного порядка. Линейный порядок всегда будет достигаться за конечное число элементарных расширений, так как общее число пар элементов из *A,* в силу конечности *A,* конечно. Теорема доказана.

Замечая, что транзитивное замыкание $s(\rho)$ определено однозначно, независимо от выбора порядка элементарных транзитивных расширений, получаем еще один факт.

Т е о р е м а 2. *Если* $\rho \in \rho_1$ *и* ρ_1 *— линейный порядок, то*

$$\rho \subseteq s(\rho) \subseteq \rho_1.$$

Таким образом, множество расширений до линейного порядка ациклического отношения ρ совпадает с множеством расширений до линейного порядка отношения $s(\rho)$. Это обстоятельство является причиной особого внимания к классу антирефлексивных, антисимметричных и транзитивных отношений. Бинарные отношения, обладающие этими тремя свойствами, называются отношениями *частичного порядка.*

Из теоремы 2 вытекает способ сокращенного задания отношений частичного порядка. Пусть ρ — отношение частичного порядка. Определим соответствующее ему отношение «покрытия»

$$\rho': \langle a, b\rangle \in \rho' \Rightarrow «\langle a, b\rangle \in \rho$$

и не существует $c \in A$ такого, что $\langle a, c\rangle \in \rho$ и $\langle c, b\rangle \in \rho$».

Тогда $s(\rho')=\rho=s(\rho)$. Следовательно, *отношение частичного порядка синтаксически определяется соответствующим отношением покрытия относительно правила транзитивного замыкания.* Легко понять, что им же оно определяется и семантически относительно класса отношений частичного порядка *R:* $\rho = (\rho')_R$.

5. Примеры, (а) Рассмотрим на $\{0, 1\}^n$ отношение

$$\langle x_1, \ldots, x_n\rangle \leqslant \langle y_1, \ldots, y_n\rangle \Rightarrow x_1 \leqslant y_1 \& \ldots \& x_n \leqslant y_n.$$

Легко проверить, что оно является отношением нестрогого частичного порядка, а отображение ν из примера (а, п.6) $\{0, 1\}^n$ на линейно упорядоченное множество натуральных чисел позволяет продолжить частичный порядок до линейного (называемого *лексикографическим)* следующим образом:

$$\langle x_1, \ldots, x_n\rangle < \langle y_1, \ldots, y_n\rangle \Rightarrow \nu(x_1, \ldots, x_n) < \nu(y_1, \ldots, y_n).$$

(б) Пусть $E_k=\{0,1,\ldots,k-1\}$. На $E_{m_1} \times E_{m_2} \times \ldots \times E_{m_n}$ определим, аналогично предыдущему, отношение частичного порядка:

$$\langle x_1, \ldots, x_n\rangle \leqslant \langle y_1, \ldots, y_n\rangle \Rightarrow x_1 \leqslant y_1 \& \ldots \& x_n \leqslant y_n.$$

Для $\langle x_1, \ldots, x_n\rangle \in E_{m_1} \times \ldots \times E_{m_n}$ положим

$$\nu(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i \left(\prod_{j=i+1}^{n} m_j \right), \quad \text{где} \quad \prod_{j=n+1}^{n} m_j \Rightarrow 1,$$

и снова получим расширение до линейного порядка. Это прямое обобщение примера (а): там ν представляло собой отображение, обратное двоичному представлению натуральных чисел, здесь — представлению натуральных чисел в системе счисления с переменным основанием (или произвольным постоянным основанием в случае $m_1 = m_2 = \ldots = m_n$).

6. В приложениях важное значение имеют вопросы реализации нумераций, устанавливающих линейный порядок в множествах. Встречаются такие нумерации в весьма разнообразных модификациях.

10.3. Отношения эквивалентности и разбиения

1. Для бинарного отношения ρ на множестве A через ρ(a) обозначается окрестность элемента a: $\rho(a) \Rightarrow \{b \mid \langle a, b\rangle \in \rho\}$. В графе G_ρ этого отношения множество ρ(a) есть множество всех вершин, соседних с a, т. е. таких, в которые из a ведут ребра. Множество всех окрестностей

$$A/\rho \Rightarrow \{\rho(a) \mid a \in A\}$$

называется *фактор-множеством A* по отношению ρ.

Система непустых множеств $A_1,\ldots,A_k$ называется *разбиением* множества A, если $A_i \cap A_j = \varnothing$ при $i \neq j$ и $A_1 \cup \ldots \cup A_k = A$. В этом случае $|A| = |A_1| + \ldots + |A_k|$, а если все классы разбиения равномощны, то $|A| = k \cdot |A_i|$ для любого $i = 1, k$.

Бинарное отношение называется *отношением эквивалентности,* если оно рефлексивно, симметрично и транзитивно.

Теорема 1. *Если ρ — отношение эквивалентности на A, то A/ρ есть разбиение.*

Доказательство. Ввиду рефлексивности ρ имеем $a \in \rho(a)$ для любого $a \in A$. Поэтому $A = \bigcup_{a\in A} \{a\} \subseteq \bigcup_{a\in A} \rho(a) \subseteq A$, откуда

$$\bigcup_{a\in A} \rho(a) = A.$$

Остается показать, что если $\rho(a) \cap \rho(b) \neq \varnothing$, то $\rho(a)=\rho(b)$. Пусть $c \in \rho(a) \cap \rho(b)$ и $d \in \rho(a)$. Тогда $\langle b, c\rangle \in \rho$ (так как $c \in \rho(b)$), $\langle c, a\rangle \in \rho$ (так как $\langle a, c\rangle \in \rho$ и ρ симметрично) и $\langle a, d\rangle \in \rho$ по условию. Ввиду транзитивности ρ заключаем, что

$<b,\ d> \in \rho$, т. е. $d \in \rho(b)$, следовательно, $\rho(a) \subseteq \rho(b)$. Аналогично проверяется, что $\rho(b) \subseteq \rho(a)$, откуда следует $\rho(a)=\rho(b)$ ч. т. д.

Заметим, что A/ρ может быть разбиением и тогда, когда ρ не является отношением эквивалентности. Например, если $A = \{a,\ b\},\ \rho = \{\langle a,\ b\rangle,\ \langle b,\ a\rangle\}$, то $A/\rho=\{\{a\},\{b\}\}$, но ρ антирефлексивно.

Тем не менее именно факторизация по отношениям эквивалентности, особенно в случае, когда классы эквивалентности равномощны, оказывается наиболее эффективным приемом анализа множеств.

Упрощение задания отношения эквивалентности достигается следующим образом. Для эквивалентности ρ граф G_ρ состоит из компонент связности $G^{(1)}, G^{(2)}, \ldots, G^{(k)}$, которые все являются полными графами на классах эквивалентности. Возьмем в каждой компоненте $G^{(i)}$ остов $T^{(i)}$ (связывающее дерево), и пусть ρ_i — отношение такое, что

$$G_{\rho_i} = T_i,\ \text{а}\ \rho' = \bigcup_{i=1}^{k} \rho_i.$$

Аналогично элементарному ранзитивному расширению определим элементарное рефлексивное и элементарное симметричное расширение отношений. Пусть $s(\rho)$— рефлексивное, симметричное и транзитивное замыкание отношения ρ.

Теорема 2. *Отношение эквивалентности* ρ *синтаксически определяется любым из своих остовных отношений* ρ' *относительно правила* s *рефлексивного, симметричного и транзитивного замыкания:* $s(\rho') = \rho$. *Им же оно определяется и семантически относительно класса отношений эквивалентности* R: $(\rho')_n = \rho$.

2. С разбиениями связано несколько комбинаторных функций. В первую очередь это *число сочетаний* из n элементов по $k: |C_n^k| = \binom{n}{k} \Rightarrow \frac{n!}{k!\,(n-k)!}$ — *биномиальные коэффициенты, число всех разбиений n-элементного множества:* $B(n)$ и некоторые другие.

(а) Сочетания из n элементов по k суть k-элементные подмножества n-элементного множества, скажем, $E_n = (0, 1,\ldots, n - 1)$. Рассмотрим на множестве $P_{n,k}$ k-элементных перестановок E_n отношение эквивалентности ρ: перестановки эквивалентны в том и только том случае, если они состоят из одних и тех же элементов. Тогда между множеством C_n и множеством $P_{n,k}/\rho$, очевидно, имеется взаимно однозначное соответствие: сочетанию A соответствует класс $P_{A,k}$

эквивалентности, состоящий из всех перестановок состава A. Например, при $n = 4$, $k = 2$ соответствие выглядит так:

$$\begin{array}{cc} C_4^2 & P_{4,2}/\rho \\ \{0,1\} \to & \{\langle 0,1\rangle, \langle 1,0\rangle\}, \\ \{0,2\} \to & \{\langle 0,2\rangle, \langle 2,0\rangle\}, \\ \{0,3\} \to & \{\langle 0,3\rangle, \langle 3,0\rangle\}, \\ \{1,2\} \to & \{\langle 1,2\rangle, \langle 2,1\rangle\}, \\ \{1,3\} \to & \{\langle 1,3\rangle, \langle 3,1\rangle\}, \\ \{2,3\} \to & \{\langle 2,3\rangle, \langle 3,2\rangle\}. \end{array}$$

Все классы равномощны: $|P_{A,k}| = k!$, поэтому $|P_{n,k}| = k!\,|C_n^k|$, откуда $|C_n^k| = \binom{n}{k} = \frac{n!}{k!\,(n-k)!}$.

Термин «биномиальный коэффициент» связан с *биноминальной теоремой:*

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k},$$

которая имеет комбинаторную природу. Каждый член $x^k y^{n-k}$, появляющийся в правой части после раскрытия скобок, соответствует некоторому подмножеству k скобок из n:

$$(x+y)\ldots(x+y) \quad (n \text{ раз})$$

- тому, из которого выбран в произведение x. Поэтому после приведения подобных членов коэффициент при $x^k y^{n-k}$ должен быть равен $|C_n^k|$, что и констатирует теорема.

(б) Комбинаторная функция *сочетаний с повторениями,* т. е. неупорядоченных *выборок* из *n по k* элементов, в которые каждый элемент может входить любое число раз и в которых объем выборки равен суммарному числу вхождений в нее всех элементов (число вхождений элемента может быть и нуль), выражается также биномиальным коэффициентом: $|D_n^k| = \binom{n+k-1}{n-1}$.

Действительно, между сочетаниями с повторениями и решениями уравнения $x_0 + x_1 + \ldots + x_{n-1} = k$ в неотрицательных целых числах есть очевидное взаимно однозначное соответствие: решению $\langle x_0, x_1, \ldots, x_{n-1}\rangle$ соответствует выборка, в которую 0 входит x_0 раз, 1 — x_1 раз и т. д. Решения этого уравнения находятся во взаимно однозначном соответствии с решениями $y_0 + y_1 + \ldots + y_{n-1} = k + n$ в положительных целых числах:

$$\langle x_0, x_1, \ldots, x_{n-1}\rangle \to \langle x_0 + 1, x_1 + 1, \ldots, x_{n-1} + 1\rangle.$$

Последние в свою очередь находятся во взаимно однозначном соответствии с множеством всех двоичных последовательностей длины $n + k — 1$, включающих ровно $n — 1$ вхождений единицы, т. е. с множеством всех характеристических функций множества C_{n+k-1}^{n-1}:

$$\langle y_0, y_1, \ldots, y_{n-1}\rangle \to 0^{y_0-1}10^{y_1-1}1\ldots 10^{y_{n-1}-1}.$$

Прослеженная цепочка индуцирует взаимно однозначное соответствие между D_n^k и C_{n+k-1}^{n-1}, откуда $|D_n^k| = |C_{n+k-1}^{n-1}|$, ч. т. д.

(в) Полиномиальная теорема.

$$(x_1 + \ldots + x_k)^n = \sum_{n_1+\cdots+n_k=n} \binom{n}{n_1 \ldots n_k} x_1^{n_1} \ldots x_k^{n_k},$$

где сумма в правой части берется по всем решениям уравнения $n_1 + \ldots + n_k = n$ *в неотрицательных целых, а* $\binom{n}{n_1 \ldots n_k} \Rightarrow \frac{n!}{n_1! \ldots n_k!}$ *— полиномиальные коэффициенты.*

Аналогично примеру (а) можно показать, что

$$\binom{n}{n_1 \ldots n_k}$$

является комбинаторной функцией множествавсех размещений n-элементного множества по k ящикам таких, что в первый ящик попадает n_1 элементов, во второй — n_2 элементов и т. д.

(г) Для комбинаторной функции $B(n)$ — числа разбиений n-элементного множества — нет простого выражения в элементарных функциях. Легко непосредственно определить несколько первых значений: $B(1) = 1$, $B(2) = 2$, $B(3) = 5$, $B(4) = 15$, $B(5) = 52$,... По определению полагают $B(0) \Rightarrow 1$. Для вычисления $B(n)$ можно использовать рекуррентность

$$B(n+1) = \sum_{i=0}^{n} \binom{n}{i} B(n-i). \qquad (11)$$

Соотношение (11) получается так. Пусть $\mathfrak{B}_{n+1}$ — множество всех разбиений E_{n+1} и $n \in A \subseteq E_n$. Через $\mathfrak{B}_{n+1}^A$ обозначим подмножество тех разбиений, у которых A есть класс разбиения, содержащий элемент n. Если $|A|=k$, то класс $\mathfrak{B}_{n+1}^A$ определяется выборкой (сочетанием!) из E_n по $k—1$ элементов, так как k-м элементом A является всегда n. Легко понять, что $\mathfrak{B}_{E_n}^A$ находится во взаимно однозначном соответствии с $\mathfrak{B}_{E_n \setminus A}$, а различные классы $\mathfrak{B}_{n+1}^A$ составляют в совокупности $\mathfrak{B}_{n+1}$:

$$\mathfrak{B}_{n+1} = \bigcup_{k=0}^{n} \bigcup_{A \in C_n^k} \mathfrak{B}_{n+1}^A, \tag{12}$$

откуда

$$|\mathfrak{B}_{n+1}| = \sum_{k=0}^{n} \sum_{A \in C_n^k} |\mathfrak{B}_{n+1}^A| = \sum_{k=0}^{n} \binom{n}{k} |\mathfrak{B}_{E_n \setminus A}|$$

и (11) доказано.

(д) Сходными рассуждениями получаются рекуррентные соотношения для комбинаторной функции *S(n, k)* — числа разбиений E_n ровно на *k* непустых классов:

$$S(n+1, k) = \sum_{j=k-1}^{n} \binom{n}{j} S(j, k-1), \tag{13}$$

$$S(n+1, k) = S(n, k-1) + kS(n, k). \tag{14}$$

Начальные значения для вычисления: $S(n, n) = S(n, 1) = 1$ при любом $n \geq 1$.

Для *S(n, k)* существует и явное выражение через элементарные функции:

$$S(n, k) = \frac{1}{k!} \sum_{j=0}^{k} (-1)^j \binom{k}{j} (k-j)^n. \tag{15}$$

Это можно доказать, например, с помощью (14) индукцией по *n*.

3. С отношениями эквивалентности связана важная идея производящих функций. Это понятие помогает при решении вопросов перечисления элементов как конечных, так и бесконечных множеств. Пусть каждому элементу множества *А* сопоставлены количественные признаки $P_1, \ldots, P_k$, т. е. имеется функция *f*, которая каждому $a \in A$ ставит в соответствие набор натуральных чисел $f(a) = \langle P_1(a), \ldots, P_k(a) \rangle$ (у людей такими признаками могут быть, например, возраст, рост, вес и т. п.). Каждому признаку P_i поставим в соответствие переменную x_i $(i = \overline{1, k})$, элементу $a \in A$ произведение

$$x_1^{P_1(a)} \ldots x_k^{P_k(a)},$$

а всему множеству *А* — функцию

$$F_A(x_1, \ldots, x_k) = \sum_{a \in A} x_1^{P_1(a)} \ldots x_k^{P_k(a)}. \tag{16}$$

Функция (16) называется *производящей функцией* множества *А* по признакам $P_1, \ldots, P_k$. Переменные и операции, фигурирующие в (16), формальные, т. е. символы, имена, которые надо интерпретировать, придать им определенные значения, смысл. Лишь после интерпретации производящая функция приобретает конкретное содержание, становится рабочим инструментом. Данная символика объясняется тем, что в наиболее распространенной интерпретации x_i — действительные или комплексные переменные, сложение, умножение и возведение в степень — элементарные функции анализа. Термин «производящая функция» отражает способ использования этого понятия — компактное задание информации о множестве с учетом того, что в алгебрах функций часто можно указать нетривиальные тождественные преобразования формул. Потому к интерпретации предъявляется лишь требование, чтобы *представление функций формулами вида* $\sum\Pi$ — *«сумма произведений», как в* (16),— *было единственным* с точностью до перестановок сомножителей в произведениях и слагаемых в сумме. Впрочем, изредка оказываются полезными и некоммутативные производящие функции.

Иногда производящая функция F_A перечисляет все множество A — если f устанавливает взаимно однозначное соответствие между A и $\{f(a) \mid a \in A\}$ (т. е. если признаки однозначно идентифицируют элемент из A, являются как бы его паспортом). Чаще F_A перечисляет A/ρ, где

$$\langle a_1, a_2\rangle \in \rho \Rightarrow \langle P_1(a_1), \ldots, P_k(a_1)\rangle = \\ = \langle P_1(a_2), \ldots, P_k(a_2)\rangle \qquad (17)$$

есть отношение эквивалентности на A. В последнем случае эквивалентным элементам A соответствуют одинаковые слагаемые в (16), поэтому после приведения подобных членов в (16) производящая функция может сохранить информацию о мощности классов эквивалентности (17) в виде коэффициентов при слагаемых.

Обратимся к примерам производящих функций.

(а) Производящая функция подмножеств n-элементного множества A по признаку $P(B) \Rightarrow |B|$, как следует из биномиальной теоремы в п. 2(а), есть

$$(1 + x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k.$$

Производящая функция сочетаний с повторениями по тому же признаку, согласно п. 2(б), есть

$$F_n(x) = \sum_{k=0}^{\infty} \binom{n+k-1}{n-1} x^k.$$

Используя рекуррентность

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1},$$

легко получить соотношение $F_n(x) = xF_n(x) + F_{n-1}(x)$.

откуда $F_n(x) = \frac{F_{n-1}(x)}{(1-x)}$ и ввиду $F_1(x) = \sum_{k=0}^{\infty} \binom{k}{0} x^k = = \sum_{k=0}^{\infty} x^k = \frac{1}{(1-x)}$ находим

$$F_n(x) = \frac{1}{(1-x)^n}.$$

Этот же вывод легко получить и прямыми комбинаторными рассуждениями, наподобие п. 2(*а*).

(б) Эквивалентные преобразования с целью упрощения формульного задания комбинаторных функций во многих случаях используют производящие функции.

Так, из примеров *(а)* получаем

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n, \quad \sum_{k=0}^{\infty} \binom{n+k}{k} 2^{-k} = \frac{1}{(1-1/2)^{n+1}} = 2^{n+1}.$$

Аналогично, из полиномиальной теоремы п. 2(в), полагая $x_1 = \ldots = x_k = 1$, получаем тождество

$$\sum_{n_1+\ldots+n_k=n} \binom{n}{n_1 \ldots n_k} = k^n.$$

(в) Производящие функции часто оказываются полезными для эквивалентных преобразований рекуррентных схем задания комбинаторных функций к формульному представлению через элементарные функции. Для числа разбиений *B(n),* используя рекуррентность (11), такое представление можно получить в виде бесконечного ряда.

Пусть

$$B(x) = \sum_{n=0}^{\infty} \frac{B(n)}{n!} x^n$$

— производящая функция последовательности $\{B(n)/n!\}_{n=0}^{\infty}$. Учитывая, что $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$, получаем

$$B'(x) = \sum_{n=1}^{\infty} \frac{B(n)}{(n-1)!} x^{n-1} = \sum_{n=0}^{\infty} \frac{B(n+1)}{n!} x^n =$$

$$= \sum_{n=0}^{\infty} \left[\sum_{k=0}^{n} \binom{n}{k} B(n-k) \right] \frac{x^n}{n!} =$$

$$= \sum_{n=0}^{\infty} \left[\sum_{k=0}^{n} \frac{x^k}{k!} \cdot \frac{B(n-k)}{(n-k)} x^{n-k} \right] = e^x B(x).$$

Таким образом, для $B(x)$ получено дифференциальное уравнение $B'(x)=e^x \bullet B(x)$ с начальным условием $B(0)=1$. Интегрируя, получаем

$$B(x)= e^{(e^{x-1})} = \frac{1}{e} \bullet e^{(e^x)} \qquad (18)$$

Разложим теперь (18) в ряд по степеням x:

$$B(x) = \frac{1}{e} \sum_{k=0}^{\infty} \frac{e^{kx}}{k!} =$$

$$= \frac{1}{e} \sum_{k=0}^{\infty} \frac{1}{k!} \sum_{n=0}^{\infty} \frac{k^n x^n}{n!} = \frac{1}{e} \sum_{n=0}^{\infty} \left(\sum_{k=0}^{\infty} \frac{k^n}{k!} \right) \frac{x^n}{n!}.$$

Отсюда получаем искомое выражение для числа разбиений:

$$B(n) = \frac{1}{e} \sum_{k=0}^{\infty} \frac{k^n}{k!}. \qquad (19)$$

Представление о порядке роста $B(n)$ дают следующие оценки.

Теорема 1. $B(n) \leq n!$

Доказательство. Доказательство проведем индукцией по n. Для $n=0$ имеем $B(0) = 1 = 0!$ Пусть для $1, 2, ..., n$ утверждение справедливо. Тогда при $n \geq 2$ получаем

$$B(n+1) \leqslant \sum_{i=0}^{n} \binom{n}{i} (n-i)! = n! \sum_{i=0}^{n} \frac{1}{i!} < n!\, e < (n+1)!,$$

ч. т. д.

Теорема 2.

$$B(n) \geqslant n^{n(1-\varepsilon(n))},$$

где $\varepsilon(n) \to 0$ *при* $n \to \infty$.

Доказательство. Мы покажем, что для любого m найдется $n_0 = n_0(m)$ такое, что при $n > n_0$

$$B(n) > n^{n\left(1-\frac{1}{\log_m n}-\frac{\log_m \log_m n}{\log_m n}\right)}. \qquad (20)$$

Согласно (19) имеем, учитывая, что $k! < \frac{k^k}{e}$ при $k \geq 3$:

$$B(n) = e^{-1}\sum_{k=1}^{\infty}\frac{k^n}{k!} > e^{-1}\sum_{k=3}^{\infty}\frac{k^n}{k!} > \sum_{k=3}^{\infty}k^{n-k}.$$

Пусть $f(x) = x^{n-x}$. При n достаточно большом имеем $B(n) > f\left(\frac{n}{\log_m n}\right)$, так как на отрезке $\left[\frac{n}{\log_m n}-1, \frac{n}{\log_m n}\right]$ функция $f(x)$ убывает (на этом отрезке

$$f'(x) = x^{n-x}\cdot\left(\frac{n}{x}-1-\ln x\right) < 0).$$

Следовательно,

$$B(n) > \left(\frac{n}{\log_m n}\right)^{n-\frac{n}{\log_m n}} > n^{n\left(1-\frac{1}{\log_m n}-\frac{\log_m \log_m n}{\log_m n}\right)},$$

ч. т. д.

Аналогично можно показать, что

$$S_k(x) = \sum_{n=0}^{\infty} S(n+k, k)\, x^n = \frac{1}{(1-x)(1-2x)\ldots(1-kx)}$$

и $S(n, k) \sim \frac{k^n}{k!}$ при фиксированном k и $n\to\infty$.

(г) Рассмотрим систему функций $f_1(n),\ldots, f_m(n)$, определенную условиями $f_i(n)=0$ при $n<0$, $f_i(0)=c_i$; и системой линейных рекуррентных соотношений

$$f_i(n) = \sum_{j=1}^{m}\sum_{k=1}^{b_j} a_{ijk} f_j(n-d_{ijk}) \qquad \left(i=\overline{1, m}, d_{ijk} > 0\right) \qquad (21)$$

для положительных n. Пусть $F_i(x) = \sum_{n=0}^{\infty} f_i(n)\, x^n$ — производящая функция последовательности $\{f_i(n)\}_{n=0}^{\infty}$ (i = 1, m). Умножая каждое равенство (21) на x^n и суммируя по n = 1, 2,..., получаем следующую систему линейных функциональных уравнений для производящих функций $F_i(x)$:

$$F_i(x) - c_i = \sum_{j=1}^{m} P_{ij}(x)\, F_j(x) \qquad \left(i=\overline{1, m}\right), \qquad (22)$$

где $P_{ij}(x) = \sum_{k=1}^{\delta_j} a_{ijk} x^{d_{ijk}}$. По правилу Крамера $F_i(x) = \frac{\delta_i(x)}{\delta(x)}$, где

$$\delta(x) = \det \begin{vmatrix} 1-P_{11} & -P_{12} & \dots & -P_{1m} \\ -P_{21} & 1-P_{22} & \dots & -P_{2m} \\ \dots & \dots & \dots & \dots \\ -P_{m1} & -P_{m2} & \dots & 1-P_{mm} \end{vmatrix},$$

а $\delta_i(x)$ получается подстановкой в матрицу, определяющую $\delta(x)$, вместо i-го столбца начальных условий $\|c_j\|$.

Пусть $d_i(x)$=НОД($\delta_i(x)$, $\delta(x)$), нормированный условием $d_i(0)=\delta(0)=1$ (находится с помощью алгоритма Евклида). Дробно-рациональную функцию $F_i(x)$ представим в виде несократимой дроби $F_i(x)=P_i(x)/\Delta_i(x)$, где $P_i(x)=\delta_i(x)/d_i(x)$, $\Delta_i(x) = \delta(x)/d_i(x)$. Пусть $x_1, \dots, x_s$ — корни многочлена $\Delta_i(x)$ кратностей $k_1, \dots, k_s$ соответственно. Тогда

$$F_i(x) = \frac{c_i' P_i(x)}{\left(1-\frac{x}{x_1}\right)^{k_1} \dots \left(1-\frac{x}{x_s}\right)^{k_s}}$$

и после разложения на простейшие дроби

$$\frac{1}{\left(1-\frac{x}{x_1}\right)^{k_1} \dots \left(1-\frac{x}{x_s}\right)^{k_s}} = \sum_{j=1}^{s} \sum_{l=1}^{k_j} A_{jl} \left(1-\frac{x}{x_j}\right)^{-l},$$

учитывая разложение $\frac{1}{(1-x)^n}$ из примера (а), получаем

$$F_i(x) = c_i' P_i(x) \sum_{j=1}^{s} \sum_{l=1}^{k_j} A_{jl} \sum_{t=0}^{\infty} \binom{l+t-1}{l-1} \frac{x^t}{x_j^t} =$$

$$= c_i' P_i(x) \sum_{t=0}^{\infty} \left(\sum_{j=1}^{s} q_j(t) x_j^{-t} \right) x^t,$$

Где $q_j(t) \Rightarrow \sum_{l=1}^{kj} \binom{l+t-1}{l-1} A_{jl}$ — полином от t степени не выше k_j—1.

Если

$$P_i(x) = \sum_{r=0}^{N} \alpha_r \cdot x^r,$$

то для $f_i(n)$ находим

$$f_i(n) = c_i' \sum_{r=0}^{N} \alpha_r \cdot \sum_{j=1}^{s} q_j(n-r)\, x_j^{-(n-r)} = \sum_{j=1}^{s} \pi_j(n)\, x_j^{-n}, \tag{23}$$

где

$$\pi_j(n) \Rightarrow c_i' \sum_{r=0}^{N} \alpha_r \cdot q_j(n-r)\, x_j^r$$

— полиномы от n.

Если некоторые из x_j — комплексные числа, то (23) можно преобразовать к действительной форме $\sum_j \psi_j(n) |x_j|^{-n}$, где $\psi_j(n)$ — полиномы от n, $\cos n\varphi$, $\sin n\varphi$.

При практической реализации этого плана нахождения формульного представления функций $f_i(n)$ наибольшую трудность может вызвать поиск корней $\Delta_i(x)$. Рассмотрим пример: пусть

$$f_1(n) = (17/4) f_1(n-1) - (9/17) f_2(n-1),$$
$$f_2(n) = 9 f_1(n-1) - (4/17) f_2(n-2),$$
$$f_1(0) = f_2(0) = 1.$$

Находим

$$\delta = \det \begin{vmatrix} 1-(17/4)\,x & (9/17)\,x \\ -9x & 1+(4/17)\,x^2 \end{vmatrix} = 1 - (17/4)\,x + 5x^2 - x^3 = (1-2x)^2 (1 - x/4),$$

$$\delta_1 = \det \begin{vmatrix} 1 & (9/17)\,x \\ 1 & 1+(4/17)\,x^2 \end{vmatrix} = \frac{1}{17}(4x^2 - 9x + 17),$$

$$\delta_2 = \det \begin{vmatrix} 1-(17/4)\,x & 1 \\ -9x & 1 \end{vmatrix} = \frac{1}{4}(19x+4), \quad d_1(x) = d_2(x) = 1,$$

$$F_1(x) = \frac{16}{17\cdot 49}\cdot\frac{(x+3)(4x^2-9x+17)}{(1-2x)^2} + \frac{1}{17\cdot 49}\cdot\frac{4x^2-9x+17}{(1-x/4)},$$

$$F_2(x) = \frac{4}{49}\cdot\frac{(x+3)(19x+4)}{(1-2x)^2} + \frac{1}{4\cdot 49}\cdot\frac{(19x+4)}{(1-x/4)}.$$

Остается найти коэффициенты при x^n. После несложных преобразований окончательно находим

$$f_1(n) = \frac{1}{833}\left[(756n + 788)\cdot 2^n + 45\cdot 4^{-n}\right],$$

$$f_2(n) = \frac{1}{49}(189n + 29)\cdot 2^n + \frac{20}{49}\cdot 4^{-n},$$

откуда, в частности, $f_1(n) \sim \frac{108}{119}\cdot n\cdot 2^n$, $f_2(n) \sim \frac{27}{7}\cdot n\cdot 2^n$.

(д) Любая функция алгебры логики $f(x_{1,}...,x_n)$ является производящей функцией множества $N \subseteq \{0,\ 1\}^n$, если представление (16) интерпретировать как совершенную дизъюнктивную нормальную форму и $P_i(x_1,..., x_n) = x_i$, а $x^1 = x$ и $x^0 = \overline{x}$. При этом соответствие между элементами N_j и конъюнкциями (16) взаимно однозначное.

Отметим еще две интерпретации (16) в классе функций алгебры логики, для которых также выполняется требование единственности представления в форме $\sum\Pi$: *полиномы Жегалкина* и *функции Яблонского.* Для функций Яблонского правая часть (16) есть сокращенная ДНФ.

4. Стандартным приемом факторизации универса при анализе конечных систем подмножеств являются *диаграммы Венна.*

Пусть $S = \{A_1, ..., A_n\}$ — система множеств в универсе U. Тип элемента $x \in U$ относительно S определяется тем, каким множествам из S этот элемент принадлежит, а каким — нет. Положим $<a, b> \in \rho \Rightarrow$ «для любого $i = \overline{1,\ n}$: $a \in A_i \leftrightarrow b \in A_i$». Очевидно, что ρ есть отношение эквивалентности на U и U/ρ содержит не более 2^n классов эквивалентности ρ (число возможных типов элементов, так как тип элемента $x \in U$ характеризуется двоичным набором $<s_1(x), ..., s_n(x)>$, где $s_i(x) = 1$, если $x \in A_i$, и $s_i(x) = 0$ в противном случае).

Диаграмма Венна для системы n множеств представляет собой разбиение прямоугольника на 2^n клеток — по клетке для каждого типа элементов. На рис. 2 построены диаграммы Венна для n = 1, 2, 3.

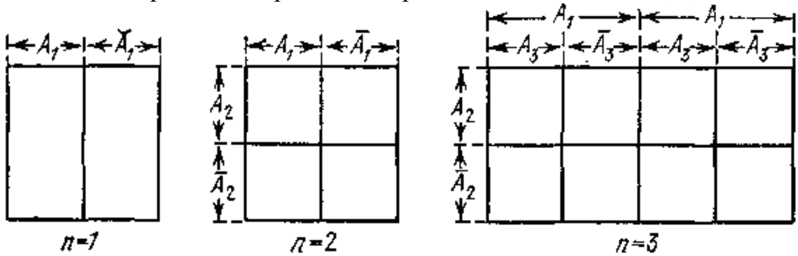


Рис. 2.

Из приведенных примеров легко понять, как построить диаграмму Венна для любого n: она получается из диаграммы для $(n$ — 1)-го множеств после того, как мы разделим пополам все вертикальные (или горизонтальные) полосы и отнесем к A_n все полосы с нечетными номерами, считая слева (сверху), а все полосы с четными номерами отнесем к $\overline{A}_n$. Тем самым каждая клетка предыдущей диаграммы разобьется на две части, одна из которых относится к A_n, а другая — к $\overline{A}_n$.

Полезно представить себе диаграмму Венна в виде ящика, в который можно разложить все элементы универса. Ящик состоит из 2^n ячеек, в каждую из них складываются однотипные элементы, элементы различных типов попадают в различные ячейки в соответствии с указателями. Таким образом, многие задачи логического анализа совокупностей множеств (равным образом — свойств) могут быть сведены к стандартной комбинаторной задаче о размещениях элементов данного множества по ящикам.

Пример. Дано множество $A \subseteq U=\{u_1 \ ..., \ u_n\}$. Сколько можно составить из элементов U пар множеств $<X_1, X_2>$ таких, что $X_1 \subseteq A \subseteq X_2$?

Рассмотрим диаграмму Венна для n = 3 (рис. 3).

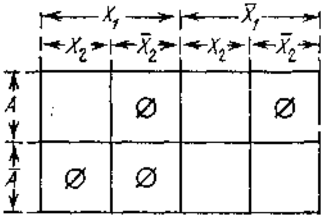


Рис. 3.

Так как условие $X \subseteq Y$ равносильно $X \cap \overline{Y} = \varnothing$, а между парами множеств $<X_1, X_2>$ и размещениями универса по ящикам диаграммы, такими что A размещается в верхних ящиках, а $\overline{A}$ — в нижних, есть взаимно однозначное соответствие, то парам, удовлетворяющим $X_1 \subseteq A \subseteq X_2$, соответствуют размещения, при которых ящики, отмеченные на диаграмме символом $\varnothing$, пусты. Следовательно, искомые пары перечисляются всевозможными размещениями элементов A по двум ящикам AX_1X_2 и $A\overline{X}_1X_2$ — независимо — элементов $\overline{A}$ по двум ящикам $\overline{A}\ \overline{X}_1X_2$ и $\overline{A}\ \overline{X}_1\ \overline{X}_2$. Число таких размещений равно $2^{|A|} \cdot 2^{|\overline{A}|} = 2^{|A|+|\overline{A}|} = 2^{|U|} = 2^n$.

10.4. Независимые множества в графах

1. В период формирования теории графов в XVIII — XIX веках одним из основных поставщиков задач о графах были игры и головоломки. Так и задача нахождения максимальных независимых множеств вершин в графах впервые возникла в шуточном, развлекательном варианте: в 1854 году Гаусс предложил читателям берлинского шахматного журнала найти все позиции, в которых на шахматной доске стоит максимально возможное число попарно не атакующих друг друга ферзей. Это число равно восьми, одна из позиций, решающих задачу Гаусса, изображена на рис. 4 (ферзь атакует любую клетку, расположенную с ним на одной линии — вертикали, горизонтали или диагонали).

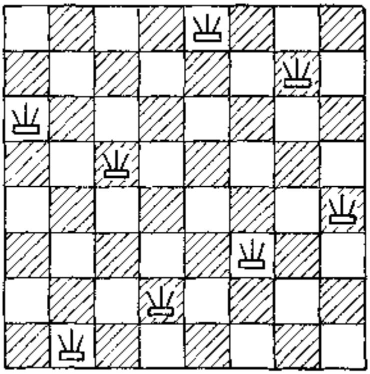


Рис. 4.

В XX веке содержание многих задач теории графов обогатилось, так как обнаружилась их связь с естественнонаучными моделями. Так случилось и с задачей о независимых множествах вершин. На ее прямое отношение к вопросам помехоустойчивого кодирования указал Шеннон в 1956 году.

2. В этом параграфе мы рассматриваем гиперграфы и графы только для симметричных отношений, т. е. ребра мы рассматриваем как неупорядоченные множества.

Пусть $G = <V, R>$ — гиперграф со счетным множеством вершин V и множеством ребер R (подмножеств V). Множество $W \subseteq V$ называется *независимым,* если оно не содержит в себе ни одного ребра.

Множество $W \subseteq V$ называется *опорой,* если для любого $r \in R$ имеет место $r \mid W \neq \varnothing$.

Независимое множество называется *тупиковым,* если всякое его расширение уже не является независимым множеством. Опора называется *тупиковой,* если всякое ее подмножество уже не является опорой. Через $H(G)$ обозначим класс всех независимых множеств G, $H'(G)$ — класс всех его опор, $H_T(G)$ и $H'_T(G)$ — соответственно классы всех тупиковых независимых множеств и тупиковых опор G.

Понятия независимого множества и опоры гиперграфа двойственны: непосредственно из определений вытекает следующая теорема.

Теорема 1. *Множество $W \in H(G)$ тогда и только тогда, когда $\overline{W} = V/W \in H'(G)$. Множество $W \in H_T(G)$ тогда и только тогда, когда $\overline{W} = V \setminus W \in H'_\tau(G)$.*

Через

$$\alpha(G) = \max_{W \in H(G)} |W|$$

обозначают *число независимости* гиперграфа G,

$$G,\ \alpha'(G) = \min_{W \in H'(G)} |W|$$

— *число опоры* G. Для конечных гиперграфов из теоремы 1 следует, что

$$\alpha(G) = |V| - \alpha'(G). \tag{24}$$

У бесконечного гиперграфа число независимости может быть конечным, бесконечным или не существовать.

Пример, когда для графа G числа независимости не существует, извлекается из следующей теоремы Диксона.

Пусть в графе $G_k = \langle V,\ R\rangle$: $V \Rightarrow N^k$, где $N = \{0, 1, 2,...\}$ есть множество всех натуральных чисел с нулем, и для

$X = \langle x_1, \ldots, x_k\rangle$ и $Y = \langle y_1, \ldots, y_k\rangle$ в N^k $(X \neq Y)$

$\langle X,\ Y\rangle \in R \Rightarrow$ «X и Y покоординатно сравнимы, т. е.

$$x_1 \leqslant y_1, \ldots, x_k \leqslant y_k \text{ или } y_1 \leqslant x_1, \ldots, y_k \leqslant x_k».$$

Теорема 2. *В N^k не существует бесконечного множества попарно несравнимых векторов.*

Доказательство. Индукция по k. При $k = 1$ утверждение очевидно: любые два числа сравнимы. Пусть оно справедливо при всех $k < m$. Рассмотрим $H \in H(G_m)$. Пусть $H_{ij} \Rightarrow \{X|X \in H,\ x_i = j\}$ — множество всех векторов из H, у которых значение i-й координаты фиксировано и равно j. По предположению индукции это множество конечно, пусть

$h_{ij} = |H_{ij}| < \infty$,

Положим $m_i \Rightarrow \min\{x_i \mid X \in H\}$ и $H' \Rightarrow \bigcup_{i=1}^{m} H_{im_i}$.

Множество H' непусто и конечно: $|H'| \leqslant \sum_{i=1}^{m} h_{im_i}$. Следовательно, для каждого $i=1,m$ существует $M_i \Rightarrow \max\{x_i | X \in H'\} < \infty$ и множество $H'' \Rightarrow \bigcup_{i=i}^{m} \bigcup_{j=m_i}^{M_i} H_{ij}$ тоже конечно:

$$|H''| \leqslant \sum_{i=1}^{m} \sum_{j=m_i}^{M_i} h_{ij}.$$

Остается показать, что $H=H''$ и, следовательно, конечно. По построению $H'' \subseteq H$ и для любого $X \in H''$ имеет место $X \leqslant M \Rightarrow \langle M_1, \ldots, M_m \rangle$. Но если $Y \in H$ и $Y \notin H''$, то, по построению H'', для любого i имеем $y_i > M_i$, т. е. $Y > M$. Тогда для любого $X \in H''$ было бы $X \le M < Y$, т. е. X и Y были бы сравнимы в H, что невозможно. Таким образом, $H = H''$ и теорема доказана.

Следствие. $\alpha(G_1) = 1$, *и при* $k \geq 2$ $\alpha(G_k)$ *не существует.*

Действительно, для любого t = 1, 2, .. ., в N^k можно указать независимое множество

$$\{\langle i,\ t-1-i,\ 0, \ldots, 0\rangle \mid i = \overline{0, t-1}\}$$

мощности t, а бесконечного независимого множества не существует в силу теоремы Диксона.

3. Алгоритмическая задача нахождения $\alpha(G)$ и какого-либо из независимых множеств максимальной мощности или всего класса $H_M(G)$ независимых множеств максимальной мощности для конечных графов, подобно задаче о минимизации ДНФ и многим другим экстремальным задачам, относится к классу так называемых *универсальных переборных проблем.* К такому ярлыку есть весьма веские основания. Считается, что основная часть любого алгоритма, решающего универсальную переборную проблему, состоит в переборе области решений и выборе оптимального на основе сравнения их параметров. При таких обстоятельствах особенно актуальным становится вопрос об организации перебора.

Пусть $V = \{v_1, \ldots, v_n)$. В нашем случае переборный путь состоит в перечислении всех 2^n подмножеств V, отборе из них $H(G)$ и, после сравнения попавших в $H(G)$, нахождении $H_M(G)$ и, следовательно, $\alpha(G)$. Объем перебора растет при этом очень быстро — экспоненциально — с ростом числа вершин гиперграфов. Существуют различные эвристические приемы сокращения перебора в задачах дискретной оптимизации, фигурирующие обычно под шифром метода ветвей и границ. Мы ограничимся здесь изложением систематического подхода.

Первое очевидное соображение состоит в том, что максимальные независимые множества содержатся среди тупиковых: $H_M \subseteq H_T$. Как следует из теоремы 1.2, между H_T и H'_T имеется простое взаимно однозначное соответствие. Поэтому для нахождения максимальных независимых множеств достаточно перебрать H_T или H'_T. Кроме этого, для H'_T мы можем написать функцию Яблонского в качестве производящей функции, перечисляющей характеристические функции H'_T по признакам

$$P_i(x_1, \ldots, x_n) = \begin{cases} 1, & \text{если } x_i = 1, \\ 0 & \text{в противном случае} \end{cases}$$

при интерпретации $x^0 \Rightarrow 1$.

Теорема 1. *Функция Яблонского*

$$f_G(x_1, \ldots, x_n) = \underset{r \in R}{\&} \left(\bigvee_{v_i \in r} x_i \right) \qquad (25)$$

является производящей для H'_T.

Доказательство. Легко проверяются следующие факты:

(а) $f_G(x_1,..,. x_n) = 1$ в том и только том случае, если вектор $<x_1,..., x_n>$ есть набор значений характеристической фукции опоры $\{v_i \mid x_i = 1)$.

(б) Функция алгебры логики f_0 монотонна, следовательно, ее сокращенная ДНФ не содержит отрицаний.

(в) $x_{i_1} \& \ldots \& x_{i_k}$ есть конъюнкция минимального ранга в том и только том случае, если $[v_{i_1}, \ldots, v_{i_k}]$ — тупиковая опора. Отсюда следует утверждение теоремы. После раскрытия скобок в (25) получается выражение, не содержащее отрицаний. Поэтому сокращенная ДНФ получается из него применением только правил поглощения:

$$A \vee AB = A, \quad 1 \cdot A = A \cdot 1 = A, \quad A \cdot A = A \vee A = A.$$

При раскрытии скобок полезно также иметь в виду тождество

$$\underset{i=1}{\overset{n}{\&}} (x \vee x_i) = x \vee x_1 x_2 \ldots x_n.$$

Пример. Для графа G, изображенного на рис. 5, имеем

$$f_G(x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6,\ x_7,\ x_8) =$$
$$= (x_1 \vee x_2)(x_1 \vee x_4)(x_2 \vee x_3)(x_2 \vee x_7)(x_3 \vee x_4)(x_4 \vee x_5)$$
$$(x_4 \vee x_6)(x_5 \vee x_6)\&(x_5 \vee x_8)(x_6 \vee x_8)(x_6 \vee x_7) =$$
$$= x_2x_4x_5x_6 \vee x_2x_4x_6x_8 \vee x_2x_4x_5x_7x_8 \vee x_1x_2x_3x_5x_6 \vee$$
$$\vee x_1x_3x_5x_6x_7 \vee x_1x_3x_4x_5x_7x_8 \vee x_1x_3x_4x_6x_7x_8.$$

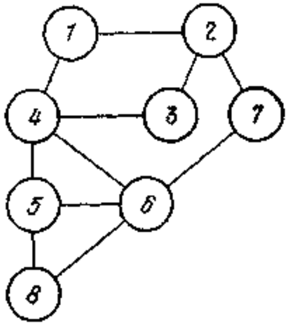


Рис. 5.

Таким образом, $\alpha(G) = 4$ и $|H_T| = |H'_T| = 7$. Множества H_T и H'_T перечислены в таблице 2.

Таблица 2

H_T	1, 3, 7, 8	1, 3, 5, 7,	1, 3, 6	4, 7, 8
H'_T	2, 4, 5, 6	2, 4, 6, 8	2, 4, 5, 7, 8	1, 2, 3, 5, 6
H_T	2, 4, 8	2, 6	2, 5	
H'_T	1, 3, 5, 6, 7	1, 3, 4, 5, 7, 8	1, 3, 4, 6, 7, 8	

Рассмотрим графы $G_n^{(2)} = \langle E_2^n, R\rangle$, т. е. подграфы G_n, порожденные подмножеством $E_n^2 = \{0,\ 1\}^n \subset N^n$.

Теорема 2.

$$\alpha\left(G_n^{(2)}\right) = \binom{n}{[n/2]}.$$

Доказательство. Пусть $H \in H\left(G_n^{(2)}\right)$ и $H = \bigcup_{i=0}^{n} H_i$, где H_i— множество всех векторов из H веса i (для $X = <x_1,..., x_n>$ вес $\|X\| = \sum_{i=1}^{n} x_i \Big)$, $h_i = |H_i|$.

Возьмем произвольный вектор X такой, что $\|X\| = k$. Перечислим все линейно упорядоченные цепочки, проходящие через X, у которых соседние элементы отличаются только в одной координате, а длина равна $n + 1$. От X вверх цепь можно продолжить $(n — k)!$ способами: добавляем единицу в любой из $(n — k)$ нулевых разрядов X, затем — в любой из оставшихся $(n — k — 1)$ нулевых и т. д. Аналогично, вниз такую цепь можно продолжить, независимо от верхнего участка, $k!$ способами. Поэтому общее число рассматриваемых цепей равно $k!(n — k)!$. Но различным элементам H соответствуют попарно непересекающиеся множества цепей. Так как всего имеется $n!$ цепей, имеем $\sum_{k=0}^{n} h_k k!\,(n-k)! \leqslant n!$, откуда

$$\sum_{k=0}^{n} h_k \binom{n}{k}^{-1} \leqslant 1. \qquad (26)$$

А так как $\binom{n}{k}$ достигает максимума при $k = [n/2]$ (целая часть $n/2$), умножая (26) на $\binom{n}{[n/2]}$, получаем оценку

$$|H| = \sum_{k=0}^{n} h_k \leqslant \sum_{k=0}^{n} h_k \binom{n}{k}^{-1} \binom{n}{[n/2]} \leqslant \binom{n}{[n/2]},$$

т. е.

$$\alpha\left(G_n^{(2)}\right) \leqslant \binom{n}{[n/2]}.$$

Равенство в утверждении теоремы следует из того, что множество всех векторов веса $[n/2]$ независимо и имеет мощность $\binom{n}{[n/2]}$. Теорема доказана.

Приведем усиление теоремы 2.

Теорема 3. $H_{\text{м}}\left(G_n^{(2)}\right)$ *при четном n состоит из единственного множества, упомянутого в доказательстве теоремы* 2, *а при нечетном n — из двух, вторым является множество всех векторов веса* $(n+1)/2$.

Из теоремы 2 следует оценка числа тупиковых независимых множеств произвольного гиперграфа, что то же самое —длины сокращенной ДНФ $f_G(x_i,..., x_n)$:

$$|H_{\tau}(G)| = l_c(f_G) \leqslant \binom{n}{[n/2]} \sim \frac{2^n}{\sqrt{2^{-1}\pi n}}. \qquad (27)$$

Если G — граф, то оценка (27) может быть усилена:

$$|H_{\tau}(G)| = l_c(f_G) \leqslant (\sqrt[3]{3})^n, \quad \sqrt[3]{3} \approx 1{,}442\ldots \qquad (28)$$

4. Рассмотрим еще один бесконечный граф $G=(A^+, R)$, где R — отношение префиксности между словами

$$A^+ = \bigcup_{i=1}^{\infty} A^i: \langle x,y\rangle \in R \Rightarrow$$

«существует z такое, что $x = yz$ или $y = xz$». Случай $|A|=1$ тривиален: любые два слова находятся в отношении префиксности, G есть полный граф и $\alpha(G) = 1$.

Будем далее предполагать, что $A=E_n$, $n\geq 2$. Тогда множество $\{1, 01,..., 0^i1,...\}$ независимо и бесконечно, т. е. $\alpha(G) =\infty$.

Независимые множества графа G называют *префиксными кодами,* тупиковые префиксные коды называют также полными. В этом пункте исследуются свойства префиксных кодов.

Для слова $x = x_1, \ldots x_k, \in A^k$ через $|x| = k$ будем обозначать длину этого слова, а для множества слов $W = \{w_1, w_2,.., w_i,...\}$ через $D(W)$ — спектр длин слов W, $D(W)=\{|w_1|,|w_2|,...,|w_i|,...)$, а $d(W)$ =max ($|w_1|$ i = 1, 2,....} при $|W|$ ∞.

Непосредственно из определений вытекает следующая теорема.

Теорема 1. *Если $W\in H(G)$, то*

(а) *если* α — *префикс по крайней мере одного из слов W, то* $W_\alpha \Rightarrow \{\beta|\alpha\beta\in W)$ $\in H(G)$;

(б) *если $w_i\in W$ и $A_1 \subseteq A$, то* $(W\backslash w_i) \cup wA_i \in H(G)$;

(в) *если* $\alpha A \subseteq W$, *то* $(W\backslash\alpha A) \cup\{\alpha\}\in H(G)$;

(г) $W\in H_T(G)$ *в том и только том случае, если для любого* $\alpha\in A^+$ *найдется* $w_i\in W$ *такое, что* $(\alpha,w_i) \in R$.

Непосредственным применением критерия тупиковости кода (см. (г) в теореме 1) получаются все утверждения следующей теоремы.

Теорема 2. *Пусть $W\in H_T(G)$. Тогда:*

(а) *если* α — *префикс по крайней мере одного из слов W, то* $W_\alpha\in H_T(G)$;

(б) если $W=0\bullet W_0 \cup ... \cup (n-1)W_{n-1}$, *то каждое множество* W_i $(i = 0, 1, \ldots,n - 1)$ *есть тупиковый префиксный код;*

(в) *единственный тупиковый префиксный код с наименьшим числом слов и одновременно наименьшей возможной суммой длин всех слов (равными n) есть A;*

(г) *если* α — *префикс по крайней мере одного из слов* $W(|W|< \infty)$ *и* $|\alpha| = d(W) - 1$, *то* $\alpha A \subseteq W$;

(д) *если* $w_i \in W$, *то* $W_{(i)} \Rightarrow (W \backslash \{w_i\}) \cup w_i A \in H_T(G)$;

(е) *если* $\alpha A \subseteq W$, *то* $W' \Rightarrow (W \backslash \alpha A) \cup \{\alpha\} \in (G)$.

Если $W \in = H_T(G)$ и конечен, то в утверждении (е) теоремы 2 имеем

$$|W'| = |W| - n + 1 < |W|,$$
$$\sum_{w \in W'} |w| < \sum_{w \in W} |w|.$$

С учетом (в) и (г) это позволяет во многих случаях доказывать утверждения о тупиковых префиксных кодах индукцией по $|W|$, $d(W)$, или по $\sum_{w \in W} |w|$.

Теорема 3. *Пусть* $W \in H_T(G)$ *и* $|W| = m < \infty$.

Тогда:

(а) $n - 1$ *является делителем* $m - 1$;

(б) $d(W) \leqslant \frac{m-1}{n-1}$.

Доказательство. (а) При $m = n$ это следует из (в) теоремы 2. Если при $|W| < m$ утверждение справедливо и $m>n$, то $d(W) > 1$ и согласно (г) $d(W) =\{d_i, ..., d_{m-n}, d, ..., d\}$. Тогда по (е) $\{d_1 ..., d_{m-n}, d(W)\}$ — спектр длин слов тупикового префиксного кода с числом слов, меньшим m. По предположению индукции $n - 1$ является делителем $m - n$, но $m - 1 =(m - n) + (n - 1)$ и $n - 1$ является делителем $m - 1$, ч. т. д.

(б) Индукция по $d(W)$. При $d(W) = 1$ это следует из (в) теоремы 2. Предположим утверждение справедливым при $d<d(W)$. Согласно (г), W можно представить в виде $W = W_1 \cup W_2 \bullet A$, где $W_2 \bullet A \subseteq A^{d(W)}$, $d(W_1)<d(W)$. По (е) $W_3 = W_1 \cup W_2 \in H_T$ и $d(W_3)<d(W)$. По предположению $d(W) - 1 \leqslant \frac{|W_1| + |W_2| - 1}{n-1}$, следовательно,

$$d(W) \leqslant \frac{|W_1| + |W_2| - 1 + n - 1}{n-1} \leqslant \frac{|W| - 1}{n-1},$$

так как $|W_1| + |W_2| + n - 1 \leq |W|$. Теорема доказана.

Теорема 4. *Если* $W \in H(G)$, *то* $D(W)$ *удовлетворяет неравенству Мак-Миллана*

$$\sum_{w \in W} n^{-|w|} \leqslant 1. \qquad (29)$$

Доказательство. Пусть $W = \{w_1, ..., w_k\}$, $|w_i|=d_i$ и $N>d(W)$. Тогда

$$w_1A^{N-d_1} \cup \ldots \cup w_kA^{N-d_k} \subseteq A^N,$$

причем $w_iA^{N-d_i} \cap w_jA^{N-d_j} = \varnothing$ при $i \neq j$.

Следовательно.

$$\left| \bigcup_{i=1}^{k} w_iA^{N-d_i} \right| = \sum_{i=1}^{k} \left| w_iA^{N-d_i} \right| = \sum_{i=1}^{k} n^{N-d_i} \leqslant \left| A^N \right| = n^N,$$

откуда следует (29). Теорема доказана, так как для бесконечного W утверждение следует из его справедливости для любого конечного подмножества.

Пусть $W \in H(G)$ и $W=W_1 \cup ... \cup W_i,...$, где $W_i=W \cap A^i$. Тогда спектр длин W можно задать в виде последовательности $D'(W) \Rightarrow \langle\sigma_1, \sigma_2, ..., \sigma_i ...\rangle$, где $\sigma_i = |W_i|$, а неравенство Мак-Миллана переписать в форме

$$\sum_{i=1}^{\infty} \sigma_i n^{-i} \leqslant 1. \qquad (30)$$

Теорема 5. *Если выполнено* (30), *то существует* $W \in H(G)$, *удовлетворяющий условию* $D'(W) = \langle\sigma_1, \sigma_2, ..., \sigma_i ...\rangle$.

Доказательство. Искомый префиксный код строится поэтапно: выбираем произвольно $W_1 \subseteq A$ так, что $|W_1| = \sigma_1$, и, после того как построены $W_1, W_2, ..., W_{k-1}$ на k-м этапе выбиваем произвольно $W_k \subseteq$ $\subseteq \Delta_k \Rightarrow A^k \setminus (W_1 \cdot A^{k-1} \cup \ldots \cup W_{k-1} \cdot A^1)$ так, что $|W_k|=\sigma_k$, и так до бесконечности. Корректность построения, очевидно, зависит от того, будет ли справедливо неравенство $|\Delta| \geq \sigma_k$ при любом k = 1, 2,... Но мы имеем

$$|\Delta_k| = |A^k| - \sum_{i=1}^{k-1} |W_i| \cdot |A^{k-i}| = n^k - \sum_{i=1}^{k-1} \sigma_i \cdot n^{k-i} =$$

$$= n^k \left(1 - \sum_{i=1}^{k-1} \sigma_i n^{-i}\right) \geqslant \sigma_k$$

(последнее неравенство следует из того, что

$$1 - \sum_{i=1}^{k} \sigma_i n^{-i} \geqslant 1 - \sum_{i=1}^{\infty} \sigma_i n^{-i} \geqslant 0,$$

Откуда $1 - \sum_{i=1}^{k-1} \sigma_i n^{-i} \geqslant \sigma_k n^{-k}$, ч. т. д.

Теорема 6. *Если* $W \in H(G)$ *и* $|W| < \infty$, *то существует* $W' \in H_T(G)$ *такой, что* $W \subseteq W'$ *и* $d(W) = d(W')$.

Доказательство. Пусть $W = W_1 \cup W_2$, где $W_2 \subseteq A^{d(W)}$, а $d(W_1)<d(W)$ (возможно, что $W_1 = \varnothing$.Тогда

$$W_2 \Leftarrow A^{d(W)} \setminus W_t A^+ \quad \text{и} \quad W' \Rightarrow W_t \cup (A^{d(W)} \setminus W_t A^+)$$

— искомый тупиковый префиксный код. Теорема доказана.

Следующая теорема устанавливает спектральную характеристику конечных множеств из $H_T(G)$.

Теорема 7. *Для того чтобы в $H_T(G)$ существовал W, удовлетворяющий условию $D(W)=\{d_1,...,d_m\}$, необходимо и достаточно выполнение*

$$\sum_{i=1}^{m} n^{-d_i} = 1. \tag{31}$$

Доказательство. *Необходимость*. Если $W' \in H_T(G)$, то по теореме 4 выполнено неравенство Мак-Миллана, которое мы возьмем в форме (20):

$\sum_{i=1}^{k} \sigma_i n^{-i} \leqslant 1$. Но если $\sum_{i=1}^{k} \sigma_i n^{-i} < 1$, то при построении $W=W_1 \cup ... \cup W_k$, как в доказательстве теоремы 5, получим $|\Delta_k| > \sigma_k$, т. е. $W_k \subseteq \Delta_k$ и $W_k \neq \Delta_k$.

Следовательно, $W' = W_1 \cup ... \cup W_{k-1} \cup \Delta_k$ есть независимое расширение W, что противоречит предположению о тупиковасти W. Поэтому выполняется (31).

***Достаточность*.** Из (31) по теореме 5 следует существование W, для которого $W \in H(G)$ и $D(W) = \{d_1, ..., d_m\}$. Но, согласно теореме 4, $W \in H_T(G)$, так как всякое его расширение нарушает необходимое условие (29). Теорема доказана.

Описание структуры множеств $W \in H(G)$ дается так называемой структурной функцией

$$f_w(z_0, ..., z_{n-1}) \Rightarrow \sum_{w \in W} z_0^{|w|_0} z_1^{|w|_1} ... z_{n-1}^{|w|_{n-1}}, \tag{32}$$

где $|w|_i \Rightarrow$ число вхождений i в слово w». Функция $f_w(z_0,..., z_{n-1})$ есть производящая функция, перечисляющая слова множества W по их составу, т. е. по признакам $P_i(w) \Rightarrow |w|_i$. В случае конечного W сумма (32) содержит конечное число слагаемых и ее называют структурным полиномом множества W. Описание структуры конечных тупиковых префиксных кодов дает

Теорема 8. *Полином $f(z_0,..., z_{n-1})$ является структурным полиномом некоторого тупикового префиксного кода в том и только том случае, если выполнены условия:*

(а) $f(0,..., 0) = 0$;

(б) *коэффициенты f — неотрицательные целые;*

(в) $f - 1 = (z_0 + ... + z_{n-1} - 1)g$ *для некоторого полинома* $g(z_0, ..., z_{n-1})$, *коэффициенты которого тоже неотрицательные целые, причем, если* $f = f_W$, *где* $W \in H_T(G)$, *то* $g = f_{\pi(W)} + 1$, где $\pi(W)$ — *множество всех префиксов слов* W.

Доказательство. ***Необходимость***. 1 Пусть $f=f_W$ и $W \in H_T(G)$. Выполнение (а) и (б) очевидно по смыслу производящей функции, (в) докажем индукцией по $|W|$. При $|W|=n$ по теореме 2 (в) $W=A$ и $f_W = z_0 + ... + z_{n-1}$, т. е. (в) выполнено с $g = 1 + f_{\pi(A)}$, так как $\pi(A) = \varnothing$ и $f_{\pi(A)} = \varnothing$. Пусть $|W| = k > n$ и для случаев, когда $|W| < k$, (в) проверено. По теореме 2 (г, е) для $\alpha \in \pi(W)$ длины $d(W) - 1$ имеем $\alpha A \subseteq W$ и $W' = (W/\alpha A) \cup \{\alpha\} \in H_T(G)$. Но $|W'| = k - n + 1 < k$ и по предположению индукции $f_{W'} = 1 + (z_0 + \ldots + z_{n-1} - 1) \cdot g_1$, где $g_1 = f_{\pi(W')} + 1$. Ввиду $f_W = f_{W'} - f_\alpha + f_{\alpha A}$ и $f_{\alpha A} = f_\alpha \cdot f_A$ имеем

$$f_W - 1 = (z_0 + \ldots + z_{n-1} - 1) \cdot g_1 + (z_0 + \ldots + z_{n-1} - 1) f_\alpha = (z_0 + \ldots + z_{n-1} - 1)(f_{\pi(W)} + 1),$$

так как $\pi(W) = \pi(W') \cup \{\alpha\}$ и $f_{\pi(W)} = f_{\pi(W')} + f_\alpha$, т. е. (в) справедливо и для W. Необходимость доказана.

Достаточность. Индукция по $t = \deg(f)$ (степени полинома f).

Если $t = 1$, то $g = \mathrm{const}$, $f = (z_0 + ... + z_{n-1}) \cdot g - (g - 1)$ и ввиду (а) $g = 1$, поэтому $f = (z_0 + ... + z_{n-1}) = f_A$ и утверждение проверено. Предположим, что оно справедливо для всех полиномов степени, меньшей t $(t>1)$. В силу (в) $f-1$ можно представить в виде

$$f - 1 = (z_0 + \ldots + z_{n-1} - 1)(h + \bar{g}),$$

где $g = h + \bar{g}$ и h — однородный полином степени $t - 1$, а $\deg(\bar{g}) < t-1$.

Рассмотрим полином $\bar{f} = f - (z_0 + \ldots + z_{n-1} - 1)h = 1 + (z_0 + \ldots + z_{n-1} - 1) \cdot \bar{g}$. Покажем, что для $\bar{f}$ (а) — (в) выполнены.

1) $\bar{f}\,(0,..., 0)=0$ очевидно.

2) Коэффициенты $\bar{f}$ — неотрицательные целые, так как отрицательные члены в $\bar{f}$ могли бы появиться только из слагаемого $-(z_0 + .. . + z_{n-1})h$ и их степень была бы равна t, тогда как $\deg(\bar{f}) = 1 + \deg(\bar{g}) \leq t - 1$.

3) $\bar{f} - 1 = f - 1 - (z_0 + \ldots + z_{n-1} - 1)h = (z_0 + \ldots + z_{n-1} - 1) \cdot \bar{g}$, где коэффициенты $\bar{g}$ — неотрицательные целые.

Так как deg ($\overline{f}$)<t, по предположению индукции существует $\overline{W}\in H_T(G)$ такой, что $\overline{f}=f_{\overline{W}}$. Пусть $\overline{W}=\{w_1, \ldots \ldots, w_N\}$, а $\overline{f}$ записан в виде суммы членов с единичными коэффициентами $\overline{f}=f_1+f_2+\ldots+f_N$ так, что $f_i=f_{w_i}$ и пусть $f_1+\ldots+f_j=h$, а $V_h=\{w_1, \ldots, w_j\}$ (то, что каждый член h входит в данное представление $\overline{f}$, сразу следует из определения $\overline{f}$). На основании теоремы 2 (д) заключаем, что $W=(\overline{W}\setminus\overline{W}_h)\cup\overline{W}_h\cdot A\in H_\tau(G)$. При этом

$$f_W=f_{\overline{W}}-f_{\overline{W}_h}+f_{\overline{W}_h\cdot A}=\overline{f}-h+(z_0+\ldots+z_{n-1})h=f$$

и g=1+$f_{\pi(W)}$, так как $g=\overline{g}+h$ и $\pi(W)=\pi(\overline{W})\cup\overline{W}_h$. Теорема доказана.

Индуктивное доказательство достаточности условий (а) — (в) подсказывает способ построения кода по заданной структурной функции. Рассмотрим, например, полином от двух переменных:

$$f=z_0z_1+2z_0^2z_1+z_0z_1^2+z_0^3+2z_0z_1^3+z_1^4+z_0^3z_1^2+z_0^2z_1^3.$$

Выполнение (а), (б) очевидно. После деления f—1 на z_0+z_1 — 1 получаем

$$f-1=(z_0+z_1-1)\left(z_0^2z_1^2+z_1^3+z_0z_1^2+\right.$$
$$\left.+z_1^2+z_0^2+z_0z_1+z_0+z_1+1\right),$$

т. е. (в) тоже выполнено и

$$1)\quad h=h_0=z_0^2z_1^2,\ \text{а}\ \overline{f}-1=(z_0+z_1-1)\left(z_1^3+z_0z_1^2+\right.$$
$$\left.+z_1^2+z_0^2+z_0z_1+z_0+z_1+1\right).$$

Следующие шаги выполняются по одному и тому же шаблону:

$$2)\ h_1=z_1^3+z_0z_1^2,\ \overline{\overline{f}}-1=(z_0+z_1-1)\left(z_1^2+z_0^2+\right.$$
$$\left.+z_0z_1+z_0+z_1+1\right);$$

$$3)\ h_2=z_1^2+z_0^2+z_0z_1;\ \overline{\overline{\overline{f}}}-1=(z_0+z_1-1)(z_0+z_1+1);$$

$$4)\ h_3=z_0+z_1;\ \overline{\overline{\overline{\overline{f}}}}-1=z_0+z_1-1.$$

Отсюда $\overline{\overline{\overline{\overline{f}}}}=z_0+z_1=f_{\{0,1\}}$ и $W_4=\{0, 1\}$. Теперь остается от W_4. вернуться к искомому коду, проходя шаги в обратном направлении согласно правилу

$$W_i=\left(W_{i+1}\setminus W_{h_i}\right)\cup W_{h_i}\cdot A.$$

Результат получится $W=W_0$:

$$W_4 = \{0, 1\}, \quad W_{h_3} = \{0, 1\},$$

$$W_3 = \{00, 01, 10, 11\}, \quad W_{h_2} = \{11, 00, 01\},$$

$$W_2 = \{10, 000, 001, 010, 011, 110, 111\}, \quad W_{h_1} = \{111, 011\},$$

$$W_1 = \{10, 000, 001, 010, 110, 1110, 1111, 0110, 0111\},$$

$$W_{h_0} = \{0110\},$$

$$W = W_0 = \{10, 000, 001, 010, 110, 1110, 1111, 0111, 01100, 01101\}.$$

Сделаем два замечания к сказанному о префиксных кодах.

Замечание 1. Мощность класса

$$H_T^{(m)}(G) = \{W \mid W \in H_T, \ |W| = m\}$$

m-элементных тупиковых префиксных кодов согласно теореме 3(а) отлична от нуля только при $m = (n\text{-}1)k+1$, $k = 1, 2, \ldots$

Как показано выше, по этой подпоследовательности мощность множества различных спектров кодов из $H^m{}_T(G)$ с ростом m стремится к бесконечности с экспоненциальной скоростью.

Замечание 2. Неравенство Мак-Миллана является частным случаем необходимого условия, вытекающего из теоремы 8, которому должен удовлетворять префиксный код W:

(А) $f_W(z_0, \ldots, z_{n-1}) \leq 1$ при любых $z_0, \ldots, z_{n-1}$ таких, что $z_i > 0$, $i = \overline{0, n-1}$, и $\sum_{i=0}^{n-1} z_i = 1$,

и более слабого условия

(Б) $f_W(z_0, \ldots, z_{n-1}) \leq 1$ при некоторых $z_0, \ldots, z_{n-1}$ таких, что

$$z_i > 0, \ i = \overline{0, n-1}, \text{ и } \sum_{i=0} z_i = 1:$$

оно получается при $z_0 = \ldots = z_{n-1} = 1/n$.

Легко проверить, что для конечных префиксных кодов следующие условия равносильны: равенство в (А), равенство в (Б) и $W \in H_T(G)$. При этом, очевидно, для любых префиксных кодов, в том числе и для бесконечных, из первого условия следует второе, а из второго — третье. Однако в общем случае никакие два из этих условий неравносильны.

5. *Хроматическое число графа G* можно определить как наименьшее число независимых множеств, на которое можно разбить множество

его вершин. Существуют графы, для которых число независимости связано с хроматическим числом дополнительного графа

$$\overline{G} = \langle V, \overline{R} \rangle \; (\overline{R} \Rightarrow V^2 \setminus (R \cup \{v_i, v_i\}_{i=1}^n)).$$

Например, если ρ — отношение эквивалентности на множестве *A,* то для его графа G_ρ имеет место

$$\chi(\overline{G}_\rho) = \alpha(G_\rho) = |A/\rho|. \qquad (33)$$

10.5. Комбинаторная теория полугрупп

1. *Полугруппой П =< П, °>* называется множество *П*, замкнутое относительно бинарной ассоциативной операции ° (знак операции в формулах обычно опускается, так как основная операция единственная, а скобки опускаются, так как операция ассоциативная и *(xy)z = x(yz) = xyz).* Содержание комбинаторной теории полугрупп в широком смысле составляет изучепие дискретных полугрупп с содержательно определенной операцией — полугрупп преобразований с операцией суперпозиции, матричных полугрупп и т. п. В более узком смысле *комбинаторная теория полугрупп есть теория свободной полугруппы, ее подполугрупп и представлений полугрупп образующими и гомоморфизмами свободной полугруппы над этими образующими.*

Свободной полугруппой над алфавитом *A* называется множество A^+ с операцией умножения слов, состоящей в приписывании слов, т. е. *x•y=xy.* В некоторых случаях к A^+ удобно добавить единичный элемент — «пустое» слово λ, длину которого полагают равной нулю, и для любого слова *x:* $x\lambda = \lambda x = x$. Полугруппу $A^* \Rightarrow A^+ \cup (\lambda)$ называют иногда свободным моноидом над алфавитом *A*.

Отображение полугруппы < $П_1$, °> в полугруппу < $П_2$, *>, скажем, φ(*x*), называется гомоморфизмом, если

$$\varphi(x_1 \circ x_2) = \varphi(x_1) * \varphi(x_2). \qquad (34)$$

Если при этом φ устанавливает взаимно однозначное соответствие между полугруппами, то оно называется ***изоморфизмом***.

Существует взаимно однозначное соответствие между гомоморфизмами полугруппы и некоторыми отношениями эквивалентности на ней, называемыми ***конгруэнтностями***. Отношение эквивалентности ρ на полугруппе *П* называется *стабильным слева,* если из $\langle x, y \rangle \in \rho$ следует, что для любого z: $\langle zz, xy \rangle \in \rho$; *стабильным справа,* если из $\langle x, y \rangle \in \rho$ следует, что для любого z: $\langle xz, yz \rangle \in \rho$. Отношение эквивалентности ρ называется *конгруэнтностью,* если оно стабильно и слева, и справа.

Теорема 1. *Если* φ — *гомоморфизм полугруппы П, то отношение* ε_φ: $<a, b> \in \varepsilon_\varphi \Rightarrow \varphi(a) = \varphi(b)$ *есть конгруэнтность. Если* ρ *есть отношение конгруэнтности на полугруппе П, то П/*ρ *есть полугруппа, в которой* $X \circ Y$ *есть класс Z конгруэнтности* ρ, *который содержит целиком* $X \circ Y = \{ab \mid a \in X,\ b \in Y\}$, *а отображение* φ_ρ, *сопоставляющее каждому* $a \in П$ *класс* $\varphi_\rho(a)$, *содержащий элемент а, есть гомоморфизм.*

Теорема 2. *Если П — конечно порожденная полугруппа и* $A=\{a_1,...,a_n\}$ — *множество имен образующих П, то на* A^+ *существует отношение конгруэнтности* ρ *такое, что П изоморфна* A^+/ρ.

Средством описания отношений на полугруппе являются:

соотношения — равенства слов в алфавите имен образующих;

смешанные соотношения — равенства слов, включающие, кроме имен образующих, символы переменных x_i, значениями которых могут быть любые слова;

тождества — частный вид смешанных соотношений, когда отсутствуют вхождения собственных имей.

Соотношение и тождество называются нетривиальными, если слова в левой и правой частях графически различны.

Дополнительные возможности описания отношений представляют *условные тождества* (называемые также *квазитождествами)* — законы вида

$$\forall x_1, \ldots x_N \quad \left[\underset{i=1}{\overset{k}{\&}} \alpha_i(x_1, \ldots, x_N) = \beta_i(x_1, \ldots, x_N) \right] \to \to \alpha(x_1, \ldots, x_N) = \beta(x_1, \ldots, x_n), \qquad (35)$$

с помощью которых формулируются формальные, синтаксические правила вывода соотношений.

В любой полугруппе выполняется квазитождество

$$\forall x_1, x_2, y, z \qquad (x_1 = x_2) \to (yx_1z = yx_2z). \qquad (36)$$

Пусть *s* — правило вывода соотношений согласно (36), т. е. соотношение вида $yx_1z = yx_2z$ выводится из $x_1=x_2$ по правилу *s*. Если $s^*(\psi) = \rho$, то говорят, что *П = П<A*| ψ > — представление полугруппы *П* образующими *А* и определяющими соотношениями ψ над этими образующими. Если существует такое представление с конечным множеством ψ, то *П* называется конечно определенной.

Теорема 3. *Если полугруппа конечно определена по отношению к какому-либо множеству образующих, то она конечно определена и по отношению к любому другому конечному множеству образующих.*

Заметим, что $s^*(\psi)$ есть наименьшее отношение конгруэнтности на полугруппе, содержащее ψ. Поэтому ψ семантически определяет отношение $\rho=s^*(\psi)$ в классе отношений конгруэнтности на полугруппе.

Теорема 4. *Если полугруппу П можно изоморфно вложить в группу, то в П выполняется квазитождество Мальцева: для любых $x, y, z, u, x', y', z', u'$*

$$(yx=y'x' \mathbin{\&} yz=y'z' \mathbin{\&} uz=u'z') \rightarrow ux=u'x'. \qquad (37)$$

В дальнейшем мы будем использовать для соотношений вида $\alpha=\beta$ равноценную запись в виде дробей: $\frac{\alpha}{\beta}$ или $\frac{\beta}{\alpha}$. В частности, это удобно в связи с тем, что с умножением $\frac{\alpha_1}{\beta_1}\cdot\frac{\alpha_2}{\beta_2}=\frac{\alpha_1\alpha_2}{\beta_1\beta_2}$ соотношения в любой полугруппе, учитывая (36), сами образуют полугруппу. Используя такую запись, правилу вывода соотношений на основе квазитождества Мальцева можно придать следующую форму: из соотношений $\frac{yx}{y'x'}$, $\frac{yz}{y'z'}$, $\frac{uz}{u'z'}$ выводится $\frac{ux}{u'x'}$, или, символически,

$$\begin{tikzcd}
\frac{x}{x'} & \frac{y}{y'} \arrow[l] \arrow[r] & \frac{z}{z'} & \frac{u}{u'} \arrow[l] \arrow[lll, dashed, bend right]
\end{tikzcd},$$

подразумевая, что $\frac{\alpha}{\beta}\rightarrow\frac{\gamma}{\delta}$ означает соотношение $\frac{\alpha\gamma}{\beta\delta}$, $\frac{\alpha}{\beta}\leftarrow\frac{\gamma}{\delta}$ — соотношение $\frac{\gamma\alpha}{\delta\beta}$, сплошные стрелки означают условия, а пунктирная — вывод из условий.

Через $\left[\frac{\alpha}{\beta}\right.$ будет обозначаться результат максимального сокращения соотношения $\frac{\alpha}{\beta}$ слева, например,

$$\left[\frac{ab}{aba}\right.=\left[\frac{c}{ca}\right.=\frac{\lambda}{a},\quad \left[\frac{ab}{ac}\right.=\left[\frac{bb}{bc}\right.=\frac{b}{c}.$$

Двойственное обозначение для сокращения справа — $\left.\frac{\alpha}{\beta}\right]$.

Проблема конечной определенности полугрупп имеет большое значение, но не следует его переоценивать.

Теорема 5. *Существуют конечно определенные полугруппы, для которых проблема равенства слов алгоритмически неразрешима* (т. е. проблема распознавания по паре слов, равны эти слова в полугруппе или нет, иными словами — представляют один и тот же элемент полугруппы или нет?).

Таким образом, с представлениями полугрупп образующими и определяющими соотношениями связаны принципиальные проблемы, в отношении решения которых дело обстоит качественно хуже, чем с универсальными переборными проблемами.

2. Рассмотрим некоторые вопросы, связанные с *разложениями слов на множители.* Такое разложение можно задавать *скобочной* или *поименной* записью (вместо скобок можно использовать какой-либо разделительный знак, скажем, вертикальную черту). Например, скобочной записи α=(01)(010)(010)(01), где α=0101001001, соответствует поименная $b_1b_2b_2b_1$, если некоторым словам присвоить имена так, что $b_1 \Rightarrow 01$, $b_2 \Rightarrow 010$. Очевидно, что поименная запись однозначно определяет соответствующую ей скобочную: $b_1b_2b_2 \Rightarrow$
$\Rightarrow$ (01)(010)(010)...

Число разложений слова длины N в произведение i непустых слов равно числу способов расставить i—1 разделительных знаков в (N — 1) промежутке между буквами слова и есть поэтому $\binom{N-1}{i-1}$, $i = \overline{1, N}$. Следовательно, общее число таких разложении слова длины N равно $\sum_{i=1}^{N}\binom{N-1}{i-1} = 2^{N-1}$.

Сложнее перечислить все разложения слова в произведение непустых сомножителей из заданного множества. Пусть $V = \{v_1, ..., v_m, \dots\}$ — множество слов, $B =\{b_1, ..., b_m, \dots\}$ — алфавит имен слов из V в том же порядке. Пусть $F_v(\alpha)$ — множество всех поименных разложений слова α в произведение слов из V. Положим

$$J(V,\ \alpha) \Rightarrow \{j \mid \alpha = v_j\alpha_j\} \tag{38}$$

и по определению $F_V(\lambda) \Rightarrow \{\ \lambda\ \}$, Легко проверить, что в таком случае для вычисления $F_V(\alpha)$ при $\alpha \neq \lambda$ справедлива рекуррентная формула:

$$F_V(\alpha) = \bigcup_{j \in J(V,\alpha)} b_j F_V(\alpha_j). \tag{39}$$

Например, для $V = \{01,\ 100, 010,\ 1001\}$ и $\alpha = 0101001001$ находим

$$F_V(\alpha) = b_1F_V(01001001) \cup b_3F_V(1001001) =$$
$$= b_1b_1F_V(001001) \cup b_1b_3F_V(01001) \cup$$
$$\cup\, b_3b_2F_V(1001) \cup b_3b_4F_V(001) = b_1b_3b_1F_V(001) \cup$$
$$\cup\, b_1b_3b_3F_V(01) \cup b_3b_2b_2F_V(1) \cup \{b_3b_2b_4\} =$$
$$= \{b_1b_3b_3b_1,\ b_3b_2b_4\}.$$

Установим условия, при которых в полугруппе V^+ свободной полугруппы A^+ выполняется свойство однозначности разложения на множители по базису V (т. е. условия того, что для любого $\alpha \in A^+$ имеет место $|F_V(\alpha)| \leq 1$). Имеется несколько однотипных критериев. Далее, в этом пункте предполагается, что V — базис, т. е. для слов $v \in V$ имеет место $|F_V(v)| = 1$.

Теорема 1. *Для выполнения свойства единственности разложения слов на простые (неразложимые) сомножители в подполугруппе V^+ необходимо и достаточно, чтобы для любых α, $\beta,\gamma,\delta \in V^+$ из соотношений $\alpha x = \beta$ и $x\gamma=\delta$ следовало $x \in V^*$.*

Доказательство. ***Необходимость.*** Пусть свойство единственности выполнено для V^+, и пусть $\alpha x = \beta$ и $x\gamma=\delta$, где α, $\beta,\gamma,\delta \in V^+$. Тогда для слова $\alpha x\gamma$ существуют разложение вида $\alpha x|\gamma$ (так как $\alpha x = \beta \in V^+$ и $\gamma \in V^+$) и разложение вида $\alpha|x\gamma$ (так как $x\gamma = \delta \in V^+$ и $\alpha \in V^+$). Но, ввиду единственности, это одно и то же разложение, включающее оба разделительных знака: $\alpha|x|\gamma$, откуда $x \in V^*$.

Достаточность. Предположим, что единственности нет. Тогда существует слово $\alpha\beta\gamma$ такое, что $\beta \notin V^*$, и есть разложения $\alpha|\beta\gamma$, в котором между β и γ нет разделительного знака, и $\alpha\beta|\gamma$, в котором между α и β нет разделительного знака. Но в таком случае мы имели бы $\alpha\beta=u \in V^+$, $\beta\gamma=v \in V^+$, $\alpha \in V^+$, $\gamma \in V^+$, но $\beta \notin V^*$, что противоречит условию. Теорема доказана.

Легко видеть, что условие в теореме 1 эквивалентно следующему:

$$\text{для любого } x\text{: } V^+x \cap V^+ \neq \varnothing \text{ и } xV^+ \cap V^+ \neq \varnothing \text{ влечет } x \in V^*. \quad (40)$$

Рассмотрим еще два частных случая условия (40):

$$\text{если } \alpha,\ \beta,\ \gamma \in V^+,\ x\alpha = \gamma,\ \beta x = \gamma,\ \text{то } x \in V^* \quad (41)$$

(эквивалентно: $xV^+ \cap V^+ \cap V^+x \neq \varnothing$ влечет $x \in V^*$), и

$$\text{если } \alpha,\ \beta,\ \gamma \in V^+,\ \alpha x = \beta,\ x\alpha = \gamma,\ \text{то } x \in V^* \quad (42)$$

(эквивалентно: если $\alpha\beta \in V^+$ и $\beta\alpha \in V^+$, то либо $\alpha \in V^*$ и $\beta \in V^*$, либо $\alpha \notin V^*$ и $\beta \notin V^*$).

В действительности каждое из этих условий эквивалентно (40) и может служить критерием единственности разложения на множители. Пусть, например, выполнено (42) и не выполнено (40): $\alpha, \beta, \gamma, \delta \in V^+$, $\alpha x = \beta, x\gamma = \delta$ и $x \notin V^*$. Умножим первое соотношение слева на γ, а второе справа на α. Получаем $(\gamma\alpha)x = \gamma\beta$, $x(\gamma\alpha) = \delta\alpha$, $\gamma\alpha$, $\gamma\beta$, $\delta\alpha \in V^+$, $x \notin V^*$, — противоречие с (42). Аналогично проверяется равносильность (41) и (40).

Условие префиксности кода V — достаточное для единственности разложения слов V^+ на простые множители. Это легко вывести непосредственно из (39) и формально, записав его в виде

$$Vx \cap V^+ \neq \varnothing \to x \in V^* \tag{43}$$

и сравнив с (40). Оно не является необходимым.

Теорема 1 и эквивалентные ей критерии (41), (42) не являются алгоритмическими. На вопрос о том, какова длина кратчайшего из неоднозначно разложимых слов, если такие существуют, они не дают ответа.

Еще одни критерий единственности разложения слов на множители по множеству слов V можно сформулировать в терминах структурных функций. Здесь V может и не быть базисом. Пусть $A = \{a_1, ..., a_n\}$ — алфавит V, т. е. множество всех букв, имеющих вхождение в слова V, в

$$f_V(z_1, \ldots, z_n) = \sum_{v \in V} \left(\prod_{i=1}^{n} z_i^{|v|_i} \right)$$

— структурная функция V, которую мы рассматриваем в области $z_1 > 0$, $z_2 > 0$, ..., $z_n > 0$. Тогда, очевидно, для любого N по смыслу производящей функции имеет место

$$f_{V^N}(z_1, \ldots, z_n) \leqslant (f_V(z_1, \ldots, z_n))^N. \tag{44}$$

Учитывая, что строгое неравенство в (44) имеет место только при условии, что существует слово, допускающее два разложения в произведение N слов из V, и что если существует слово, допускающее два разложения на множители из V, скажем, $b_{i_1} \ldots b_{i_p}$ и $b_{j_1} \ldots b_{j_q}$, то найдется и слово, допускающее два разложения на одинаковое число сомнножителей из V, именно

$$b_{i_1} \ldots b_{i_p} b_{j_1} \ldots b_{j_q} \text{ и } b_{j_1} \ldots b_{j_q} b_{i_1} \ldots b_{i_p},$$

получаем теорему.

Теорема 2. *Для выполнения свойства единственности разложения на множители из V необходимо и достаточно, чтобы для любого* N — 1, 2, ... *имело место*

$$f_{V^N}(z_1, \ldots, z_n) = (f_V(z_1, \ldots, z_n))^N.$$

3. Далее в этом параграфе будем рассматривать свободные полугруппы и $\mathfrak{I}$-полугруппы — полугруппы, изоморфные конечно порожденным подполугруппам свободных полугрупп. Представления $\mathfrak{I}$-полугрупп образующими в A^+ (т. е. задания множествами слов V таким образом, что $\Pi = V^+$) называются *словесными представлениями*.

Пусть $\mathfrak{I}$-полугруппа $\Pi = V^+$ задана словесным представлением $V = \{v_1, \ldots, v_m\}$ и $B = \{b_1 \ldots, b_m\}$ — алфавит имен образующих в том же порядке. Пусть f - естественный гомоморфизм B^+ на V^+:

$$f(b_i) = v_i \quad (j = \overline{1, m}), \quad f(b_{i_1} \ldots b_{i_k}) = f(b_{i_1}) \ldots f(b_{i_k})$$

$$\text{и} \quad R(V) = \left\{ \frac{\alpha}{\beta} \middle| f(\alpha) = f(\beta) \right\}.$$

R(V) есть отношение конгруэнтности на B^+, соответствующее гомоморфизму f, поэтому $\Pi = \Pi\langle B | R(V)\rangle$. Пусть

$$R_1(V) \Rightarrow R(V) \backslash R(V)^2,$$

т. е. $R_1(V)$ есть множество всех соотношений $R(V)$, которые не могут быть разложены в произведение соотношений $R(V)$.

Теорема 1. *$R(V) = R_1(V)^*$, причем $R_1(V)$ — множество свободных образующих $R(V)$.*

Доказательство. Если $\frac{\alpha}{\beta} = \frac{\alpha_1}{\beta_1}\frac{\alpha_2}{\beta_2} \in R(V)$ и $\frac{\alpha_1}{\beta_1}, \frac{\alpha_2}{\beta_2} \in R(V)$, то $\left| f\left(\frac{\alpha}{\beta}\right) \right| = \left| f\left(\frac{\alpha_1}{\beta_1}\right) \right| + \left| f\left(\frac{\alpha_2}{\beta_2}\right) \right|$ (для соотношения $\frac{\alpha}{\beta} \in R(V)$ через $f\left(\frac{\alpha}{\beta}\right)$ обозначаем слово $f(\alpha)$, равное $f(\beta)$), где $\left| f\left(\frac{\alpha_1}{\beta_1}\right) \right| > 0$ и $\left| f\left(\frac{\alpha_2}{\beta_2}\right) \right| > 0$.

Поэтому каждое соотношение $R(V)$ представимо в виде произведения соотношений из $R_1(V)$, число которых заведомо не больше числа букв в соотношении. Единственность такого представления следует из того, что

$\frac{\alpha}{\beta} = \frac{\alpha_1}{\beta_1}\frac{\alpha_2}{\beta_2} \in R(V)$ и $\frac{\alpha_1}{\beta_1} \in R(V)$ влечет выполнение $\frac{\alpha_2}{\beta_2} \in R(V)$, так как если $f(\alpha_1\alpha_2) = f(\alpha_1)f(\alpha_2) = f(\beta_1)f(\beta_2) \Rightarrow f(\beta_1\beta_2)$ и $f(\alpha_1) = f(\beta_1) = \gamma$, то после сокращения слева на γ получаем $f(\alpha_2) = f(\beta_2)$, что равносильно $\frac{\alpha_2}{\beta_2} \in R(V)$. Теорема доказана.

Следствие. $\Pi = \Pi\langle B | R_1(V)\rangle$.

Замечание. Когда целесообразно, естественный, гомоморфизм f продолжают на свободный моноид B^*, полагая $f(\lambda) = \lambda$.

4. В этом пункте рассмотрим бескоэффициентные уравнения в свободной полугруппе, проще говоря — *уравнения в словах.* Пусть ρ:

$$f_i(x_1, \ldots, x_m) = g_i(x_1, \ldots, x_m), \; i \in J, \tag{45}$$

— система уравнений в словах (f_i, g_i — слова в алфавите неизвестных). Набор слов $X_0 = \langle x^0_1, \ldots, x^0_m \rangle$ называется решением системы (45), если при подстановке левая и правая части каждого уравнения совпадают графически. Набор слов называется *унтер-решением* системы (45), если $R(X_0) \subseteq s^*(\rho)$, т. е. если в полугруппе X_0^* нет никаких нетождественных соотношений, кроме, быть может, соотношений ρ и их следствий. Так, всякое множество свободных образующих, в том числе всякий префиксный код, является унтер-решением.

Унтер-решение называется *экстремальным,* если его спектр является минимальным элементом в множестве спектров всех унтер-решений системы (т. е. в этом множестве нет спектра, который им мажорируется по отношению покоординатного сравнения).

Пусть $Y(\rho)$ — множество всех унтер-решений системы ρ, $Y^0(\rho)$ — множество всех экстремальных унтер-решений этой системы.

Теорема 1. *Для любой системы уравнений* ρ *множество ее экстремальных унтер-решений* $Y^0(\rho)$ *конечно.*

Утверждение теоремы следует непосредственно из теоремы Диксона.

Теорема 2. *Не существует алгоритма, вычисляющего по произвольной конечной системе уравнений в словах* ρ *множество ее экстремальных унтер-решений* $Y^0(\rho)$.

Доказательство. С произвольной системой уравнений ρ над переменными $X = \{x_1, \ldots, x_m)$, включающей уравнения $x_i x_1 = x_1 x_i = x_i$ для всех i=l, 2. …, m, связываем полугруппу $\Pi = \Pi\langle X|\rho\rangle$, в которой x_1 является единицей, и полугруппу $\Pi_0 = \Pi\langle X \cup \{x_0\} | \rho_0\rangle$, где

$$\rho_0 = \rho \cup \{x_1 x_0 = x_0, \; x_0 x_1 = x_0\}$$

(Π есть свободное произведение $\Pi * \{x_0^i\}_{i=0}^{\infty}$ полугрупп с общей единицей — классом, включающим x_1).

Покажем, что $\langle a, \lambda, \ldots, \lambda\rangle \in Y^0(\rho_0)$ в том и только том случае, если Π — единичная (т. е. одноэлементная) полугруппа.

Пусть Π — единичная. Тогда $\langle a, \lambda, \ldots, \lambda\rangle \in Y^0(\rho_0)$, так как $\alpha \Rightarrow \alpha_1 a \alpha_2 a \ldots \alpha_k a \alpha_{k+1} = a^k$ в Π_0, и если

$$\beta \Rightarrow \beta_1 a \beta_2 a \ldots \beta_l a \beta_{l+1} = a^k \text{ в } \Pi_0,$$

то $k = l$ и $\alpha_1 = \beta_1, \ldots, \alpha_{k+1} = \beta_{k+1}$ в Π. Следовательно, $\langle\alpha, \beta\rangle \in \rho_0$.

Обратно: если $\langle a, \lambda, \ldots, \lambda\rangle \in Y^0(\rho_0)$, то из $\alpha \neq \beta$ в Π следует $\langle \alpha a, \beta a\rangle \notin \rho_0$ в Π_0, но в Π_0 имеем по предположению $a\alpha = \beta a = a$ — противоречие. Следовательно, для любых α,β пара $\langle \alpha, \beta\rangle \in \rho$ в Π, т. е. Π — единичная.

Но проблема распознавания единичности для конечно определенных полугрупп алгоритмически неразрешима, откуда следует утверждение теоремы.

Что касается решений уравнений в словах, проблема состоит в нахождении общего решения, описывающего с точностью до каких-либо преобразований все решения. Некоторые из относящихся сюда результатов содержатся в следующих пунктах.

5. Рассмотрим уравнения в словах с двумя неизвестными. Для слова α обозначим через $\varepsilon(\alpha)$ примитивный корень этого слова, т. е. кратчайшее из слов β, для которых верно $\alpha\{\beta\}^+$. В представлении $\alpha = \varepsilon(\alpha)^\tau$ натуральное число $\tau = \tau(\alpha)$ — показатель слова α. Слово α называется простым, если $\varepsilon(\alpha)=\alpha$, в противном случае — периодическим ($\tau(\alpha)>1$). Очевидно, что для любого слова α: $\varepsilon(\varepsilon(\alpha)) = \varepsilon(\alpha)$, т. е. $\varepsilon(\alpha)$ — простое (если бы $\varepsilon(\alpha) = \beta^i$, $i \geqslant 2$, то имели бы $\alpha = \beta^{i\tau}$, $|\beta| < |\varepsilon(\alpha)|$).

Теорема 1. *Пусть* $\delta = \varphi(\alpha, \beta) = \psi(\alpha, \beta)$, $\alpha \neq \lambda$, $\beta \neq \lambda$, *и* $\varphi(x, y) \neq \psi(x, y)$. *Тогда существует слово* γ *такое, что* $\alpha = \gamma^i$ *и* $\beta = \gamma^j$ *для некоторых натуральных чисел* i, j.

Доказательство. Индукция по $|\delta| = k$. При $k = 1$ очевидно, что $\alpha = \beta = \gamma$. Пусть утверждение верно при $k < m$ и $|\delta| = m$. Если $\alpha \neq \beta$, то α и β находятся в отношении префиксности, скажем, $\alpha=\beta\bullet\omega$. Пусть $\varphi = x\varphi_1(x, y)$, а $\psi = y\psi_1(x, y)$. Тогда из

$$\beta\omega\varphi_1(\beta\omega, \beta) = \beta\psi_1(\beta\omega, \beta)$$

следует

$$\delta' = \omega\varphi_1(\beta\omega, \beta) = \psi_1(\beta\omega, \beta) \text{ и } |\delta'| < m.$$

Поэтому по предположению индукции $\omega=\gamma^i$, $\beta=\gamma^j$ и, следовательно, $\alpha = \gamma^{i+j}$, $\beta = \gamma^j$. Теорема доказана. В качестве следствий этой теоремы получаем:

$$\varepsilon(\alpha^i) = \varepsilon(\alpha) \text{ для любого слова } \alpha \text{ и натурального } i; \tag{46}$$

$$\text{если } \alpha = \beta^i, \text{ то } \beta = \varepsilon(\alpha)^{\tau(\alpha)/i}, \ \tau(\alpha) = i \cdot \tau(\beta). \tag{47}$$

(46) докажем индукцией по $|\alpha|$. При $|\alpha| = 1$ имеем $\varepsilon(\alpha^i) = \varepsilon(\alpha) = \alpha$.

Пусть $|\alpha| = k$. Тогда $\alpha^i = \left(\varepsilon\left(\alpha^i\right)\right)^{\tau(\alpha^i)}$, $\alpha = \varepsilon(\alpha)^{\tau(\alpha)}$ из $\alpha^{i+1} = \alpha \bullet \alpha^i = \alpha^i\alpha$ и по теореме 1, учитывая, что $\varepsilon(\alpha^i)$ и $\varepsilon(\alpha)$ — простые слова, заключаем, что $\varepsilon(\alpha^i) = \varepsilon(\alpha)$. (46) доказано.

Если $\alpha = \beta^i$, то $\varepsilon(\alpha) = \varepsilon(\beta^i) = \varepsilon(\beta)$, откуда

$$\alpha = (\varepsilon(\alpha))^{\tau(\alpha)} = (\varepsilon(\alpha)^{t(\beta)})^{i} = \varepsilon(\alpha)^{it(\beta)}$$

и (47) доказано.

Пусть $\alpha = a_1a_2...a_N$. Через $\alpha^{(i)} \Rightarrow a_ia_{i+1}\dots a_{i+N-1}$ обозначается циклический сдвиг слова α на i — 1 разрядов вправо (индексы в такой записи понимаются как наименьшие положительные значения по модулю

$$N = |\alpha|,\ \alpha^{(1)} = \alpha^{(N+1)} = \alpha).$$

Пусть $C(\alpha) = \{\alpha^{(i)} | i = \overline{1,\ N}\}$ — циклокласс, порожденный словом α. Легко проверить, что для любого слова α и любого i

$$\varepsilon(\alpha^{(i)}) = (\varepsilon(\alpha))^{(i)}, \qquad (48)$$

$$C(\alpha) = \{(\varepsilon(\alpha^{(j)}))^{\tau(\alpha)} | j = \overline{1,\ |\varepsilon(\alpha)|}\}, \qquad (49)$$

$$|C(\alpha)| = |\varepsilon(\alpha)|. \qquad (50)$$

6. Задача нахождения решений уравнений в словах двойственна задаче перечисления соотношений между образующими $\mathfrak{I}$-полугрупп.

Словесное представление $\mathfrak{I}$-полугруппы V называется минимальным, если для любого изоморфного словесного представления V' и изоморфизма f, индуцируемого соответствием $f(v_i) = v^1_i$, имеет место либо $|f(v_i)| = |v_i|$ для всех $i = 1, 2,\dots,|V|$, либо $|f(v_i)| > |v_i|$ для некоторого i, т. е. если V не может быть «сжат».

Например, *{aba, ab, ba}* не минимально, так как сжатие, определенное соответствием

$$aba \to a,$$

$$ab \to ab,$$

$$ba \to ba,$$

есть изоморфизм.

Теорема 1. *Пусть* $V = \{v_1, \dots, v_m\}$, $W = \{w_1, \dots, w_n\}$ *и для* $i = \overline{1,\ m}$ $v_i = \varphi_i(w_1, \dots, w_n)$, *причем каждое* x_j $(j = \overline{1,\ n})$ *входит по крайней мере в одно из слов* $\varphi_i(x_1, \dots, x_n)$, $i = \overline{1,\ m}$. *Тогда, если W не минимально, то и V не минимально.*

Доказательство. Если ψ есть изоморфизм полугруппы W^+, то его ограничение на подполугруппу V^+ есть тоже изоморфизм. Ясно, что если ψ сжимал W, то при условии теоремы ψ сжимает и V. Теорема доказана.

Теорема 2. *Если выполняется свойство единственности разложения на множители по* $V = \{v_1, ..., v_m\}$, *то V минимально в том и только том случае, если* $|v_1| = |v_2| = \dots = |v_m| = 1$, *т. е. V есть алфавит.*

Это — непосредственное следствие предыдущей теоремы.

Теорема 3. *Пусть A, B— алфавиты.* $A \cap B = \varnothing$, $T_n(A, B) = \{B, AB, A^2B, \ldots, A^{n-1}B, A^n\}$ $(n > 1)$. *Тогда:*

(а) *$T_n(A, B)$ — префиксный код;*

(б) *если* $V \subseteq T_n(A, B)^*$ *и* $A(V) \cap A \neq \varnothing$, *то V не минимально;* (для мложества слов V через $A(V)$ обозначается алфавит всех букв, входящих в слова V, $A'(V)$ и $A''(V)$ — алфавиты всех первых и всех последних букв V соответственно).

(в) *если для слова* α: $A(\alpha) \subseteq A \cup B$, *то* $\alpha \notin T_n(A,B)^*$ *в том и только том случае, если* $\alpha \in A^m$ *и n не является делителем m либо если* $\alpha = \alpha' b \alpha''$, *где* $b \in B$ *и* $\alpha'' \in A^m$, $m > 0$. *и n не является делителем m.*

Доказательство, (а) очевидно, (б) следует из того, что свойство единственности разложения на множители по множеству сохраняется для любого подмножества, и теорем 1, 2, (в) легко проверяется индукцией по | α |. Теорема доказана.

Теорема 4. *Если V минимально, то*

$$A(V) = A'(V) = A''(V).$$

Доказательство. Если, скажем, $A''(V) \neq A(V)$, то для любого $n > 1$ имеем $V \subset T_n(A(V) \setminus A''(V),\ A''(V))^*$ и

$$A(V) \cap (A(V) \setminus A''(V)) \neq \varnothing.$$

Тогда по теореме 3 V не минимален. Рассуждение для $A'(V)$ совпадает с этим с точностью до обращения всех слов в множествах V и T. Теорема доказана.

Из теоремы 4 вытекает важное свойство решений систем уравнений в словах. Если разложения на множители по словам $V = \{v_1, \ldots, v_m\}$ имеют свойство единственности, $B = \{b_1, \ldots, b_m\}$, то V^+ изоморфна свободной полугруппе с m образующими. В этом случае естественный изоморфизм $f = f_V$ полугруппы B^+ на V^+ называется свободным изоморфизмом. Если при этом V — префиксным код или обращение префиксного кода, то f_V называется префиксным или соответственно обратно-префиксным изоморфизмом. Легко проверить, что классы свободных изоморфизмов и префиксных изоморфизмов замкнуты относительно суперпозиции отображений. Если множество $W \subseteq B^+$, то его образ $f(W)$ при свободном изоморфизме называется его производным.

Как следует из теоремы 4, если V минимально, то $|A'(V)| = |A''(V)| \leq |V|$, а если свойство единственности разложения на множители по V не выполняется, то $|A'| < |V|$. Учитывая это, получаем следующую переформулировку теоремы 4.

Теорема 5. *Пусть* $\mathfrak{D}$— *множество всех решений системы уравнений в словах* (45) *в (т — 1)-буквенном алфавите, а* $\mathfrak{Y}$— *множество всех ее унтер-решений в т-буквенном алфавите. Тогда множество всех решений* (45) *и множество всех ее унтер-решений являются соответственно производными от* $\mathfrak{D}$ *и* $\mathfrak{Y}$.

7. Здесь покажем, что за редким исключением средством синтаксического задания $\mathfrak{F}$-полугрупп могут быть только соотношения между именами образующих, и выявим соответствующее исключение.

Теорема 1. *Если в* $\mathfrak{F}$-*полугруппе выполняется нетривиальное смешанное соотношение, то она коммутативна и все тождества, которые в ней выполняются, являются следствиями XY = YX, а любое смешанное соотношение перестановками сомножителей и последующим сокращением приводится к обычному соотношению.*

Доказательство. Пусть $\alpha x_i\varphi = \beta x_j\psi$ — неразложимое смешанное соотношение в $\mathfrak{F}$-полугруппе, V — минимальное словесное представление этой полугруппы $\{v_1, \ldots, v_m\}$, $B = \{b_1, \ldots, b_m\}$ и f — естественный гомоморфизм B^* на V^+. Соотношение однородно (т. е. всякая переменная входит одинаковое число раз в левую и правую части), так как при любых подстановках значений переменных длина слова в левой части должна равняться длине слова в правой части.

Возможны случаи:

(1) $\alpha=\beta$ — пустое слово и $i \neq j$;

(2) α или β отлично от λ, тогда $f(\alpha)$ и $f(\beta)$ находятся в отношении префиксности, скажем, $f(\alpha) = f(\beta)\bullet\gamma$ и $\gamma \neq \lambda$. Поскольку вместо любой переменной можно подставлять любое слово из V и равенство должно выполняться, все слова V должны начинаться на одну и ту же букву (первую букву γ в случае (2)). Тогда по теореме 4.6 $|A(V)| = 1$ и V^+ изоморфна коммутативной подполугруппе полугруппы натуральных чисел по сложению. Любое однородное тождество, очевидно, выводится из $XY = YX$, а в смешанном соотношении все переменные перестановками можно перевести в каждой части в префикс и, после сокращения, вхождений переменных не останется ввиду однородности. Теорема доказана.

Тождественные соотношения вообще накладывают сильные ограничения на класс полугрупп, в которых они выполняются. Например, в класс коммутативных полугрупп, т. е. полугрупп, в которых выполняется тождество $XY = YX$, всякая конечно порожденная полугруппа конечно определена, т. е. для ее синтаксического задания к тождеству коммутативности достаточно добавить конечное число соотношений.

В то же время в классе $\mathfrak{F}$-полугрупп существуют такие, которые не являются конечно определенными. Вопрос о сложности описания соотношений в $\mathfrak{F}$-полугруппах изучался разными авторами. Ими получен алгоритмический критерий конечной определенности $\mathfrak{F}$-полугруппы, заданной словесным представлением. С его помощью можно показать, что, например, $\mathfrak{F}$-полугруппа, заданная образующими *{a, ab, ba, bb},* не является конечно определенной. Подобный пример с менее чем 4 образующими невозможен: как показано в одной из работ, *все $\mathfrak{F}$-полугруппы с тремя и менее образующими конечно определены.* В ряде работ найдены все представления $\mathfrak{F}$-полугрупп с тремя образующими неприводимыми системами определяющих соотношений.

8. **Примеры.** (а) Найдем представление $\mathfrak{F}$-полугруппы, порожденной образующими *V={a, ba, ab},* определяющими соотношениями над образующими

$$x \Rightarrow a,\ y \Rightarrow ba,\ z \Rightarrow ab.$$

1) Покажем, что

$$R_1(V) = \left\{ \frac{x}{z} \left(\frac{y}{z} \right)^i \frac{y}{x} \right\}_{i=0}^{\infty} \tag{51}$$

Для любых $\alpha,\ \beta \in \{a,\ b\}^*$, как легко проверить, имеем

$$F_V(ba\alpha) = yF_V(\alpha),$$

$$F_V(\beta ab) = F_V(\beta)z,$$

$$F_V(\alpha aa\beta) = F_V(\alpha a)F_V(a\beta),$$

$$F_V(\alpha bb\beta) = F_V(\alpha b)F_V(b\beta).$$

Отсюда следует, что если $\frac{\mu}{\nu}$ неразложимо и $\mu,\ \nu \in F_V(\alpha)$, то α не содержит вхождений *aa* и *bb,* т. е. имеет вид α = $(ab)^i a$ для некоторого натурального числа *i* Индукцией по *i* находим $F_V\left((ab)^i a\right) = \{z^j x y^{i-j}\}_{j=0}^{i}$. Из этого множества можно выбрать, очевидно, только одну несократимую пару $xy^i = z^i x$, откуда следует (51).

2) Любое из соотношений (51) выводится из $xy = zxx$:

$$xy^{i+1} = zxy^i = z^2xy^{i-1} = \ldots = z^ixy = z^{i+1}x.$$

В результате получаем

$$V^+ = \Pi\langle\{x,\ y,\ z\}|xy = zx\rangle.$$

(б) Рассмотрим обратную задачу: описать все решения уравнения

$$xy = zx. \tag{62}$$

Так как x и z находятся в отношении префиксности, должно выполняться $x = zx_1$ либо $z= xz_1$, где x_1 и z_1 отличны от λ (решения, в которых одно из слов пустое, исчерпываются, как легко проверить, следующими наборами <λ, λ, λ>, <α, λ, λ > и < λ, α, α >; решения, соответствующие $x_1 = \lambda$ в первом случае и $z_1 = \lambda$ во втором случае — < α, α, α >, где α — произвольное слово).

В первом случае подстановка в (52) дает $zx_1y = zzx_1$ и после сокращения на z слева — $x_1y=zx_1$. Следовательно, решение, соответствующее этому случаю, получается из некоторого решения $\langle x, y, z\rangle$ того же уравнения (52), но с меньшей суммой длин слов, операцией

$$\langle x_1,\ y,\ z\rangle \to \langle zx_1,\ y,\ z\rangle. \tag{53}$$

Во втором случае получаем аналогично $xy = xz_1x$, откуда $y = z_1x$, т. е. решение имеет вид

$$\langle x,\ z_1x,\ xz_1\rangle, \tag{54}$$

где x, y, z — произвольные непустые слова. Проверкой убеждаемся, что (54) является решением (52) при любых x, z_1 следовательно, любое решение (52) получается из (54) с произвольными x, z_1 операциями (53):

$$\langle x,\ z_1x,\ xz_1\rangle \to \langle (xz_1)x,\ z_1x,\ xz_1\rangle \to \ldots$$
$$\ldots \to \langle (xz_1)^i x,\ z_1x,\ xz_1\rangle \to \ldots$$

Итак, общее решение можно записать в виде

$$\langle (\alpha\beta)^i\alpha,\ \beta\alpha,\ \alpha\beta\rangle,\quad \langle \alpha,\ \lambda,\ \lambda\rangle, \tag{55}$$

где α, β — произвольные слова, i — произвольное неотрицательное целое (< α, α, α > получается из (55) при β = λ, i = 0, <λ, λ, λ>— при α = β = λ, < λ, β, β > — при α = λ, i = 0).

Нетривиальные решения (т. е. соответствующие случаю α ≠λ, β ≠ λ и ε(α)≠ε(β)) все изоморфны — это можно проверить, найдя представление полугруппы, порожденной любым из них, и убедившись, что оно совпадает с представлением в примере (а). Это не случайность, а свойство широкого класса уравнений в словах: например, показано, что нетривиальное решение любого неоднородного уравнения с тремя неизвестными единственно с точностью до изоморфизма.

(в) Рассмотрим граф $G_\rho^{(N,n)}$ отношения эквивалентности ρ на

$$A^N \quad (|A| = n): \ (\alpha, \ \beta) \in \rho \Rightarrow C(\alpha) = C(\beta).$$

Согласно (33) $\alpha\left(G_\rho^{(N,n)}\right) = |A^N/\rho|$. Получим явное выражение для $\alpha(G_\rho^{(N,n)})$.

Пусть $P_n(N)$ — множество всех простых слов A^N. Тогда согласно п. 5 имеем

$$A^N = \bigcup_{d|N} \{\alpha^{N/d} \,|\, \alpha \in P_n(d)\} \tag{56}$$

(объединение по всем делителям d числа N, включая 1 и N). Так как каждое слово α представимо в виде $\alpha = \varepsilon(\alpha)^{\tau(\alpha)}$, где ε(α) — простое (примитивный корень), единственным способом, множества в правой части (56) попарно не пересекаются и мы имеем

$$|A^N| = n^N = \sum_{d|N} |P_n(d)|.$$

Отсюда, применяя теоретико-числовую формулу обращения арифметических функций (если где

$$G(N) = \sum_{d|N} F(d), \text{ то } F(N) = \sum_{d|N} \mu(d)\, G\left(\frac{N}{d}\right),$$

где функция Мёбиуса μ(N) вычисляется так: μ(1) = 1, если $p_1,.., p_k$ — различные простые, то $\mu(p_1,.., p_k)=(—1)^k$, $\mu(N)=0$ в остальных случаях), получаем

$$|P_n(N)| = \sum_{d|N} \mu(d)\, n^{N/d},$$

$$\alpha\left(G_\rho^{(N,n)}\right) = |A^N/\rho| = \sum_{d|N} \frac{1}{d} |P_n(d)| = \sum_{d|N} \frac{1}{d} \sum_{k|d} \mu(k)\, n^{d/k}.$$

Так, если N — простое число, то

$$\alpha\left(G_\rho^{(N,n)}\right) = \sum_{k|1} \mu(k)\, n^{1/k} + \frac{1}{N} \sum_{k|N} \mu(k)\, n^{N/k} =$$

$$= \mu(1)\cdot n + \frac{1}{N}\left(\mu(1)\cdot n^N + \mu(N)\cdot n\right) =$$

$$= n + \frac{1}{N}\left(n^N - n\right).$$

(г) Хотя массовая проблема перечисления всех экстремальных унтер-решений систем уравнений в словах алгоритмически неразрешима, для многих конкретных систем и даже для больших классов она может быть решена. Характер рассуждений, приводящих к решению, мы проиллюстрируем на примере системы уравнений с четырьмя неизвестными. Пусть $\rho = s^*(\rho_1)$, где ρ_1:

$$x_1x_2 = x_3x_1,$$
$$x_1x_4x_1 = x_3x_2$$

— система уравнений в словах над алфавитом {0, 1}.

Пусть $X = \langle x_1, x_2, x_3, x_4\rangle \in Y^0(\rho)$ и спектр длин его слов есть

$$D_X = \langle d_1, d_2, d_3, d_4\rangle, \ d_i = |x_i|, \ i = \overline{1, 4}.$$

Установим несколько необходимых условий.

1) Все слова X отличны от λ и попарно различны. Действительно, легко проверить, что слова длины два или меньше участвуют в ρ только в одном соотношении $x_1x_2 = x_3\,x_1$. Но если какое-либо из слов $x_i = \lambda$, то выполняется $x_i^2 = x_i$, а соотношение $x_i = x_j$ при $i \neq j$ тоже не входит в ρ. В частности, все $d_i > 0$, $i = 1, 2, 3, 4$.

2) Если $1 \notin D_x$, то $<2, 2, 2, 2> \leq D_x$ и в таком случав D_x может быть спектром экстремального унтер-решения только при $D_x = <2, 2, 2, 2>$ и, следовательно, $X = (00, 01, 10, 11>$, которое является префиксным кодом и действительно принадлежит множеству $Y(\rho)$. Таким образом, остается рассмотреть такие X, что $D_x \ni =1$.

3) Две единицы не могут входить в D_x, так как если $x_i = 0$, $x_j = 1$ и $k \neq i, j$, то x_k выражается через x_i и x_j.

4) Если $1 \in D_X$ и $2 \notin D_X$, то $<1, 2, 3, 3> <D_X$ в некотором порядке, следовательно, D_x не является спектром экстремального унтер-решения, так как $<1, 2, 3, 3>$ — спектр тупикового префиксного кода: $<1, 01, 001, 000> \in Y(\rho)$. Таким образом, кроме $<2, 2, 2, 2>$ и $<1, 2, 3, 3>$ нас могут интересовать только спектры, не мажорирующие эти, т. е. имеющие состав $<1, 2, 2, i>$, где $i>2$.

5) Тройки слов с длинами {1, 2, 2} могут быть <1, 00, 01>, <1, 00, 10>, <1, 01, 10> (тройки, которые получаются из них переобозначениями 0 и 1, отличаются от этих несущественно). Две первые из них — префиксный код и его обращение, третья удовлетворяет соотношению $ab = ca$ при $a \Rightarrow 1$, $b \Rightarrow 01$, $c \Rightarrow 10$. Учитывая 1), эта тройка может входить в унтер-решение только при $x_1 = a$, $x_2 = b$, $x_3 = c$.

Выясним, возможно ли это, т. е. можно ли подобрать слово x_4 так, чтобы $<1, 01, 10, x_4> \in Y(\rho)$. Если да, то найдем кратчайшее из таких слов и получим экстремальное унтер-решение. Словом длины $2x_4$ быть не может: 01 и 10 входят в X, $x_4 \Rightarrow 11$ влечет $x_4 = x^2_1$, а если $x_4 \Rightarrow 00$, то $x_3x_4x_1 = x_1x_4x_2$ и это соотношение не входит в ρ, так как к $x_3x_4x_1$ ни одна из подстановок ρ, неприменима. Аналогично обнаруживаем, что ни одно слово длины 3 не может быть взято в качестве x_4:

$$\begin{array}{llll} 000 - & x_1x_4x_2 = x_3x_4x_1, & 100 - & x_4x_1 = x_3x_2, \\ 001 - & x_1x_4 = x_3x_2, & 101 - & x_4 \quad = x_1x_2, \\ 010 - & x_1x_4 = x_3^2, & 110 - & x_4 \quad = x_1x_3, \\ 011 - & x_4 = x_2x_1, & 111 - & x_4 \quad = x_1^3. \end{array}$$

Среди слов длины 4 находим $x_4 \Rightarrow 0001$ такое, что с вхождением x_4=0001 никакое неразложимое соотношение невозможно и <1, 01, 10, 0001 > $\in$ $Y^0(\rho)$. Если d_1= 1, d_2 = d_3 = 2, то надо проверить еще, не существуют ли унтер-решения <1, 00, 01, x_4>, <1, 01, 00, x_4>, <1, 00, 10, x_4> или <1, 10, 00, x_4> с более коротким x_4. Но если X допускает какое-либо нетождественное соотношение, то, как следует из вида ρ_1, x_1 и x_3 должны быть в отношении префиксности, а x_1 и x_2 — в отношении суффиксности. Это не выполняется в каждом из перечисленных выше случаев, следовательно, ни одни ни них нереализуем, разве что можно найти такое x_4, что X имеет свойство единственности разложения на множители. Такое тоже невозможно, как легко убедиться непосредственной проверкой.

6) Перечислим остальные спектры, реализуемость которых подлежпг проверке:

$$\begin{array}{ll} \langle 1, 2, i, 2\rangle, & \langle 2, 2, 1, i\rangle, \\ \langle 1, i, 2, 2\rangle, & \langle 2, i, 1, 2\rangle, \\ \langle 2, 1, 2, i\rangle, & \langle i, 2, 1, 2\rangle, \\ \langle 2, 1, i, 2\rangle, & \langle 2, 2, i, 1\rangle, \\ \langle i, 1, 2, 2\rangle, & \langle 2, i, 2, 1\rangle, \\ & \langle i, 2, 2, 1\rangle. \end{array}$$

Аналогичными рассуждениями можно убедиться, что ни один из них нереализуем. Таким образом, матрица спектров экстремальных унтер-решений системы ρ состоит из <1, 2, 2, 4>, <2, 2, 2, 2> и всех перестановок <1, 2, 3, 3>.

Замечание. Используя пример (б), можно показать, что общее решение системы уравнений ρ_1 можно записать в виде $<(\alpha\beta)^i\alpha, \beta\alpha, \alpha\beta, \beta\beta>$, где α, β — произвольные слова, i — произвольное неотрицательное целое.

(д) Пусть каждому элементу $i \in E_n$ = {0, 1, ..., n — 1} сопоставлено действительное число $w(i)$, его «вес», и f = f_W — естественный гомоморфизм $(E_n)^+$ на подполугруппу W^+ коммутативной полугруппы действительных чисел по сложению:

$$f(i) = w(i) \text{ и } f(x_1, ..., x_N) = f(x_1)+...+ f(x_N).$$

Множество $R_{W,n} \Rightarrow f^{1}(0)$, если непусто, образует, очевидно, подполугруппу $(E_n)^+$.

Слово $\alpha \in R_{W,n}$ назовем неотрицательным, если для любого его префикса α' имеет место $f(\alpha') \geq 0$, и положительным, если для любого его префикса α имеет место $f(\alpha') > 0$. Пусть $R^0_{W,n}$ — подполугруппа всех неотрицательных слов $R_{W,n}$. Очевидно, что единственным неприводимым порождающим множеством $R^0_{W,n}$ (в данном случае — множеством всех слов, неразложимых в произведение) является бесконечное множество $R^1_{W,n}$ всех положительных слов из $R^0_{W,n}$, причем оно является множеством свободных образующих.

Вложение полугруппы $R^0_{W,n}$ в $R_{W,n}$ имеет важное свойство, устанавливаемое ниже.

Теорема 1. *Если* $\alpha = a_1a_2...a_N \in R_{W,n}$, *то существует такое* f, *что циклический сдвиг* $\alpha^{(j)} = a_ja_{j+1}... a_Na_1... a_{j-1}$ *содержится в* $R^0_{W,n}$.

Доказательство. Пусть $h_i \Rightarrow f(a_1 \ldots a_i) =$
$= w(a_1) + \ldots + w(a_i)$ и $h(\alpha) = \{h_i\}_{i=1}^N$. Тогда

$$h_i(\alpha^{(j)}) = \begin{cases} f(a_j \ldots a_{j+i-1}), & \text{если } j+i-1 \leqslant N, \\ f(a_j \ldots a_N a_1 \ldots a_{j+i-N-1}), & \text{если } j+i-1 > N, \end{cases}$$
$$= \begin{cases} f(a_1 \ldots a_{j+i-1}) - f(a_1 \ldots a_{j-1}), \\ f(a_1 \ldots a_N a_1 \ldots a_{i+j-N-1}) - f(a_1 \ldots a_{j-1}) \end{cases}$$
$$= h_{i+j-1} - h_{j-1}.$$

Отсюда

$$h(\alpha^{(j)}) = \{h_{i+j-1} - h_{j-1}\}_{i=1}^N,$$

причем первое слагаемое при $i = 1, 2, \ldots, N$ пробегает множество значений $h(\alpha)$. Таким образом, достаточно взять j так, чтобы $h_{j-1} = \min\limits_{1 \leq i \leq N} h_i$, и получим требуемый сдвиг. Теорема доказана.

Положим

$$f(i) = w(i) = 1 - i \cdot \frac{N}{k} \ (i = \overline{0, n-1}), \ \|\alpha\| \Rightarrow \sum_{i=1}^{N} a_i.$$

Тогда

$$f(\alpha) = |\alpha| - \|\alpha\| \cdot \frac{N}{k},$$

$$\alpha \in R_{W,n} \leftrightarrow |\alpha| - \|\alpha\| \cdot \frac{N}{k} = 0 \leftrightarrow \frac{|\alpha|}{\|\alpha\|} = \frac{N}{k}.$$

Возьмем $n = 3$, $N = 3$, $k = 2$, $\alpha = 101200110 \in R_{W,3}$. Имеем $h_1 = —1/2$, $h_2 = 1/2$, $h_3 = 0$, $h_4 = -2$, $h_5 = -1$, $h_6 = 0$, $h_7 = —1/2$, $h_8 = —1$, $h_9 = 0$ и для получения неотрицательного сдвига надо взять $j = 5$: $\alpha^{(5)} = 001101012$ и для него $h_1 = 1$, $h_2 = 2$, $h_3 = 3/2$, $h_4 = 1$, $h_5=2$, $h_6 = 3/2$, $h_7 = 5/2$, $h_8 = 2$, $h_9 = 0$.

Теорема 2. Если *НОД* $(|\alpha|, \|\alpha\|) = 1$ *и* $\alpha \in R_{W,n}$, *то в циклоклассе* $C(\alpha)$ *существует единственное неотрицательное слово и оно положительно.*

Доказательство. Если $\alpha^{(j)} = \alpha_1\alpha_2$ — неотрицательный сдвиг α, то из $\alpha_1, \alpha_2 \in R_{W,n}$ следует (в предположении, что $\frac{N}{k}$ — несократимая дробь), что

$$\frac{|\alpha_1|}{\|\alpha_1\|} = \frac{N}{k} = \frac{|\alpha_2|}{\|\alpha_2\|}, \quad |\alpha_i| = p_i N, \quad \|\alpha_i\| = p_i k \quad (i = 1, 2).$$

Тогда $p_i \geqslant 1, \ |\alpha| = (p_1 + p_2)N, \ \|\alpha\| = (p_1 + p_2)k$ — противоречие и теорема доказана.

10.6. Регулярные множества слов

1. Одним из элементов модели языка является описание правил построения слов, грамматики. Теория *регулярных множеств* составляет наиболее завершенный раздел развитых исследований в области математического моделирования языков, рассматриваемых как подмножества свободных полугрупп над конечными алфавитами. Регулярные языки — наимеиьший класс языков, содержащий все конечные языки и замкнутый относительно основных комбинаторных операций над языками.

2. Основной способ задания грамматик регулярных множеств — графический. Пусть $G = \langle Q, R\rangle$ — конечный граф (как правило — ориентированный, может быть с кратными ребрами), $F \in (A^+)^R$ — функция, сопоставляющая ребрам G слова из A^+ (функция переходов), P — некоторое множество путей в графе G. Система $\Gamma = \langle G, F, P\rangle$ называется *источником,* $G = G(\Gamma)$, $F = F(\Gamma)$ и $P = P(\Gamma)$ — элементы источника Γ. Вершины графа G обычно называют *состояниями* источника Γ.

Говорят, что путь $p = r_1, \ldots r_k$ в графе G порождает слово $F(p) = F(r_1) \ldots F(r_k)$. Считается, что пустой путь Λ, не содержащий ни одного ребра, начинается и кончается в каждой вершине и порождает пустое слово λ. Множество всех слов, которые порождаются путями из $P,$ т. е.

$$\mathfrak{L}(\Gamma) \Rightarrow \{\alpha \mid \exists\, p \in P \quad F(p) = \alpha\},$$

называется *языком, порожденным источником* Г.

Пусть $Q_1 \subseteq Q$, $Q_2 \subseteq Q$ и $P(Q_1, Q_2)$ — множество всех путей в *G*, которые начинаются в вершинах Q_1 и кончаются в вершинах Q_2. Источник называется *регулярным,* если в нем выделены $q_0 \in Q$, $Q' \subseteq Q$ и множество путей определено как $P = P(q_0, Q')$.

Множество слов $\mathfrak{L} \subseteq A^*$ называется *регулярным,* если $\mathfrak{L} = \mathfrak{L}(\Gamma)$ для некоторого регулярного источника Г. Регулярное множество $\mathfrak{L} \subseteq A^*$ называется *конечно перечислимым,* если оно порождается некоторым регулярным источником Г таким, что $P_\Gamma = P(q_0, Q)$ (т. е. у которого множеством заключительных состояний *Q'* является множество всех состояний).

3. Имеется тесная связь между регулярными источниками и *конечными автоматами.* Для источника Г через $\varphi_\Gamma(a, q)$ обозначим (вообще говоря, многозначное и не полностью определенное) отображение множества $A \times Q$ в 2^Q:

$$\varphi_\Gamma(a, q) = \{q' \mid \exists r = \langle q, q' \rangle \in R \quad F(r) = a\},$$

называемое *функцией следования*. Функция φ_Γ естественно доопределяется для любых $\alpha \in A^*$ и $Q_1 \subseteq Q$:

$$\varphi_\Gamma(\lambda, Q_1) = Q_1,$$
$$\varphi_\Gamma(\alpha, q) = \{q' \mid \exists p \in P(q, q') \quad F(p) = \alpha\},$$
$$\varphi_\Gamma(\alpha, Q_1) = \bigcup_{q \in Q_1} \varphi_\Gamma(\alpha, q).$$

Источник называется *простым,* если $F \in A^n$, т. е. значениями *F* являются для всех ребер только буквы. Для простого источника это доопределение можно задать индуктивно по длине слов α*:* если для $\alpha \in A$ множества $\varphi_\Gamma(\alpha, Q_1)$ заданы, то для $\alpha = \alpha_1\alpha_2$ имеет место

$$\varphi_\Gamma(\alpha_1\alpha_2, Q_1) = \varphi_\Gamma(\alpha_2, \varphi_\Gamma(\alpha_1, Q_1)). \qquad (57)$$

Регулярный источник Г, для которого $|\varphi_\Gamma(a, q)| \le 1$ при любых $a \in A$ и $q \in Q$, называется *автоматным* или, для краткости, *А-источником,* поскольку при этом условии Г представляет собой диаграмму Мура конечного автомата $\langle A, Q, \varphi_\Gamma, q_0, — \rangle$ без выхода, с входным алфавитом *A* и множеством состояний *Q*, у которого q_0 — начальное состояние, а φ_Γ — функция следующего состояния. Соответствующий автомат полностью определен, если $|\varphi_\Gamma(a, q)| = 1$ при любых $a \in A$ и $q \in Q$, в противном случае имеем частичный автомат. Легко заметить, что область определения частичного автомата всегда представляет собой конечно перечислимое множество. Полезно иметь в виду важное свойство полностью определенных *А* -источников: *множество путей в* *графе такого источника находится во взаимно однозначном соответствии с множеством слов A*.*

Регулярный источник может быть задан графом, ребрам которого приписаны значения функции переходов F, либо таблицей функции F (начальное и заключительные состояния как-либо помечаются). В других случаих полезен *индуктивный способ задания,* особенно для A-источников. Индуктивный способ состоит в задании начального состояния q_0 и *функции следования* некоторым правилом. При этом множество Q состояний определяется как множество всех состояний, достижимых из начального, а заключительные состоянии выделяются из Q некоторым свойством. Например, условия $q_0 = 0$, $A = E_2 = \{0, 1\}$, $\varphi_\Gamma\ (a, q) = a + q + 1(\mathrm{mod} 4)$ и $Q' = \{q | q \equiv 3(\mathrm{mod}\ 4)\}$ задают регулярный источник, представленный также таблицей 3 и графически — на рис. 6.

Таблица 3

A	Q	φ_Γ	A	Q	φ_Γ
0	0	1	1	0	2
0	1	2	1	1	3
0	2	3	1	2	0
0	3	0	1	3	1

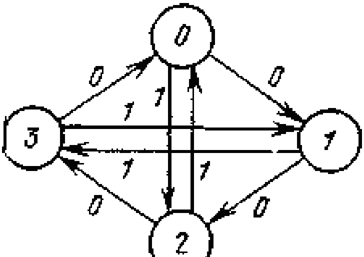


Рис 6.

Недостатком такого способа задания регулярного источника может оказаться необходимость обосновывать конечность множества его состояний.

4. Источники Γ_1 и Γ_2 называются *эквивалентными,* если $\mathfrak{L}(\Gamma_1) = \mathfrak{L}(\Gamma_2)$, т. е. если они порождают один и тот же язык

(источник Г называется *сокращенным,* если через любое из его состояний проходит по крайней мере один путь из $P(q_0, Q')$. Вполне очевидно, что *любой источник эквивалентен некоторому сокращенному источнику*). Одно пз эквивалентных преобразований источников показано на рис. 7. Оно применяется к ребру, которому приписано слово, разложимое в произведение αβ, и состоит в том, что вводится новая вершина q' и данное ребро заменяется путем из двух ребер, причем первому приписано слово α, а второму — β.

Рис. 7.

Конечным числом таких преобразований произвольный источник можно преобразовать к простому источнику.

Теорема 1. *Для любого регулярного источника существует эквивалентный ему простой регулярный источник.*

Теорема 2. *Для любого регулярного источника существует эквивалентный ему полностью определенный А-источник.*

Доказательство. Пусть Г = $<G, F, P>$ — произвольный простой регулярный источник. Полностью определенный A-источник $\Gamma_1 = <G_1, F_1, P_1>$ зададим индуктивно:

$$Q_1 \equiv \{q_M \mid M \subseteq Q\}, \quad P_1 = P\ (q_{\{q_0\}}, \{q_M \mid M \cap$$
$$\cap\, Q' \neq \varnothing\}), \quad \varphi_{\Gamma_1}(a, q_M) = q_{\varphi_\Gamma(a, M)}.$$

Покажем, что для любого $\alpha \in A^*$ имеет место

$$\varphi_{\Gamma_1}(\alpha, q_{\{q_0\}}) = q_{\varphi_\Gamma(\alpha, q_0)}. \tag{58}$$

Индукция по длине α. Если $|\alpha| \leq 1$, то (58) выполняется по определению. Пусть (58) справедливо для слов длины, меньшей k, и $|\alpha| = k > 1:\ \alpha = \beta a$, где $a \in A$, $|\beta| < k$. Тогда, используя (57), получаем

$$\varphi_{\Gamma_1}(\alpha, q_{\{q_0\}}) = \varphi_{\Gamma_1}(\alpha, \varphi_{\Gamma_1}(\beta, q_{\{q_0\}})) = \varphi_{\Gamma_1}(a, q_{\varphi_\Gamma(\beta, q_0)}) =$$
$$= q_{\varphi_\Gamma(a, q_{\varphi_\Gamma}(\beta, q_0))} = q_{\varphi_\Gamma(\beta a, q_0)} = q_{\varphi_\Gamma(\alpha, q_0)}$$

и (58) доказано. Но мы имеем

$$\alpha \in \mathfrak{L}(\Gamma) \leftrightarrow \varphi_\Gamma(\alpha, q_0) \cap Q' \neq \varnothing \leftrightarrow \varphi_{\Gamma_1}(\alpha, q_{\{q_0\}}) =$$
$$= q_{\varphi_\Gamma(\alpha, q_0)} \in Q_1' \leftrightarrow \alpha \in \mathfrak{L}(\Gamma_1)$$

и, следовательно, $\mathfrak{L}(\Gamma_1) = \mathfrak{L}(\Gamma)$. Теорема доказана.

Если число состояний Г равно N, то число состояний эквивалентного ему A-источника Γ_1, который строится в доказательстве теоремы 2, не превосходит 2^N, но может таким и оказаться. Возникает вопрос: не существует ли более экономное построение? Как показал Лупанов, *для любых N и $|A| \geq 2$ существуют регулярные источники с N состояниями, для которых 2^N — наименьшее возможное число состояний эквивалентного полностью определенного A-источника.* При $N = 3$, $A = \{a, b, c\}$ источник Лупанова и эквивалентный ему полностью определенный A-источник показаны на рис. 8, *а), б).*

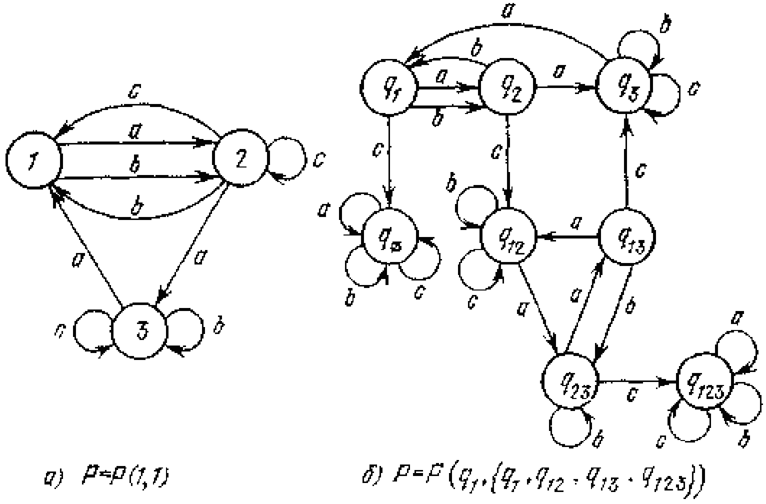


Рис. 8.

Рассмотрим вопрос об упрощении конечных автоматов в отношении уменьшения числа состояний. Покажем, что *исход минимизации определен однозначно с точностью до обозначения состояний* (т. е. с точностью до изоморфизма).

Пусть $\mathfrak{L} \subseteq A^*$ — регулярное множество и $\rho_{\mathfrak{L}}$ — отношение эквивалентности на A^*: $\langle \alpha, \beta \rangle \in \rho_{\mathfrak{L}} \Rightarrow$ «для любого $\gamma \in A^*$: $\alpha\gamma \in \mathfrak{L} \leftrightarrow \beta\gamma \in \mathfrak{L}$», а $A^*/\rho_{\mathfrak{L}} = \{\mathfrak{A}_0, \mathfrak{A}_1, \ldots \mathfrak{A}_i, \ldots\}$, где $\lambda \in \mathfrak{A}_0$. Заметим, что любой класс $\mathfrak{A}_i$ либо целиком содержится в $\mathfrak{L}$, либо не пересекается с $\mathfrak{L}$.

Пусть Г — регулярный полностью определенный A-источник и $\mathfrak{L}(\Gamma) = \mathfrak{L}$. Определим отношение эквивалентности ρ_Γ:

$$\langle \alpha, \beta \rangle \in \rho_\Gamma \Rightarrow \varphi_\Gamma(\alpha, q_0) = \varphi_\Gamma(\beta, q_0).$$

Очевидно, что классы $A^*/\rho_\Gamma = \{\mathfrak{B}_0, \mathfrak{B}_1, \ldots, \mathfrak{B}_{n-1}\}$ находятся во взаимно однозначном соответствии с состояниями $Q = \{q_0, q_1, \ldots, q_{n-1}\}$ и можно считать, что $q_i \Rightarrow \mathfrak{B}_i, \lambda \in \mathfrak{B}_0$.

Пусть $i_\mathfrak{L}(\alpha)$ — номер класса $\rho_\mathfrak{L}$, содержащего α, $i_\Gamma(\alpha)$ — номер класса ρ_Γ, содержащего α. Так как из $i(\alpha) = i(\beta)$ следует $i(\alpha\gamma) = i(\beta\gamma)$ для любого $\gamma \in A^*$, то существует функция $j(i,\alpha)$ такая, что для любых α, β

$$i(\alpha\beta) = j(i(\alpha), \beta). \tag{59}$$

Докажем, что

$$\rho_\Gamma \subseteq \rho_\mathfrak{L}. \tag{60}$$

Если $\langle\alpha, \beta\rangle \in \rho_\Gamma$, т. е. $\varphi_\Gamma(\alpha, q_0) = \varphi_\Gamma(\beta, q_0) = q_s$, то для любого γ имеем

$$\varphi_\Gamma(\alpha\gamma, q_0) = q_{j(i_\Gamma(\alpha),\gamma)} = q_{j(i_\Gamma(\beta),\gamma)} = \varphi_\Gamma(\beta\gamma, q_0),$$

$$\alpha\gamma \in \mathfrak{L} \leftrightarrow q_{j(i_\Gamma(\alpha),\gamma)} \in Q' \leftrightarrow \beta\gamma \in \mathfrak{L},$$

т. е. $\langle\alpha, \beta\rangle \in \rho_\mathfrak{L}$ и (60) доказано.

Из (60) следует, что $|A^*/\rho_\mathfrak{L}| \leqslant |A^*/\rho_\Gamma| = |Q| < \infty$, т. е. число классов $\rho_\mathfrak{L}$ конечно и каждый класс $\rho_\mathfrak{L}$ является объединением некоторых классов ρ_Γ, а так же, что если $N_\mathfrak{L}$ — наименьшее число состояний полностью определенного *А*-источиика, порождающего $\mathfrak{L}$, то $N_\mathfrak{L} \geqslant |A^*/\rho_\mathfrak{L}|$.

Покажем, что в действительности $N_\mathfrak{L} = |A^*/\rho_\mathfrak{L}|$. Для этого построим источник $\Gamma_\mathfrak{L}$ с $n = |A^*/\rho_\mathfrak{L}|$ состояниями такой, что $\mathfrak{L}(\Gamma_\mathfrak{L}) = \mathfrak{L}$. Пусть Q — множество номеров классов $A^*/\rho_\mathfrak{L}$ и $0 = i(\lambda), \varphi_{\Gamma\mathfrak{L}}(a, i) = j(i, a)$, Q' — множество всех номеров классов, содержащихся в $\mathfrak{L}$. Мы имеем $\varphi_{\Gamma\mathfrak{L}}(\alpha, 0) = i_\mathfrak{L}(\alpha)$, следовательно, $\mathfrak{L}(\Gamma_\mathfrak{L}) = \mathfrak{L}$ и $|Q| = N_\mathfrak{L}$. Итог подводит

Теорема 3. *Множество $\mathfrak{L}$ регулярно в том и только в том случае, если $A^*/\rho_\mathfrak{L}$ конечно. Если $\mathfrak{L}(\Gamma) = \mathfrak{L}$ и* Г *— полностью определенный А-источник, то $\Gamma_\mathfrak{L}$ получается из* Г *факторизацией множества состояний по отношению эквивалентности, соответствующему $\rho_\mathfrak{L}$. Если число состояний* Г *равното $N_\mathfrak{L}$,* Г *совпадает с $\Gamma_\mathfrak{L}$ с точностью до обозначения состояний.*

Алгоритмические вопросы минимизации *А*-источников решаются теми же методами, что и для конечных автоматов. (Пусть $Q(\Gamma_\mathfrak{L}) = Q(\Gamma)/\rho$ и $(q_i, q_j) \in \rho_s \Rightarrow$ «для любого α, длина которого не превосходит s, $\varphi_\Gamma(\alpha, q_i) \in Q'(\Gamma) \leftrightarrow \varphi_\Gamma(\alpha, q_j) \in Q'(\Gamma)$».

Тогда $\rho_0 \supseteq \rho_1 \supseteq \ldots$ и имеет место теорема:

$$\rho_{N_\rho - 1} = \rho.$$

).

Если $|A| = 1$, то, как следует из теоремы 2, любой регулярный источник эквивалентен A-источнику, представляющему собой граф преобразования множества Q (рис. 9).

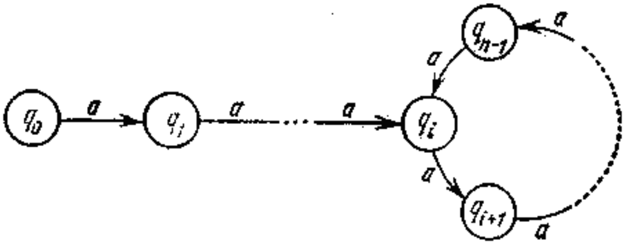


Рис. 9.

Слова A^* в этом случае можно рассматривать как коды натуральных чисел: $a^i \Rightarrow i$. Если $P = P(q_0, q_j)$, то

$$\mathfrak{L}(\Gamma) = \begin{cases} \{j\}, & \text{если } j < i, \\ \{j + k(n-i)\}_{k=0}^{\infty} & \text{в противном случае.} \end{cases}$$

Поэтому множество натуральных чисел при таком прямолинейном кодировании регулярно в том и только том случае, если оно есть объединение конечного множества и некоторого (конечного) числа арифметических прогрессий с одинаковой разностью.

Используя этот факт, легко указать примеры нерегулярных множеств. Так, $\left\{a^{i^2} \mid i = 0, 1, 2, \ldots\right\}$ не регулярно — оно бесконечно и разность между соседними по возрастанию членами может быть как угодно велика.

Непосредственной проверкой можно убедиться, что

$$\rho_{A^* \setminus \mathfrak{L}} = \rho_{\mathfrak{L}}, \tag{61}$$

$$\rho_{\mathfrak{L}_1 \cap \mathfrak{L}_2} \supseteq \rho_{\mathfrak{L}_1} \cap \rho_{\mathfrak{L}_2}. \tag{62}$$

5. Здесь покажем, что некоторые естественные операции над регулярными языками приводят снова к регулярным языкам.

Из (61), (62) следует, что если $A^*/\rho_{\mathfrak{L}}$, $A^*/\rho_{\mathfrak{L}_1}$, $A^*/\rho_{\mathfrak{L}_2}$ конечны, то и $A^*/\rho_{\overline{\mathfrak{L}}}$ и $A^*/\rho_{\mathfrak{L}_1 \cap \mathfrak{L}_2}$ тоже конечны. Учитывая теорему 2.4 и то, что

$\mathfrak{L}_1 \cup \mathfrak{L}_2 = \overline{\overline{\mathfrak{L}}_1 \cap \overline{\mathfrak{L}}_2}$, делаем первый вывод.

Теорема 1. *Класс регулярных языков замкнут относительно объединения, пересечения и теоретико-множественной разности.*

Замечание. Если $\lambda \in \mathfrak{L}$ и $\mathfrak{L}$ регулярно, то и $\mathfrak{L}\backslash\{\lambda\}$ регулярно; для любого регулярного $\mathfrak{L}$ множество $\mathfrak{L} \cup \lambda\}$ тоже регулярно.

Произведение языков $\mathfrak{L}_1$ и $\mathfrak{L}_2$ есть

$$\mathfrak{L}_1 \cdot \mathfrak{L}_2 \Rightarrow \{\alpha\beta | \alpha \in \mathfrak{L}_1, \beta \in \mathfrak{L}_2\}.$$

Теорема 2. *Класс регулярных языков замкнут относительно операции произведения.*

Доказательство. Пусть $\mathfrak{L}_1 = \mathfrak{L}(\Gamma_1)$, $\mathfrak{L}_2 = \mathfrak{L}(\Gamma_2)$ регулярны,

$$Q_1 = \{q_{10}, q_{11}, \ldots, q_{1n}\}, \quad Q_2 = \{q_{20}, q_{21}, \ldots, q_{2m}\}.$$

Построим Г такой, что $\mathfrak{L}(\Gamma) = \mathfrak{L}_1 \cdot \mathfrak{L}_2$ (достаточно это сделать в предположении, что $\lambda \notin \mathfrak{L}_1$ и $\lambda \notin \mathfrak{L}_2$, учитывая теорему 1, последующее замечание и очевидные соотношения $(\mathfrak{L}_1 \cup \{\lambda\}) \cdot \mathfrak{L}_2 = \mathfrak{L}_1 \cdot \mathfrak{L}_2 \cup \mathfrak{L}_2$.., и $\mathfrak{L}_1 \cdot (\mathfrak{L}_2 \cup \{\lambda\}) = \mathfrak{L}_1 \cdot \mathfrak{L}_2 \cup \mathfrak{L}_1$.

Возьмем $Q = Q_1 \cup Q_2$, $P = P\ (q_{10}, Q_2')$ и положим

$$\varphi_\Gamma(a, q) =$$

$$= \begin{cases} \varphi_{\Gamma_1}(a, q), & \text{если } q \in Q_1 \text{ и } \varphi_{\Gamma_1}(a, q) \cap Q_1' = \varnothing, \\ \varphi_{\Gamma_1}(a, q) \cup \{q_{20}\}, & \text{если } q \in Q_1 \text{ и } \varphi_{\Gamma_1}(a, q) \cap Q_1' \neq \varnothing, \\ \varphi_{\Gamma_2}(a, q), & \text{если } q \in Q_2. \end{cases}$$

Тогда Г — искомый регулярный источник, в чем убеждаемся непосредственной проверкой. Теорема доказана. Итерацией языка $\mathfrak{L}$ называется язык $\mathfrak{L}^* = \mathfrak{L}^+ \cup \{\lambda\}$, где $\mathfrak{L}^+ \Rightarrow \mathfrak{L} \cup \mathfrak{L}^2 \cup \cdots = \bigcup\limits_{i=1}^{\infty} \mathfrak{L}^i$.

Теорема 3. *Класс регулярных языков замкнут относительно операции итерации.*

Доказательство. Пусть $\mathfrak{L} = \mathfrak{L}(\Gamma)$. Учитывая теорему 1 и последующее замечание, можно предположить, что $\lambda \notin \mathfrak{L}$, и доказать регулярность $\mathfrak{L}^+$. Определим Γ_1: $Q_1 = Q$, $P_1 = P(q_0, Q')$,

$$\varphi_{\Gamma_1}(a, q) = \begin{cases} \varphi_\Gamma(a, q) & \text{при } \varphi_\Gamma(a, q) \cap Q' = \varnothing, \\ \varphi_\Gamma(a, q) \cup \{q_0\} & \text{в противном случае.} \end{cases}$$

Непосредственной проверкой убеждаемся, что $\mathfrak{L}(\Gamma_1) = \mathfrak{L}^+$. Теорема доказана (Легко проверить справедливость следующих тождеств:
$X(YZ) = (XY)Z$, $X(Y \cup Z) = XY \cup XZ$, $(X \cup Y)Z = XZ \cup YZ$,
$X^+ = X \cdot X^* = X^* \cdot X$ и т. д.).

6. Рассмотрим вопрос о решении некоторых уравнений в алгебре регулярных множеств.

Теорема 1. *Пусть $\lambda \notin S$, тогда уравнение*

$$X = XS \cup T \qquad (63)$$

имеет единственное решение $X = TS^*$ (в частности, $X = \varnothing$ при $T = \varnothing$, $X = S^*$ при $T = \{\lambda\}$, $X = T$ при *$S = \varnothing$).*

Доказательство. Пусть X_0 — решение (1.278), т. е. $X_0 = X_0S \cup T$. Подставляя это выражение для X_0 в правую часть (63) и применяя тождественные преобразований, получим

$$X_0 = (X_0S \cup T)S \cup T = X_0S^2 \cup TS \cup T. \qquad (64)$$

Следующую подстановку делаем в (64) и т. д. После k таких подстановок получим

$$X_0 = (X_0S \cup T)S^k \cup TS^{k-1} \cup \ldots \cup TS \cup T =$$
$$= X_0S^k \cup TS^k \cup TS^{k-1} \cup \ldots \cup TS \cup T.$$

Поэтому $T\left(\bigcup_{i=0}^{k} S^i\right) \subseteq X_0$ для любого k и, следовательно, $TS^* \subseteq X_0$.

Пусть

$$d(S) \Rightarrow \min_{\alpha \in S} |\alpha| \, (d(S) > 0,$$

так как $\lambda \notin S$ и $d(\varnothing) \Rightarrow \infty$. Для любого $\alpha \in A^*$ найдется $k = k(\alpha)$ такое, что $|\alpha| < (k + 1)d(S)$. Поэтому, если $\alpha \in X_0$, то $\alpha \notin X_0S^i$ при $i > k(\alpha) + 1$ и, следовательно, $\alpha \in T\left(\bigcup_{i=0}^{k} S^i\right) \subseteq TS^*$, т. е. $X_0 = TS^*$. Таким образом, если (63) имеет решение, то им может быть только $X_0 = TS^*$. Делаем проверку:

$$(TS^*)S \cup T = TS^+ \cup TS^0 = T(S^+ \cup S^0) = TS^*.$$

Теорема доказана.

Следствие 1. *Если $X = XS \cup T$, S и T регулярны, $\lambda \notin S$, то и X регулярно.*

Следствие 2. *Пусть $\|S_{ij}\|$ есть $N \times N$-матрица, элементы которой суть конечные подмножества A^+, не содержащие λ. Тогда система уравнений*

$$\begin{cases} X_0 = X_0S_{00} \cup X_1S_{10} \cup \cdots \cup X_{N-1}S_{N-1,0} \cup \{\lambda\}, \\ X_1 = X_0S_{01} \cup X_1S_{11} \cup \cdots \cup X_{N-1}S_{N-1,1}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ X_{N-1} = X_0S_{0,N-1} \cup X_1S_{1,N-1} \cup \cdots \cup X_{N-1}S_{N-1,N-1} \end{cases} \qquad (65)$$

*имеет единственное решение, которое может быть записано в виде формул $X_i = \Phi_i(S_{00}, S_{01}, \ldots, S_{N-1,N-1})$, $i = \overline{0, N-1}$, алгебры регулярных множеств над операциями $\cup$, $\cap$, —, •, *.*

Это легко доказать, используя исключение неизвестных по теореме 1 и индукцию по N.

В силу теорем 1 — 3 в п. 5, формула $Y = \Phi(X_1,..., X_n)$ алгебры регулярных множеств с операциями $\cup$, $\cap$, —, •, * над конечными множествами $X_1,...,X_n \subseteq A^*$ задает регулярное множество Y. Справедливо и обратное.

Теорема 2. *Всякое регулярное множество может быть задано формулой в* $\cup$, $\cap$, —, •, * *над конечными множествами слов.*

Доказательство. Пусть регулярное множество $\mathfrak{L}$ задано A-источннком Г с множеством вершин $Q = \{q_0, q_1, \dots, q_{N-1}\}$. Положим

$$X_i \Rightarrow \mathfrak{L}(\langle G, F, P(q_0, q_i)\rangle) \qquad i = \overline{0, N-1},$$
$$S_{ij} \Rightarrow \{a \mid a \in A, \exists r = \langle q_i, q_j\rangle \in R \qquad F(r) = a\},$$
$$i, j = \overline{0, N-1}.$$

Тогда X_i и S_{ij} связаны системой уравнений (65), причем выполнены условия следствия 2 теоремы 1. Следовательно,

$$X_i = \Phi_i(S_{00}, S_{01}, \dots, S_{N-1, N-1}), \qquad i = \overline{0, N-1},$$

и $\mathfrak{L} = \bigcup\limits_{q_i \in Q'} X_i$, ч.т.д.

Следствие (теорема Клини). *Класс регулярных множеств слов в алфавите A есть замыкание множества* $\{\{a\}, ..., \{a_n\}, \{\lambda\}\}$ *относительно операций объединения, умножения и итерации.*

7. Учитывая, что источник, порождающий конечное множество слов, строится тривиально, переходы от одного способа задания регулярного множества к другим из упомянутых выше можно проиллюстрировать следующей диаграммой:

8. Регулярный язык называется *связным,* если существует порождающий его регулярный источник, граф которого связен (т. е. из любой его вершины в любую другую есть ориентированный путь).

Рассмотрим A-источиики с N состояниями и алфавитом $A = \{a_1, ..., a_n\}$ из m букв ((N, m)-источники).

Теорема 1. *Число* (N, m)*-источников равно*

$$(N+1)^{Nm}.$$

Доказательство следует из взаимно однозначного соответствия между множеством (*N*, *m*)-источников и множеством полностью определенных функций *(Q* $\cup\{0\}^{A\times Q}$ (у функций этого мпожества значение 0 показывает, что соответствующая функция следования при данных значениях переменных не определена).

Теорема 2. *Доля связных (N, m)-источников в общем числе (N, m)-источников стремится к* 1 *при m→∞.*

Доказательство. Несвязные *(N, m)-источники* — это в точности те источники, у которых имеется *l* состояний $(l=\overline{1,\ N-1})$, из которых нет ни одного ребра в остальные *N* — *l* состояний. Поэтому число несвязных *(N, m*)-источпиков не превосходит

$$\sum_{l=1}^{N-1}\binom{N}{l}(N+1)^{mN-ml}(l+1)^{ml}=$$

$$=\left\{\sum_{l=1}^{N-1}\binom{N}{l}\left(\frac{l+1}{N+1}\right)^{ml}\right\}(N+1)^{mN}.$$

Отсюда с учетом теоремы 1 следует, что их доля в общем числе *(N, m*)-источников

$$\delta_m\leqslant\sum_{l=1}^{N-1}\binom{N}{l}\left(\left(\frac{l+1}{N+1}\right)^{l}\ \right)^{m},$$

а так как $\binom{N}{l}<2^N$ и $\left(\frac{l+1}{N+1}\right)^{l}\leqslant\left(\frac{N}{N+1}\right)^{N-1}\leqslant\Big(1-$ $-\frac{1}{N+1}\Big)^{N+1}<\frac{1}{e}$, имеем $\delta_m<2^N(N-1)e^{-m}$, т. е. $\lim\limits_{m\to\infty}\delta_m=0$, ч.т.д.

9. Один подкласс регулярных языков — словесные представления $\mathfrak{F}$-полугрупп — мы рассматривали ранее. Это в точности языки, допускающие формульное задание в виде *V**, где *V* — конечное множество слов. Здесь рассмотрим еще один важный подкласс — *фрагментно ограниченные языки.*

Пусть $\mathfrak{L}_1$, $\mathfrak{L}_2$ — регулярные языки и $\mathfrak{L}_1(\mathfrak{L}_2)$ — множество всех слов $\mathfrak{L}_1$, которые не содержат вхождений слов из $\mathfrak{L}_2$. (Если слово α входит в слово β, т. е. $\beta=\gamma_1\alpha\gamma_2$, то α называется *фрагментом* β *(префиксом, если* $\gamma_1=\lambda$, *суффиксом,* если $\gamma_2=\lambda$, и само β называют *несобственными фрагментами)).*

Так как $\mathfrak{L}_1(\mathfrak{L}_2)=\mathfrak{L}_1\setminus A^*\mathfrak{L}_2A^*$, в силу теорем 1—3 из п. 5 язык $\mathfrak{L}_1(\mathfrak{L}_2)$ тоже регулярен, причем из доказательства этих теорем можно извлечь эффективный способ построения порождающего источника.

Особое значение имеют фрагментно ограниченные языки с конечными множествами ограничений $\mathfrak{L}_2$, как наиболее легкий аппарат построения приближенных моделей языков. Простейшими из таких языков являются *диграммно ограниченные языки,* для которых $\mathfrak{L}_2$ есть множество *диграмм,* т. е. слов длины два.

Установим некоторые свойства фрагментно-ограниченных языков.

Во-первых, заметим, что если α есть фрагмент β, то для любых $\mathfrak{L}_1$, $\mathfrak{L}_2$ имеет место $\mathfrak{L}_1(\mathfrak{L}_2, \alpha, \beta) = \mathfrak{L}_1(\mathfrak{L}_2, \alpha)$.

Поэтому, если для языка $\mathfrak{L}$ определить его фрагментно свободное ограничение $\mathfrak{L}^\varphi$ как множество всех слов $\mathfrak{L}$, которые не содержат в качестве фрагментов других слов $\mathfrak{L}$, то для любых языков $\mathfrak{L}_1$ и $\mathfrak{L}_2$ имеем $\mathfrak{L}_1(\mathfrak{L}_2) = \mathfrak{L}_1(\mathfrak{L}_2^\varphi)$. Следовательно, не ограничивая общности, всегда можно предполагать определяющее множество ограничений для фрагментно ограниченного языка фрагментно свободным. Например, если $\mathfrak{L} = \{0^i\}_{i=0}^{\infty} \cup \{0^i10^j\}_{i,j=0}^{\infty}$ — множество всех двоичных слов, вес которых не превосходит 1, то $\mathfrak{L} = E_2^*(\mathfrak{L}_1) = E_2^*(\mathfrak{L}_1^\varphi)$, где $\mathfrak{L}_1$ — множество всех слов, вес которых не менее 2, а $\mathfrak{L}_1^\varphi = \{11, 101, 1001, \ldots, 10^i1, \ldots\}$.

Во-вторых, очевидно, что для любого множества индексов J и языков $\mathfrak{L}_i \subseteq A^*$, $i \in J$:

$$\bigcap_{i \in J} A^*(\mathfrak{L}_i) = A^*\Big(\bigcup_{i \in J} \mathfrak{L}_i\Big).$$

Отсюда следует, что для любого языка $\mathfrak{L} \subseteq A^*$ среди фрагментно ограниченных языков, содержащих $\mathfrak{L}$, имеется наименьший (их пересечение), обозначаемый $\Phi(\mathfrak{L})$. Если $\mathfrak{L} \subseteq A^*$ и $\Phi(\mathfrak{L}) = A^*$, то $\mathfrak{L}$ называется *фрагментным покрытием* A^*.

Теорема 1. *Если* $\alpha \in A^+$ *и* $\mathfrak{L} \subseteq A^*(\alpha)$, *то*

$$\lim_{k \to \infty} \frac{|\mathfrak{L} \cap A^k|}{|A^k|} = 0.$$

Доказательство. Пусть $|\alpha| = n$, $|A| = m$. Очевидно, что

$$\mathfrak{L} \cap A^k = \mathfrak{L} \cap [(A^n \setminus \{\alpha\})^{[k/n]} \cdot A^{k-n[k/n]}].$$

Поэтому

$$\frac{|\mathfrak{L} \cap A^k|}{|A^k|} \leqslant \frac{(m^n - 1)^{[k/n]} m^{k-n[k/n]}}{m^k} = \left(\frac{m^n - 1}{m^n}\right)^{[k/n]} \to 0$$

при $k \to \infty$, ч. т. д.

Теорема 2. *Пусть* $\mathfrak{L}$ *— регулярное фрагментое покрытие* A^*. *Тогда существует слово* α *такое, что для любого слова* β *найдется* γ *так, что* $\alpha\beta\gamma \in \mathfrak{L}$ (*иначе:* $\alpha A^* \subseteq \pi(\mathfrak{L})$, *где* $\pi(\mathfrak{L})$ *— множество всех префиксовслов* $\mathfrak{L}$).

Доказательство. Предположим, что такого α, как в заключении теоремы, не существует. Пусть Г — сокращенный *А*-источиик, порождающий $\mathfrak{L}$, $Q = \{q_0, q_1, \ldots, q_{N-1}\}$ — множество его состояний. Тогда для каждого *i* = 0, 1, 2,…,*N* — 1 существует слово β_i такое, что $\varphi_\Gamma(\beta_i, q_i) = \varnothing$.

Положим

$$\varphi_\Gamma(\beta_0, q_1) = q_{i(1)},$$
$$\varphi_\Gamma(\beta_0\beta_{i(1)}, q_2) = q_{i(2)},$$
$$\cdots\cdots\cdots\cdots$$
$$\varphi_\Gamma(\beta_0\beta_{i(1)}\ldots\beta_{i(s-1)}, q_s) = q_{i(s)},$$
$$\cdots\cdots\cdots\cdots$$
$$\varphi_\Gamma(\beta_0\beta_{i(1)}\ldots\beta_{i(N-2)}, q_{N-1}) = q_{i(N-1)},$$
$$\beta = \beta_0\beta_{i(1)}\ldots\beta_{i(N-1)}.$$

Возьмем произвольные слова $\alpha, \gamma \in A^*$. Пусть $\varphi_\Gamma(\alpha, q_0) = q_s$. Тогда имеем

$$\begin{aligned}\varphi_\Gamma(\alpha\beta\gamma, q_0) &= \varphi_\Gamma(\beta\gamma, q_s) = \\ &= \varphi_\Gamma(\beta_0\beta_{i(1)}\ldots\beta_{i(s-1)}\beta_{i(s)}\ldots\beta_{i(N-1)}\gamma, q_s) = \\ &= \varphi_\Gamma(\beta_{i(s)}\ldots\beta_{i(N-1)}\gamma, q_{i(s)}) = \\ &= \varphi_\Gamma(\beta_{i(s+1)}\ldots\beta_{i(N-1)}\gamma, \varphi_\Gamma(\beta_{i(s)}, q_{i(s)})) = \\ &= \varphi_\Gamma(\beta_{i(s+1)}\ldots\beta_{i(N-1)}\gamma, \varnothing) = \varnothing.\end{aligned}$$

Таким образом, $\mathfrak{L} \subseteq A^*(\beta)$ и, следовательно, $\Phi(\mathfrak{L}) \subseteq A^*(\beta)$, что противоречит предположению о том, что $\mathfrak{L}$ есть фрагментное покрытие A^*. Теорема доказана.

10. Здесь рассмотрим вопрос о роли букв в регулярных языках и возможности *алфавитной редукции.* Пусть $A = \{a_1, \ldots, a_m\}$ и $\mathfrak{L} \subseteq A^*$.

Буква $a_i \in A$ называется *фиктивной* в $\mathfrak{L}$, если отображение $\alpha \to \alpha' \Rightarrow \binom{a_i}{\lambda}\alpha$, состоящее в замене a_i пустым словом λ во всех вхождениях a_i в α, таково, что из $\alpha, \beta \in \mathfrak{L}$ и $\alpha \neq \beta$ следует $\alpha' \neq \beta'$. В противном случае буква a_i называется *существенной.*

Буквы a_i и a_j называются *контекстно различимыми* в языке $\mathfrak{L}$, если отображение $\alpha \to \alpha' \Rightarrow \begin{pmatrix} a_i \\ a_j \end{pmatrix} \alpha$ таково, что из $\alpha, \beta \in \mathfrak{L}$ и $\alpha \neq \beta$ следует $\alpha' \neq \beta'$.

Язык $\mathfrak{L}$ называется *неприводимым,* если все его буквы существенные и попарно контекстно неразличимы, в противном случае говорят, что $\mathfrak{L}$ допускает *алфавитную редукцию* (т. е. взаимно однозначное отображение $\mathfrak{L}$, состоящее в элиминации фиктивной буквы, либо в отождествлении какой-нибудь пары контекстно различимых букв).

Для регулярных языков задачи выявления фиктивных букв и контекстно различимых пар букв допускают алгоритмическое решение. Легко показать что *результат любой алфавитной редукции регулярного языка есть регулярный язык,* поэтому после конечного числа применений соответствующих алгоритмов можно найти все неприводимые языки, к которым можно свести данный регулярный язык последовательностью алфавитных редукций.

Пусть $\mathfrak{L} = \mathfrak{L}(\Gamma)$ и Г — регулярный (N, m)-источник.

Букву $a_i \in A$ назовем циклической, если в $G(\Gamma)$ есть цикл, всем ребрам которого приписана буква a_i. В противном случае букву a_i назовем ациклической. Если a_i — ациклическая, то пусть $l(a_i)$ — максимальное число ее последовательных вхождений в слова $\mathfrak{L}$.

Лемма. *Если буква a_i циклическая, то она существенная.*

Доказательство. Пусть $\alpha = \beta a_i^s \gamma \in \mathfrak{L}$, где вхождение a_i^s порождается циклом порождающего пути. Тогда $\delta = \beta\gamma \in \mathfrak{L}$, так как $\varphi_\Gamma(\alpha, q_0) = \varphi_\Gamma(\delta, q_0)$, и $\alpha' = \delta'$, т. е. a_i — существенная, ч. т. д.

Теорема 1. *Если a_i — существенная буква, то найдутся* $\alpha, \beta \in \mathfrak{L}$ *такие, что*

$$\alpha \neq \beta, \ \alpha' = \beta' \ \text{ и } \ |\alpha| \leqslant |\beta| \leqslant 2N^3.$$

Доказательство. Согласно лемме можно сразу предположить, что a_i — ациклическая. Построим A-источник Γ_1 над алфавитом

$$\left\{\frac{a_j}{a_j}\,\middle|\, j = \overline{1, m}\right\} \cup \left\{\frac{b_j^\nu}{b_j^\mu}\,\middle|\, j = \overline{1,m},\ j \neq i, \mu, \nu \leqslant l(a_i)\right\},$$

где $b_i^\nu \Rightarrow a_j a_i^\nu$.

(Символ a_i/a_i, а также еще и символы вида a_i^j/a_i^k $(j,k \geq 0)$ могут быть приписаны только ребрам, идущим из начального состояния.)

Состояния Γ_1: $\{q_0/q_0,\ q_0/q_1, \ldots, q_i/q_j, \ldots, q_{N-1}/q_{N-1}\}$, их N^2, начальное — q_0/q_0, и функция следующего состояния определяются посредством

$$\varphi_{\Gamma_1}\left(\frac{x}{y}, \frac{q_i}{q_k}\right) = \frac{\varphi_\Gamma(x, q_i)}{\varphi_\Gamma(y, q_k)}, \text{ где } \varphi_\Gamma(b_j^\nu, q_k) \Rightarrow \varphi_\Gamma(a_j a_i^\nu, q_k).$$

Множество заключительных состояний пусть будет

$$Q_1' = \left\{\frac{q_i}{q_k}\,\middle|\, q_j \in Q',\, q_k \in Q'\right\}.$$

Некоторые ребра в графе источника Γ_1 выделим (скажем, изобразим двойной стрелкой: ребро выделено, если ему приписана пара a_i^ν/a_i^μ или b_j^ν/b_j^μ с $\mu \neq \nu$).

Легко проверить, что буква a_i существенная в $\mathfrak{L}$ в том и только том случае, если в Γ_1 найдется путь из начального состояния в заключительное, содержащий выделенное ребро (такой путь порождает пару различных слов $\frac{\alpha_1}{\beta_1}$ такую, что

$$\alpha_1 \Rightarrow \alpha \in A^+ \text{ и } \beta_1 \Rightarrow \beta \in A^+, \alpha,\ \beta \in \mathfrak{L},\ \alpha \neq \beta,\ \alpha' = \beta').$$

Кратчайший из таких путей может содержать не более одного цикла, следовательно, его длина не превосходит $2N^2$ (причину, по которой может случиться, что без цикла обойтись нельзя, иллюстрирует рис. 10).

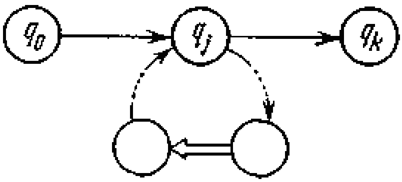


Рис. 10.

При возвращении к алфавиту A от пары слов, порожденной Γ_1, буквы b_j^ν будут превращаться в слова длины не более $l(a_i)+1 \leq N$, поэтому каждое из слов будет иметь длину не более $2N^3$. Теорема доказана. Алгоритм определения фиктивных букв следует непосредственно из теоремы 1. Самое простое — это рассмотреть все пары слов языка длины $\leq 2N^3$, но это привело бы к несообразному объему работы. В действительности часто результат дается много легче. Скажем, если язык конечно перечислим, то все буквы, имеющие вхождение хотя бы в одно слово языка, существенные: это следует из того, что в конечно перечислимое множество вместе с каждым словом входят и все его префиксы. В любом случае заметно проще построить все m источников $\Gamma_1(a_i)$ $(i = \overline{1, m})$. При этом индуктивное построение часто не придется доводить до конца, останавливаясь по получении какой-нибудь одной из искомых пар слов. Например, для языка, порожденного источником, изображенным на рис. 11, соответствующие построения представлены на рис. 12 (и построении $Г_1(a)$ нет необходимости, так как *a* — циклическая и, по лемме, существенная). Вывод: буквы *a, b, d* — существенные, *c* — фиктивная.

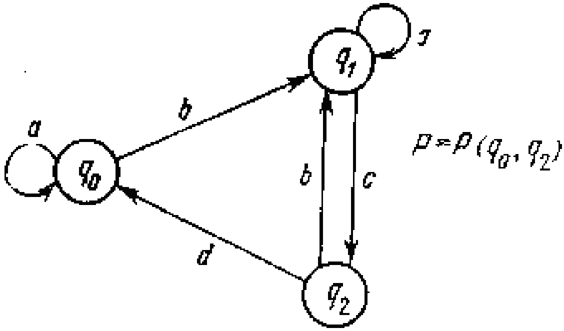


Рис. 11.

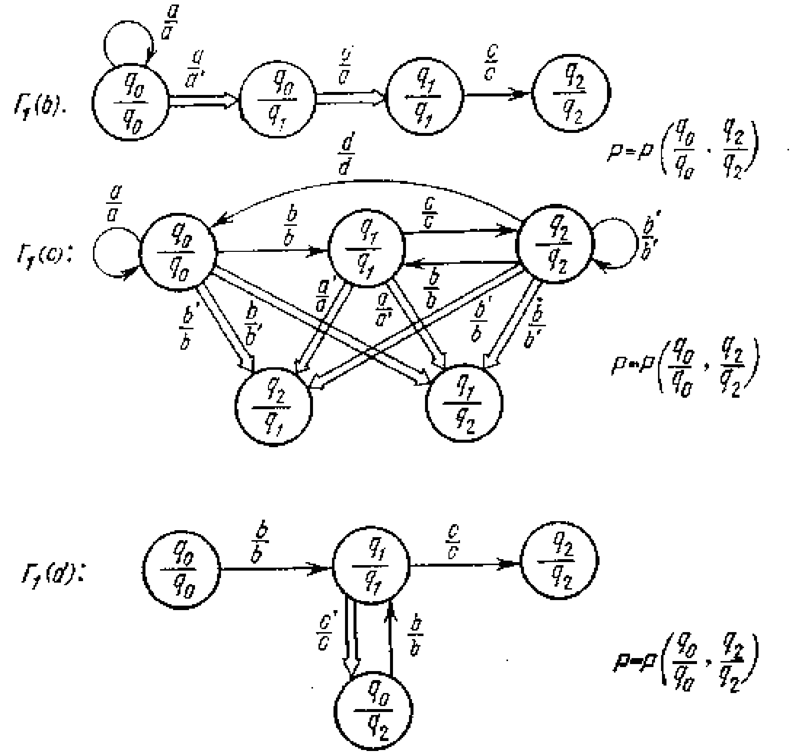


Рис. 12.

Теорема 2. *Если* a_i *и* a_j *- контекстно неразличимы, то найдутся слова* α, $\beta \in \mathfrak{A}$ *такие, что*

$$\alpha \neq \beta,\ \alpha' = \beta' \text{ и } |\alpha| = |\beta| \leqslant 2N^2.$$

Доказательство этой теоремы оставим читателю. Оно вполне аналогично предыдущему, но несколько проще построение вспомогательного источника Γ_1 (а оценка несколько лучше) за счет того, что здесь искомые слова могут быть только одинаковой длины и алфавит Γ_1 есть $\left\{\frac{a_k}{a_k} \middle| k = \overline{1, m}\right\} \cup \left\{\frac{a_i}{a_j}, \frac{a_j}{a_i}\right\}$, а выделяют те и только те ребра, которым приписаны символы

$$\frac{a_i}{a_j}, \frac{a_j}{a_i}.$$

Так, в предыдущем примере (источник на рис. 11) любая пара букв, кроме *{a, b},* контекстно различима. Для пары *{b, d}* подтверждающий это источник $\Gamma_1(b, d)$ изображен на рис. 13.

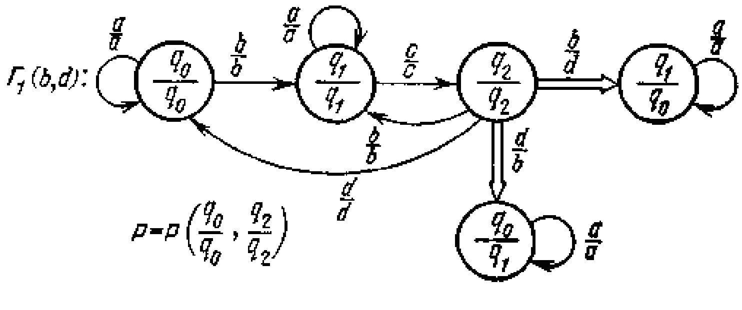


Рис. 13.

Другой источник изображен на рис. 14, *а).*

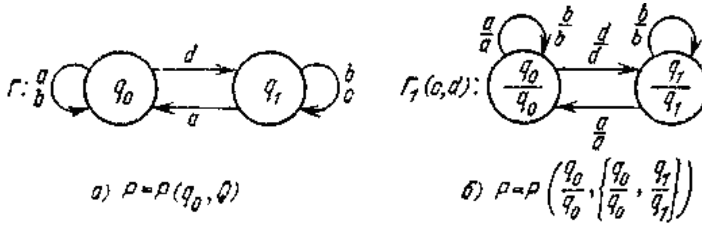


Рис. 14.

Легко проверить, что в порожденном им языке есть только две контекстно различимые буквы *c* и *d,* источник $\Gamma_1(c, d)$ показан на рис. 14, б): в нем нет ни одного выделенного ребра.

11. Регулярные источники могут быть полезны для представления отношений на множествах слов, как в п. 10, если пары слов рассматривать как слова в алфавите дробей $\left\{\frac{a_i}{\lambda}, \frac{\lambda}{a_i}\right\}_{i=1}^{m}$. Например, отношение префиксности на множестве слов $\{a, b\}^*$ допускает очевидное регулярное представление источником, изображенным на рис. 15.

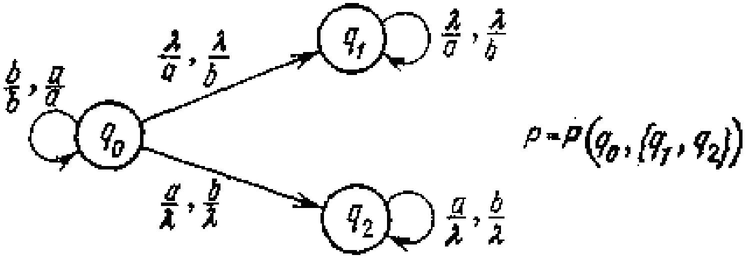


Рис. 15.

Методы теории регулярных языков обычно применимы и при изучении *регулярных отношений.* Например, для произвольного регулярного множества $\mathfrak{L} = \mathfrak{L}(\Gamma)$ отношение $\Delta_{\mathfrak{L}}$ равенства на множестве слов $\mathfrak{L}$ получаем в виде $\mathfrak{L}(\Gamma_1)$, где Γ_1 есть результат преоб разования каждого ребра Г по правилу, показанному на рис. 16.

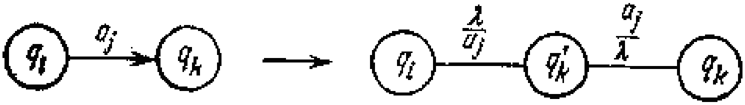


Рис. 16.

Теорема 1. *Пусть* ρ — *регулярное отношение на A^*, порождаемое источником* Г *с N_Γ состояниями,* $\mathfrak{L} = \mathfrak{L}(\Gamma_1) \subseteq A^*$ — *регулярный язык, порождаемый A-источником* Г, *с* $N_{\mathfrak{L}}$ *состояниями, и* $\rho|\mathfrak{L} \Rightarrow \rho \cap (\mathfrak{L} \times \mathfrak{L})$ — *ограничение отношения* ρ *на множество* $\mathfrak{L}$. *Тогда* $\rho|\mathfrak{L}$ — *регулярное отношение, порождаемое источником, число состояний которого не превосходит* $N_\Gamma \cdot N_{\mathfrak{L}}^2$.

Доказательство состоит в индуктивном построения источника Γ_2, порождающего $\rho|\mathfrak{L}$. Пусть

$$Q(\Gamma) = \{q_0, q_1, \ldots, q_{N_\Gamma - 1}\} \text{ и } Q(\Gamma_1) = \{r_0, r_1, \ldots, r_{N_{\mathfrak{L}} - 1}\}.$$

Возьмем $\left\langle q_0, \frac{r_0}{r_0} \right\rangle$ в качестве начального состояния Γ_2 и определим функцию следования посредством

$$\varphi_{\Gamma_2}\left(\frac{\lambda}{a}, \left\langle q_i, \frac{r_j}{r_k}\right\rangle\right) = \left\{\left\langle q_p, \frac{r_j}{\varphi_{\Gamma_1}(a, r_k)}\right\rangle \middle| q_p \in \in \varphi_{\Gamma}\left(\frac{\lambda}{a}, q_i\right)\right\},$$

$$\varphi_{\Gamma_2}\left(\frac{a}{\lambda}, \left\langle q_i, \frac{r_j}{r_k}\right\rangle\right) = \left\{\left\langle q_p, \frac{\varphi_{\Gamma_1}(a, r_j)}{r_k}\right\rangle \middle| q_p \in \in \varphi_{\Gamma}\left(\frac{a}{\lambda}, q_i\right)\right\}.$$

Состояние $\left\langle q_i, \frac{r_j}{r_k}\right\rangle$ относим к числу заключительных в Γ_2 в том и только том случае, если $q_i \in Q'(\Gamma)$ и $r_j, r_k \in Q'(\Gamma_1)$.

Остается непосредственной проверкой убедиться, что Γ_2 — искомый источник.

Упражнение. Используя теорему 1, понизить оценку в теореме 1.10 до $4N^2$.

Индивидуальные тестовые задачи

1. R — класс отношений на множестве $\{1, 2, 3, 4, 5\}$, графы которыхизображены на рис. 17. Найти графы отношений:
а) $(\varnothing)_R$; б) $\{\langle 2, 3\rangle\}_R$; в) $\{\langle 1, 2\rangle\}_R$; г) $\{\langle 2, 5\rangle, \langle 5, 1\rangle\}_R$; д) $\{\langle 4, 1\rangle, \langle 2, 3\rangle\}_R$; е) $\{\langle 4, 5\rangle\}_R$.

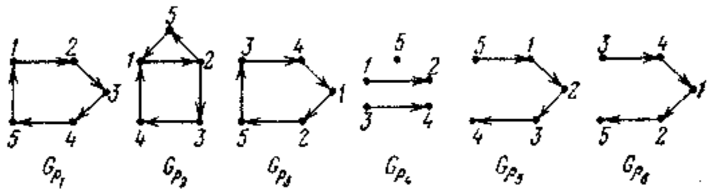


Рис. 17.

2. R — класс всех отношений эквивалентности на множестве $\{2, 3, 6, 8, 9, 12\}$, $\rho = \{\langle i, j\rangle \mid \text{НОД}\ (i, j) = 1\}$. Найти отношение $(\rho)_R$ и его граф.

3. R — класс всех отношений ρ на множестве слов $\{a, b\}^*$, для которых из $\langle \alpha, \beta\rangle \in \rho$ следует, что $\langle \alpha, \beta a\rangle \in \rho$ и $\langle \alpha, \beta b\rangle \in$

$\in \rho$, $\langle \alpha, \beta \rangle \in \tau \Rightarrow$ «α есть префикс $\{\beta\}$». Доказать, что $\tau = (=)_R$. Относительно какого правила отношение префиксности определяется отношением равенства синтаксически?

4. X, Y ... — произвольные множества. Доказать о помощью диаграмм Венна, что:

а) $X = Y$ равносильно $X \otimes Y = \varnothing$;

б) $X \subseteq Y$ равносильно $X \setminus Y = \varnothing$;

в) из условий $X \cap \bar{Z} = \varnothing$, $Y \cap \bar{W} = \varnothing$, $Z \cap W = \varnothing$, $X \cup Y = Z \cup W$ следует, что $X = Z$ и $Y = W = \varnothing$;

г) из условий $X \cap Y = Z \cup W$, $(X \cap W) \cup (Y \cap Z) = X \cup Y$ следует, что $X = Y = Z$ и $Y = W = \varnothing$;

д) из условий $X \cap Z = Y \cap W$, $(X \cap W) \cup (Y \cap Z) = Y \cup W$ следует, что $X = W$ и $Y = Z$:

е) система условий $Y \subseteq \bar{X}Z \cup X\bar{Z}$, $YW \subseteq XZ \cup \bar{X}\bar{Z}$, $XY \subseteq Z \cup W$, $YZ \subseteq X \cup W$ равносильна условию $Y = \varnothing$.

5. $A, B, C \subseteq U$, $|U| = n$. Доказать, что в универсе U

а) уравнение $A\bar{X} \cup \bar{B}Y = \bar{B}X \cup AY$ имеет $2^{2n-|A \otimes B|}$ решений;

б) уравнение $X \cup Y = \bar{Y} \cup Z$ имеет 4^n решений;

в) система уравнений $\begin{cases} A\bar{X} \cup B\bar{X} = C, \\ BX \cup A\bar{X} = \bar{C} \end{cases}$ имеет решение в том и только том случае, если $A = \bar{B}$ (если решение существует, то оно единственно и есть $X = A \otimes \bar{C}$).

6. Проверить, что если $U = P_2$ (множество всех функций алгебры логики), то на диаграмме Венна для системы замкнутых классов T_0, T_1, S, M, L пустыми будут в точности те клетки, которые на рис. 18 помечены символом $\varnothing$. Привести примеры функций каждого из остальных типов.

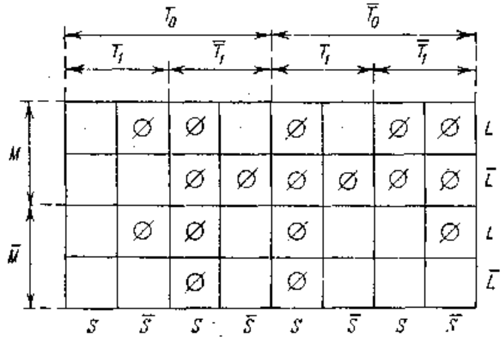


Рис. 18.

7. Определить число тупиковых и число максимальных независимых множеств для графов:
а) P_n (путь длины n);
б) C_n (цикл длины n);
в) Z_n (звезда с n вершинами);
г) $K_{i,j}$ двудольный граф.
8. Построить q-ичный префиксный код, реализующий спектр длин D:
а) $\langle 2, 2, 3, 3, 3, 5, 5, 5, 5, \rangle$, $q = 2$;
б) $\langle 2, 2, 3, 4, 4, 4, 4, 4 \rangle$, $q = 2$;
в) $\langle 1, 2, 2, 3, 3, 3, 3, 5, 5, 5 \rangle$, $q = 3$.

9. Какие из следующих полиномов являются структурными полиномами двоичных префиксных кодов?
а) $Z_1^2 + Z_1Z_2 + Z_2^2 + Z_1^2Z_2 + Z_1Z_2^2$;
б) $Z_1 + Z_1Z_2 + Z_1Z_2^2 + Z_1^2Z_2^2 + Z_2^4$;
в) $Z_1Z_2 + 2Z_1^2Z_2 + Z_1Z_2^2 + Z_1^3 + 2Z_1Z_2^3 + Z_2^4 + Z_1^3Z_2^2 + Z_1^2Z_2^3$.
10. Пусть для $\mathfrak{M} \subseteq A^*$ $F(\mathfrak{M}) = A^*\mathfrak{M} \setminus A^*\mathfrak{M}A^+$ (например, F (01, 00, 111) = {01, 00, 111, 101, 100, 1101, 1100}).
Доказать, что $F(\mathfrak{M})$ — префиксный код (возможно бесконечный) и для любого $\mathfrak{M}$ найдется $\mathfrak{M}'$ такое, что $(F(\mathfrak{M}))^2 = F(\mathfrak{M}')$.
11. Какие из следующих бесконечных префиксных кодов являются тупиковыми?
а) $\{1, 01, 001, \ldots, 0^i1, \ldots\}$,
б) $\{00, 10, 010, 0110, \ldots, 01^i0, \ldots\}$,
в) $\{0, 100, 10100, \ldots, v_i, v_{i+1}, v_{i+2} = 1\,v_iv_{i+1}, \ldots\}$.

12. Какие соотношения и какие тождества выполняются в полугруппе, порожденной преобразованием f с операцией суперпозиции преобразований?
$$\text{а)}\ f = \begin{pmatrix} 1 \to 2 \\ 2 \to 3 \\ 3 \to 3 \\ 4 \to 1 \\ 5 \to 1 \end{pmatrix}, \quad \text{б)}\ f = \begin{pmatrix} 1 \to 2 \\ 2 \to 3 \\ 3 \to 4 \\ 4 \to 2 \\ 2 \to 3 \end{pmatrix}.$$
13. Доказать, что матрицы
$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \text{ и } B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$
порождают свободную полугруппу, изоморфную $\{A, B\}^+$.
14. Доказать, что в полугруппе натуральных чисел по сложению выполняется квазитождестно Мальцева.
15. Найти все разложения слова α на множители из V, где:

а) $V = \{1, 10, 01\}$, $\alpha = 10100101$, $\alpha = (100)^i1$;

б) $V = \{11, 110, 01, 101\}$, $\alpha = 110101$, $\alpha = 1^i01^j01$;

в) $V = \{0^3, 0^5, 0^7\}$, $\alpha = 0^{10}$, $\alpha = 0^{17}$, $\alpha = 0^{21}$.

16. Найти порождающие множества и их структурные функции следующих подполугрупп свободных полугрупп;

а) полугруппа всех (0, 1)-слов, содержащих четное число вхождений нуля;

б) полугруппа всех двоичных слов, начинающихся с единицы;

в) полугруппа всех (0, 1, 2)-слов, у которых сумма входящих цифр четна;

г) полугруппа всех двоичных слов, у которых в любом префиксе число вхождений нуля не превосходит числа вхождений едипицы.

17. Доказать, что для $\mathfrak{I}$ полугруппы, заданной словесным представлением *V*, изоморфное представление образующими и определяющими соотношениями есть *П:*

а) $V = \{0^5, 0^6\}$, $\Pi\langle\{x, y\} \mid xy = yx, x^6 = y^5\rangle$;

б) $V = \{a, a(ba^2)^2b, (ba^2)^4\}$, $\Pi\langle\{x, y, z\} \mid xz^3 = (yx)^3yx^2\rangle$;

в) $V = \{0^6, 0^9, 0^{17}\}$, $\Pi\langle\{x, y, z\} \mid xy = yx, xz = zx, yz = zy, x^3 = y^2, z^3 = x^4y^3\rangle = \Pi\langle\{x, y, z\} \mid xy = yx, xz = zx, yz = zy, x^3 = y^2, z^3 = x^7y\rangle$;

г) $V = \{(ab)^2a, (ba)^3b, (ab^4)\}$ $\Pi\langle\{x, y, z\} \mid xyz = zxy, (xy)^2 = z^3\rangle$;

д) $V = \{(ab)^4a, (ba)^3b, (ab)^6a\}$, $\Pi\langle\{x, y, z\} \mid xyz = zyx, (zy)^3z = (xy)^2x\rangle$.

18. Доказать, что, если слова α' и β' имеют общий префикс длины $|\alpha| + |\beta|$, то $\alpha\beta = \beta\alpha$ и $\varepsilon(\alpha) = \varepsilon(\beta)$.

19. Если *x* — слово, *y* — бесконечная последовательность, то под произведением $x \bullet y$ понимается бесконечная последовательность *xy*. Пусть $\alpha = \beta\beta \ldots \beta \ldots = \beta^\infty$ — бесконечная периодическая последовательность. Доказать, что:

а) $\beta^i\alpha = \alpha$ для любого натурального числа *i;*

б) если $|\beta| = \tau$ — наименьший период α, а *T* — тоже период α, то τ является делителем *T.*

20. Доказать, что уравнение в словах $xz^2xy = yzyx^2$ имеет только тривиальные решения.

21. Доказать, что общее решение уравнения в словах $xyz = zyx$ есть

$$x = (\alpha\beta)^i\alpha,\ y = \beta(\alpha\beta)^j,\ z = (\alpha\beta)^k\alpha,$$

где α, β — произвольные слова, *i, j, k* — произвольные неотрицательные целые числа.

22. Доказать, что общий вид нетривиального решения уравнения в словах $xyz = zxy$ есть

$$x = (\alpha\beta)^i\alpha,\quad y = (\beta\alpha)^i\beta,\quad z = (\alpha\beta)^k,$$

где α, β — произвольные слова такие, что $\varepsilon(\alpha) \neq \varepsilon(\beta)$, $l, j, k \geqslant 0$. Показать, что среди его решений имеется бесконечное число попарно неизоморфных.

23. Найти общее решение следующих уравнений и систем уравнений в словах:

$$\text{а) } xyxz = zx^2y; \qquad \text{б) } z^3 = x^2yx^3; \qquad \text{в) } \begin{cases} xyz = yzx, \\ yx = uz. \end{cases}$$

24. Найти экстремальные унтер-решения следующих систем уравнений в словах над алфавитом $\{0, 1\}$ с правилом вывода следствий (3):

$$\text{а) } \begin{cases} z = y^2x, \\ y^3 = x^2; \end{cases} \qquad \text{б) } \begin{cases} yx = xz, \\ xu = y^2, \\ ux = z^2, \\ uy = zu. \end{cases}$$

25. Доказать, что, если ρ — конечная система уравнений в словах, то V является унтер-решением системы ρ в том и только в том случае, если выполняется свойство единственности разложения на множители по V.

26. Пусть A — алфавит, $B \subseteq A^{m_1}$, $C \subseteq A^{m_2}$. Доказать, что если $B^i = C^j$, то существует множество слов D такое, что $B = D^k$ и $C = D^l$ для некоторых целых чисел k, l.

27. Доказать, что, если $x^iy = yz^i$ при некотором $i > 0$, то $x^jy = yz^j$ при любом j.

28. Доказать, что:

а) если множество $\mathfrak{L}$ регулярно, то и его обращение $O(\mathfrak{L}) = \{x_1x_2 \ldots x_l \mid x_l \ldots x_2x_1 \in \mathfrak{L},\ x_j \in A\ \ (j \geqslant 1)\}$ тоже регулярно;

б) пусть для слова $\alpha = x_1^{i_1} \cdots x_k^{i_k}$, где x_j $(j = \overline{1, k})$— буквы алфавита A и $x_j \neq x_{j+1}$ для $j = \overline{1, k-1}$, обозначено $\alpha' = x_1 \ldots x_k$ и для множества слов $\mathfrak{L}$ пусть $\mathfrak{L}' = \{\alpha' \mid \alpha \in \mathfrak{L}\}$.

Тогда если $\mathfrak{L}$ регулярно, то и $\mathfrak{L}'$ тоже регулярно;

в) если $\mathfrak{L}_1\mathfrak{L}_2$ регулярно, где — $\mathfrak{L}_1$ и $\mathfrak{L}_2$ произвольные множества слов, то существует регулярное множество $\mathfrak{L}_3$ такое, что $\mathfrak{L}_1\mathfrak{L}_2 = \mathfrak{L}_1\mathfrak{L}_3$, и среди таких $\mathfrak{L}_3$ есть наибольшее;

г) если $\mathfrak{L}$ регулярно, то и множество $\pi(\mathfrak{L})$ всех префиксов слов из $\mathfrak{L}$ регулярно;

д) если $\mathfrak{L}$ регулярно, то и множество $\varphi(\mathfrak{L})$ всех фрагментов слов из $\mathfrak{L}$ тоже регулярно;

е) множество $\{a, a^2, a^4, \cdots\} = \{a^{2^k} \mid k = 0, 1, 2 \cdots\}$ не регулярно.

29. Перечислить классы $A^*/\rho_{\mathfrak{L}}$ для отношения правой синтаксической зквивалентиости в случаях:

а) $\mathfrak{L} = \{ab, ca, bc, bb, ba\}$, $A = \{a, b, c\}$;
б) $\mathfrak{L} = \{ab, aba, bb\}$, $A = \{a, b\}$;
в) $\mathfrak{L} = \{a^3, a^4, \ldots\} = \{a^i | i \geqslant 3)\}$, $A = \{a, b\}$.

30. Пусть Г = <*G*, *F*, *P*> — источник и вершины его графа раскрашены в два цвета, а *P* — множество всех путей из $P(q_0, Q')$ таких, что число вершин каждого цвета, через которые они проходят, одинаково. Доказать, что языки, порождаемые классом таких источников, включают в себя класс регулярных языкои как собственное подмножество, и привести пример нерегулярного $\mathfrak{L}(\Gamma)$.

31. Привести примеры регулярных языков, для которых не существует порождающих регулярных источников: а) связного и б) с одним заключительным состоянием.

32.Найти наименьшее фрагментно ограниченное приближение $\Phi(\mathfrak{L})$ для языков:

а) $\mathfrak{L} = \{aba, bbb\}$;
б) $\mathfrak{L} = \{aabba\}$;
в) $\mathfrak{L} = \{ab, a^2b, a^3b, \ldots, a^ib, \ldots\}$.

33. Определить все возможные цепочки алфавитных редукций для языков:

а) $\mathfrak{L} = \{abac, abb\}$;
б) $\mathfrak{L} = \{abac, bca, cac, bbab\}$;
в) $\mathfrak{L} = \mathfrak{L}(\Gamma_i)$ при $P = P(q_0, q_0)$

для регулярных языков, порожденных источниками, изображенными на рис. 19.

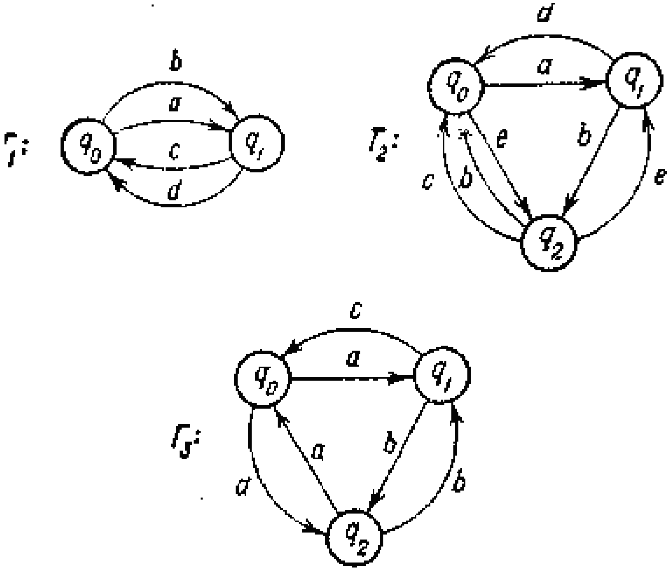


Рис. 19.

34. Даны алфавит $A = \{ a_1, ..., a_n\}$ и регулярный язык $\mathfrak{L}$. Найти регулярное представление следующих отношений на $\mathfrak{L}$:

а) $\left\{\frac{\alpha}{\beta}\,\middle|\,\alpha \text{ — префикс } \beta\right\}$; б) $\left\{\frac{\alpha}{\beta}\,\middle|\,\alpha \text{ — фрагмент } \beta\right\}$;

в) $\left\{\frac{\alpha}{\beta}\,\middle|\,\alpha \text{ — подпоследовательность } \beta\right\}$;

г) $\tau_i = \left\{\frac{\alpha}{\beta}\,\middle|\,\alpha \neq \beta, \alpha' = \beta',\right\} (i = \overline{1, n})$ (x' получается из x заменой всех вхождений a_i на λ). Используя такое представление, улучшить оценку в теореме 1 до $4N^2$;

д) $\tau_{ij} = \left\{\frac{\alpha}{\beta}\,\middle|\,\alpha \neq \beta, \alpha' = \beta'\right\} (i, j = \overline{1, n})$ (x' получается из x отождествлением всех вхождений букв a_i и a_j).

11. Приложение отношений

11.1. Законы композиции

Композиция объектов. В математике и ее приложениях большое значение имеют отношение, которые ставят в соответствие паре каких-либо объектов (a, b) третий объект c. Примерами таких отношений являются действия над числами. В общем случае отношение может представлять собой некоторую *операцию* не только между числами, но и между объектами любой природы. При этом запись $a \top b = c$, или $a \perp b = c$, означает, что a в композиции с b дает c. Символ $\top$ (или $\perp$) обозначает операцию, объекты a и b *называются операндами,* а *объект* c - *результатом операции или композицией объектов* a и b.

Обозначим множества операндов соответственно через A и B ($a \in A$ и $b \in B$), а множество результатов операции — через C ($c \in C$). Так как множество всех пар (a, b) есть прямое произведение $A \times B$, то операцию определяют как отображение множества $A \times B$ в C, т.е. $A \times B \to C$, и часто называют *законом композиции.*

Таблица Кели. Любой закон композиции $A \times B \to C$ над конечными множествами можно задавать прямоугольной матрицей *(таблицей Кели).* Строки таблицы отвечают элементам множества A, столбцы — элементами множества B. На пересечении строки и столбца, соответствующих паре (a, b), располагается элемент $c = a \top b$.

Хорошо известными примерами являются таблицы сложения и умножения одноразрядных чисел. В общем случае таблица, которая определяет бинарную операцию, имеет вид:

$\top$	b_1	b_2	b_3	b_4	b_5	...
a_1	c_{11}	c_{12}	c_{13}	c_{14}	c_{15}	...
a_2	c_{21}	c_{22}	c_{23}	c_{24}	c_{25}	...
a_3	c_{31}	c_{32}	c_{33}	c_{34}	c_{35}	...
a_4	c_{41}	c_{42}	c_{43}	c_{44}	c_{45}	...
...	...	...	...	...	...	...

Законы композиции на множестве. Множества *A, B, C,* которые принимают участие в операции $A\times B \to C$, не обязательно должны быть различными. Если $B=C=S$, то говорят, что закон композиции *определен на множестве S.*

Различают *внутренний* закон композиции $S \times S \to S$ и *внешний* закон композиции $\Omega \times S \to S$, где Ω и S — различные множества. В случае внутреннего закона говорят, что множество образует *группоид* относительно операции $\top$. В случае внешнего закона композиции элементы $\alpha \in \Omega$ называют *операторами,* а Ω — *множеством операторов* на множестве S.

Примерами внутреннего закона композиции являются сложение $a+b=c$ и умножение $ab=c$ на множестве действительных чисел, а также геометрическое суммирование векторов на плоскости. Умножение вектора на скаляр может быть примером внешнего закона композиции на множестве векторов, причем операторами являются скаляры — элементы множества действительных чисел.

Пусть S — множество диференцируемых функций $f_i(x_1, x_2, ... x_n)$ и Ω - множество операторов дифференцирования $\partial/\partial x_j$ ($j = 1, 2, ..., n$). Тогда паре $(\partial/\partial x_j, f_i)$ можно поставить в соответствие частную производную df_i/dx_j, т.е. имеем внешний закон композиции на множестве диференцируемых функций.

Ниже речь будет идти о внутренних законах композиции.

Матрица и граф группоида. Конечный группоид S относительно закона $\top$ определяется квадратной матрицей n-го порядка (n — число элементов группоида), например,

$\top$	a	b	c	d
a	b	c	a	b
b	a	b	c	a
c	b	a	d	d
d	d	b	d	b

Построение графа группоида основано на представлении бинарного соотношения $a \top b = c$ (рис. 1, *а),* где дуги графа изображают элементы $a, b, c \in S$, причем операнды образуют некоторый путь, а дуга результата операции замыкает этот путь. Если $a \top b = a$, то b изображается петлей в конечной вершине дуги a. При построении графа сначала наносят дуги для всех элементов группоида как выходящие из одной вершины, а затем последовательно изображают все бинарные соотношения.

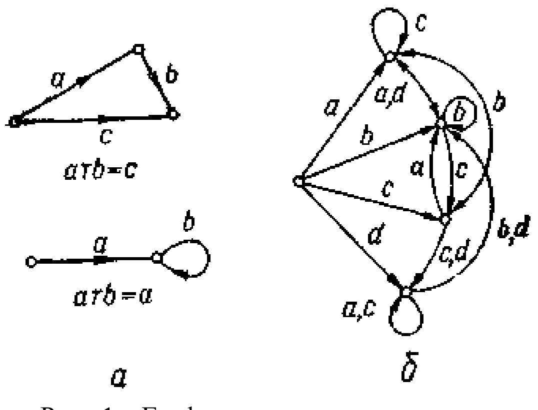


Рис. 1. Граф операции на множестве:
а — операнды *a, b* и результат операции *c; б* - граф группоида

На рис. 1, *б* изображен граф группоида, заданного приведенной выше матрицей. Дуги *a, b, c, d,* выходящие из одной вершины, соответствуют элементам группоида.Так как $a \top a = b$, $a \top b = c$, $a \top c = a$ и $a \top d = b$, то из конца дуги *a* проводят дуги *a, b, c, d* соответственно к конечным вершинам дуг *b, c, a, b.* Две параллельные дуги *a* и *d,* направленные к конечной вершине дуги *b,* условно изображают одной дугой *a, d.* Дуга

c начинается и кончается в конечной вершине дуги a, т. е. образует петлю. Аналогично изображают на графе и остальные соотношения, определяемые матрицей группоида.

Свойства внутреннего закона композиции. Операции на множестве S могут обладать некоторыми общими свойствами, которые обычно выражаются соотношениями между элементами с S:

коммутативность $a \top b = b \top a$;

ассоциативность $a \top (b \top c) = (a \top b) \top c$;

дистрибутивность слева $(a \top b) \perp c = (a \perp c) \top (b \perp c)$

и справа $c \perp (a \top b) = (c \perp a) \top (c \perp b)$.

На множестве действительных чисел сложение и умножение ассоциативны и коммутативны. Умножение дистрибутивно (слева и справа) относительно сложения, но сложение не дистрибутивно относительно умножения, так как вообще $a + bc \neq (a + b)(a + c)$.

Возведение в степень не ассоциативно $(a^b)^c \neq a^{(b^c)}$, не коммутативно $a^b \neq b^a$, но дистрибутивно справа относительно умножения, так как $(ab)^c = a^c b^c$.

Пересечение и объединение множеств взаимно дистрибутивны относительно друг друга. Если в множестве $F \subset S$ композиция любых двух элементов из F также принадлежит F, то F называется *замкнутым* относительно рассматриваемого закона композиции (подмножество четных чисел является замкнутым относительно сложения и умножения).

Регулярный, нейтральный и симметричный элементы. Закон композиции наделяет элементы множества некоторыми общими свойствами. При различных законах одни и те же элементы могут обладать различными свойствами. Поэтому имеет смысл говорить о свойствах элементов множества S относительно заданного на нем закона композиции $\top$.

Элемент a называется *регулярным,* если из соотношений $a \top x = a \top y$ и $x \top a = y$ следует $x = y$ (*сокращение* на регулярный элемент). Всякое число регулярно относительно сложения, а для умножения регулярно всякое число, кроме нуля ($0x = 0y$ не влечет $x = y$).

Нейтральным элементом $e \in S$ называют такой элемент, что для всех элементов x из S справедливо $e \top x = x \top e = x$ (если нейтральный элемент существует, то он единственен и регулярный). Среди чисел нуль — нейтральный элемент относительно сложения, а единица — относительно умножения. Пустое множество является нейтральным элементом относительно объединения, а основное множество (универсум) — относительно пересечения. На множестве всех квадратных матриц n-го порядка с числовыми элементами нулевая и единичная матрицы служат соответственно нейтральными элементами относительно сложения и умножения.

Если множество содержит нейтральный элемент e относительно закона композиции $\top$, то элемент b называется *симметричным (обратным, противоположным)* элементу a, если $a \top b = b \top a = e$; при этом a называют *симметризуемым* элементом и b обозначается через $\overline{a}$, т.е. $b = \overline{a}$. Относительно ассоциативного закона $\top$ элемент $\overline{a}$, симметричный элементу a (если он существует), единственен и регулярен.

При сложении симметричным некоторому числу x будет $-x$, а при умножении x^{-1}. Например, симметричными элементами на множестве квадратных матриц n-го порядка относительно умножения есть взаимо-обратные матрицы. Множество всех собственных подмножеств относительно объединения или пересечение не содержит симметричных элементов. Множество, в котором всякий элемент имеет симметричный, называется *симетризуемым.*

Аддитивнные и мультипликативные обозначения. Свойства законов композиции можно представить в двух формах. В аддитивных обозначениях операция $\top$ записывается символом сложения (+), а в мультипликативных — символом умножения (•). Если множество наделено двумя законами композиции, то чаще всего первый из них $\top$ считается *аддитивным,* а второй $\perp$ — *мультипликативным.* В аддитивной записи нейтральный элемент обозначается через 0 и называется *нулем,* а симметричный элементу a — через $(-a)$. В мультипликативной записи нейтральный элемент обозначается через 1 и называется *единицей,* а симметричный элементу a — через a^{-1}.

Если закон композиции ассоциативный и коммутативный, а элементы множества $x_1, x_2, ..., x_n \in S$ отмечены *операторным индексом* i, то в аддитивной записи

$$x_1 + x_2 + \ldots + x_n = \sum_{i=1}^{n} x_i$$

и в мультипликативной записи

$$x_1 \cdot x_2 \cdot \ldots \cdot x_n = \prod_{i=1}^{n} x_i \ .$$

Следует подчеркнуть, что здесь, в отличие от элементарной алгебры, знаки (+) и (•) не обязательно означают сложение и умножение чисел. Они просто заменяют в различных соотношениях символы $\top$ и $\perp$, указывая на то, что над элементами множества (не обязательно числами) выполняются некоторые операции. Эти операции могут лишь извне напоминать обычные операции сложения или умножение чисел, но по существу в общем случае — это другие операции. Удобство аддитивных и мультипликативных обозначений заключается в том, что при операциях над числами различные соотношения совпадают с общепринятой формой записи.

Алгебраические системы. Определяя на некотором множестве S один или два закона композиции и наделяя их определенными свойствами, а также задавая структуру множества относительно законов композиции (наличие нейтрального элемента и симметризуемость множества), получаем различные *алгебраические системы (структуры* или *модели).* Наиболее употребительные из них такие :

Группа — это наделенное ассоциативным законом множество, содержащее нейтральный элемент и симметризуемо относительно этого закона. Если, кроме того, закон композиции коммутативный, то группу называют *абелевой (коммутативной).*

Во всякой группе соотношения (уравнения) $a \top x=b$ и $y \top a=b$ допускают единственное решение $x = \overline{a} \top b$ *(частное справа)* и $y=b \top \overline{a}$ *(частное слева)*. Имеет место также соотношения $\overline{(a \top b)}=\overline{b} \top \overline{a}$ или $-(a+b)=-b-a$ (в аддитивной записи) и $(a \cdot b)^{-1} = b^{-1} \cdot a^{-1}$ (в мультипликативной записи).

Кольцо — это множество, которое наделено двумя законами композиции: относительно первого (аддитивного) оно образует абелеву группу, а второй закон (мультипликативный) является ассоциативным, а также дистрибутивным относительно первого закона.

Телом называют кольцо с единицей, в котором каждый отличный от нуля элемент обладает симметричным относительно второго (мультипликативного) закона.

Поле — это коммутативное тело.

11.2. Модель и отношения

Одним из основных в дискретной математике является понятие модели. *Моделью* Ψ называется совокупность множества M с заданными в нем отношениями

$$S = \{R_{11}, R_{12}, ..., R_{1n_1}, R_{21}, R_{22}, ..., R_{2n_2}, ...R_{m1}, R_{m2}, ..., R_{mn_m}\},$$

где множество M — *носитель модели,* а заданные отношения R_{ia}, $R_{ia} \subset M^i$ образуют *сигнатуру модели*

$$\Psi = \langle M,\ S\rangle.$$

Степень носителя определяет *арность отношения*. Два отношения R_α и R_β, имеющие одну и ту же степень, называются *совместимыми по объединению* или просто *совместимыми*.

Очевидно, что n-местную операцию $f_n(m_1,m_1,\ldots,m_n)=m_{n+1}$ можно рассматривать как $(n+1)$-арне отношение R_{n+1}.

Совокупность множества M с заданными в нем операциями и отношениями, будем называть *алгебраической системой.*

Частным случаем алгебраической системы есть *алгебра отношений* и ее расширение — *реляционня алгебра.*

Рассмотрим алгебру отношений, носитель которой — множество отношений, а сигнатура - операции объединения, пересечение, разности и расширенного декартова произведения отношений.

Объединением $R_\alpha \square R_\beta$ *двух совместимых отношений* R_α и R_β является множество всех кортежей, каждый из которых принадлежит хотя бы одному з этих отношений. Объединением отношений

$$R_\alpha=\{(a,b,c),(a,b,d),(b,c,e)\} \text{ и } R_\beta=\{(a,b,d),(b,d,e),(c,d,e)\}$$

является

$$R_\alpha \cup R_\beta =\{(a,b,c),(a,b,d),(b,c,e),(b,d,e),(c,d,e)\}.$$

Рассмотренные отношения являются совместимыми, так как их степени равны:

$$s(R_a)=s(R_b) =3,\ R_a,\ R_b \subset M^3,\ M=\{a,b,c,d,e\}.$$

Пересечением $R_\alpha \cap R_\beta$ *двух совместимых отношений* R_α и R_β является множество всех кортежей, которые принадлежат как отношению R_α, так и отношению R_β. Пересечением отношений R_α и R_β является

$$R_\alpha \mathrm{I}\ R_\beta =\{(a,b,c),(a,b,d),(b,c,e)\}\ \mathrm{I}\ \{(a,b,d),(b,d,e),(c,d,e)\}= \{(a,b,d)\}.$$

Разностью $R_\alpha \backslash R_\beta$ *двух совместимых отношений* R_a и R_β является множество всех кортежей, которые принадлежат отношению R_a и не приналежать отношению R_β. Так, например,

$$R_\alpha \backslash R_\beta =\{(a,b,c),(a,b,d),(b,c,e)\}/\{(a,b,d),(b,d,e),(c,d,e)\}= \{(a,b,c),(b,c,e)\}.$$

Расширенным декартовым произведением $R_a \times R_\beta$ *двух отношений* R_a и R_β является множество всех кортежей π таких, что π — конкатенация кортежа $a \in R_a$ и кортежа $b \in R_\beta$ *(конкатенация* кортежей $(a_1, a_2,...,a_n)$ и $(b_1, b_2,...,b_m)$ - кортеж $(a_1, a_2,...,a_n, b_1, b_2,...,b_m)$). Например, для рассмотренных отношений R_a и R_β расширенное декартово произведение

$R_a \times R_\beta = \{(a,b),(c,d),(a,e)\} \times \{(a,b,c),(b,d,e)\} = \{(a,b,a,b,c),(a,b,b,d,e),$ $(c,d,a,b,c),(c,d,b,d,e),(a,e,a,b,c),(a,e,b,d,e)\}$.

Понятие модели и алгебры отношений находят широкое применение при формализации реальных объектов. Рассмотрим, как используется алгебра отношений при создании информационного обеспечения — разработке реляционной базы данных.

Основой построения реляционной базы данных является двумерная таблица, каждый столбец которой соответствует домену (или атрибуту, соответствуещему части домена), строка - кортежу значений атрибутов, находящихся в отношении *R*.

Рассмотрим 5-арне отношение R_5 (испытания) (табл. 1).

Таблица 1.

R_5	D_1	D_2	D_3	D_4	D_5
1	К5-01	ТЕОРИЯ АВТОМАТОВ	ПРОФ. ПЕТРЕНКО	03. ЯНВ.	АУД. 210
2	К5-02	МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА	ПРОФ. ИВЧЕНКО	03. ЯНВ.	АУД. 211
3	К5-03	ФИЗИКА	ПРОФ. КРАВЧЕНКО	03. ЯНВ.	АУД. 211
4	К5-04	АЛГОРИТМИЧЕСКИЕ ЯЗЫКА	ПРОФ. ИВЧЕНКО	05. ЯНВ.	АУД. 210
5	К5-01	ФИЗИКА	ПРОФ. КРАВЧЕНКО	09. ЯНВ.	АУД. 210
6	К5-02	ТЕОРИЯ АВТОМАТОВ	ПРОФ. ПЕТРЕНКО	09. ЯНВ.	АУД. 211
7	К5-03	АЛГОРИТМИЧЕСКИЕ ЯЗЫКА	ПРОФ. ИВЧЕНКО	10. ЯНВ.	АУД. 211
8	К5-04	МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА	ПРОФ. ПЕТРЕНКО	10. ЯНВ.	АУД. 210

Таблица 1 определяет отношение реляционной модели данных. Отношение R_5 является подмножеством декартова произведения $D_1 \times D_2 \times D_3 \times D_4 \times D_5$, в котором сомножитель является доменом D_i. Элементами домена D_i служат значения атрибутов.

Домен D_1 (группа) содержит значение К5-01, К5-02, К5-03, К5-04:

D_1 = {К5-01, К5-02, К5-03, К5-04};

аналогично имеем домены:

D_2 (дисциплина):

D_2={ТЕОРИЯ АВТОМАТОВ, МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА, ФИЗИКА, АЛГОРИТМИЧЕСКИЕ ЯЗЫКИ};

D_3 (экзаменатор):

D_3={ПРОФ. ПЕТРЕНКО, ПРОФ. ИВЧЕНКО, ПРОФ. КРАВЧЕНКО};

D_4 (дата):

D_4 = {03 ЯНВ., 05 ЯНВ., 09 ЯНВ., 10 ЯНВ.};

D_5 (аудитория):

D_5 = {АУД. 210, АУД. 211}.

Порядок столбцов в таблице фиксирован, строки в общем случае могут располагаться произвольно. Цифры первого столбца 1, 2,..., 8 идентифицируют элементы отношения D_5.

Для преобразования отношений определим *реляционную алгебру*. Носитель реляционной алгебры прелставляет собой множество отношений, сигнатура кроме введенных операций (объединение, пересечение, разность и расширенное декартово произведение отношений) включает специальные операции над отношениями: выбор, проекцию и соединение.

Операция *выбора* представляет собой процедуру построения «горизонтального» подмножества отношений, т.е. подмножества кортежей, обладающих заданным свойством.

Пример 2. С помощью операции выбора построить отношение R'_5 (расписание экзаменов проф. Петренка). Результатом операции выбора являются строки, у которых домен D_3 представлен значениям, ПРОФ. ПЕТРЕНКО; это 1, 6, 8-я строки (табл. 2).

Таблица 2.

R_5	D_1	D_2	D_3	D_4	D_5
1	К5-01	ТЕОРИЯ АВТОМАТОВ	ПРОФ. ПЕТРЕНКО	03.ЯНВ.	АУД. 210
2	К5-02	ТЕОРИЯ АВТОМАТОВ	ПРОФ. ПЕТРЕНКО	09. ЯНВ.	АУД. 211
3	К5-04	МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА	ПРОФ. ПЕТРЕНКО	10. ЯНВ.	АУД. 210

Для определения проекций отношений множество в реляционной алгебре разбивается на два подмножества в случае бинарного отношения и на n подмножеств в случае n-арного отношения:

$$R_2 \subset M^2,\ M=A \cup B,\ A \cap B=\varnothing,\ R_2 \subset A\times B;$$

$$R_n \subset M^2,\ M=\bigcup_{i=1}^{n} A_i;\ A_{i_a} \cap A_{i_b} = \varnothing,$$

$$i_a,\ i_b\ (i_a \neq i_b) \in \{i_1, i_2,\ldots,i_n\}, R_n \subset A_1\times A_2\times\ldots\times A_n\ldots$$

Проекцией Пр(R_2/A) *бинарного отношения* R_2, $R_2 \subset A\times B$ на A называется множество элементов $\{a_i/(a_i,b_i) \in R_2\}$.

Проекцией Пр$(R_n/ A_{i_1}, A_{i_2},\ldots, A_{i_m})$ n-арного отношения $R_n \subset A_1\times A_2\times\ldots\times A_n$, $n\geq m$, на $A_{i_1}, A_{i_2},\ldots, A_{i_m}$ называется множество кортежей $(a_{i_1}, a_{i_2},\ldots, a_{i_m})$, где $a_{i_1} \in A_{i_1}, a_{i_2} \in A_{i_1},\ldots, a_{i_m} \in A_{i_1}$, каждый из которых является частью элемента n-арного отношения R_n.

Операция проекции определяет построение «вертикального» подмножества отношения, т.е. множества подмножества кортежей, получаемого выбором одних и исключением других доменов.

Пример 3. Проекция Пр $(R_5/D_2,\ D_3)$ порождает множество пар**,** каждая из которых определяет дисциплину и экзаменатора (табл. 3).

Таблица 3

R_2	D_2	D_3
	ТЕОРИЯ АВТОМАТОВ	ПРОФ. ПЕТРЕНКО
	МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА	ПРОФ. ИВЧЕНКО
	ФИЗИКА	ПРОФ. КРАВЧЕНКО
	АЛГОРИТМИЧЕСКИЕ ЯЗЫКА	ПРОФ. ИВЧЕНКО
	МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА	ПРОФ. ПЕТРЕНКО

Одинаковые строки в табл. 3 объединены в одну.

Операция соединения по двум таблицам, имеющим общий домен, позволяет построить одну таблицу, каждая строка которой образуется соединением двух строк исходных таблиц. Из заданных таблиц берут строки, содержащие одно и то же значение из общего домена; общему домену сопоставляется один столбец.

Пример 4. Найдем по двум заданным таблицам (табл. 4,*а*, 4,*б*) результат операции соединения по домену D_1 (табл. 4,*в*).

Таблица 4, *а*

R_5	D_1	D_2	D_3	D_4	D_5
	K5-01	ТЕОРИЯ АВТОМАТОВ	ПРОФ. ПЕТРЕНКО	03.ЯНВ.	АУД. 210
	K5-02	МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА	ПРОФ. ИВЧЕНКО	03. ЯНВ.	АУД. 211
	K5-03	ФИЗИКА	ПРОФ. КРАВЧЕНКО	05. ЯНВ.	АУД. 211
	K5-04	АЛГОРИТМИЧЕСКИЕ ЯЗЫКА	ПРОФ. ИВЧЕНКО	05. ЯНВ.	АУД. 210

Таблица 4, *б*

R_5	D_1	D_2	D_3	D_4	D_5
	K5-01	ФИЗИКА	ПРОФ. КРАВЧЕНКО	09. ЯНВ.	АУД. 210
	K5-04	МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА	ПРОФ. ИВЧЕНКО	10. ЯНВ.	АУД. 210
	K5-02	ТЕОРИЯ АВТОМАТОВ	ПРОФ. КРАВЧЕНКО	09. ЯНВ.	АУД. 211
	K5-03	АЛГОРИТМИЧЕСКИЕ ЯЗЫКА	ПРОФ. ИВЧЕНКО	10. ЯНВ.	АУД. 211

Таблица 4, *в*

R_5	D_1	D_2	D_3	D_4	D_5	D'_2	D'_3	D'_4	D'_5
	K5-04	ТЕО-РИЯ АВТО-МАТОВ	ПРОФ. ПЕТРЕ-НКО	03 ЯНВ.	АУД. 210	ФИЗИ-КА	ПРОФ. КРАВ-ЧЕНКО	09 ЯНВ.	АУД. 210
	K5-02	МАТЕ-МАТИ-ЧЕС-	ПРОФ. ИВЧЕ-НКО	03 ЯНВ.	АУД. 211	ТЕО-РИЯ АВТО-	ПРОФ. ПЕТРЕ-НКО	09 ЯНВ.	АУД. 211

		КАЯ ЛИНГ-ВИСТ-ИКА				МА-ТОВ			
	К5-03	ФИЗИ-КА	ПРОФ. КРАВ-ЧЕНКО	05 ЯНВ.	АУД. 211	АЛГО-РИТ-МИ-ЧЕС-КИЕ ЯЗЫКИ	ПРОФ. ИВЧЕН-КО	10 ЯНВ.	АУД. 211
	К5-04	АЛГО-РИТ-МИ-ЧЕС-КИЕ ЯЗЫКИ	ПРОФ. ИВЧЕ-НКО	05 ЯНВ.	АУД. 210	МАТЕ-МАТИ-ЧЕС-КАЯ ЛИНГ-ВИСТ-ИКА	ПРОФ. ПЕТРЕ-НКО	10 ЯНВ.	АУД. 210

Аналогично можно определить операцию соединения не только по условию «равенства», но и по другим условиям сравнения: $>,\ \geqslant,\ \neq,\ <,\ \leqslant$.

Определим, например, операцию соединения по условию «больше, чем» (>).

Соединением по условию «больше, чем» отношения R_a *по атрибуту* X *и отношение* R_b *по атрибуту* Y (атрибуты X и Y являются атрибутами одного и того же домена, общего для отношений R_a и R_b), $X > Y$, называется множество всех кортежей π_i таких, что π_i — конкатенация кортежа a_i, принадлежащего R_a, и кортежа b_i принадлежащего R_b, где X — часть a_i, Y — часть b_i и $X > Y$.

Запрос в реляционной базе данных будет выполнен тем быстрее, чем меньше операций над отношениями он содержит.

11.3. Отношения на базах данных и структурах данных

Как уже нами установлено, все вокруг определяется отношениями. Довольно лишь взять отношение s на переменных $(x_1, ..., x_n)$ так, чтобы можно было построить множество

$$\{(x_1, x_2, ..., x_n): s\ (x_1, x_2, ..., x_n) \text{ истинно}\}.$$

Пусть задан набор $(x_1, x_2, ..., x_n)$. Отношение $s(x_1, ..., x_n)$ можно *разрешить,* т.е. выяснить, $s(x_1, ..., x_n)$ истинно или ложно. Конечно, s не обязательно будет представлено «хорошей» формулой. Нетрудно показать, что вместо отношения s, определяющего множество наборов длины n, любое множество таких наборов также определяет отношение (и содержательные свойства s).

Эти два подхода эквивалентны.

Определение. При обработке данных наборы из n элементов называют *записями;* элементы этих наборов называют *полями.* Записи, определяющие отношение, обычно содержатся в файле. Если потребовать, чтобы несколько файлов содержали совокупность записей, которые удовлетворяют некоторым отношением, то мы получим (относительную) *базу данных.*

Замечание. Для случая обработки данных мы сейчас употребили термин «поле». В следующих разделах мы будем употреблять этот термин в математическом смысле, однако в данном случае это не приводит к недоразумению.

Таким образом, это дает нам первый реальный пример отношений, которые в большей мере связаны с вычислениями, в частности, с прикладными задачами. Тем не менее краткое обсуждение некоторых простейших свойств баз данных не только обеспечивает основу дальнейшего математического исследования отношений, но и проясняет некоторые факторы, понимание которых необходимо для эффективного управления системами баз данных.

Теория баз данных включает в себя изучение так называемых нормальных форм, однако обоснование некоторых из них очевидно лишь в простых случаях. Мы рассмотрим только три формы для следующих задач:

— вставить новый набор из n элементов;
— удалить набор из n элементов;
— модифицировать набор, который содержит n элементов.

Начнем с простейшей нормальной формы.

Определение. *Файлы в первой нормальной форме* (INF), или - более просто — *нормализованные файлы,* имеют записи фиксированной длины, которые состоят из элементов, взятых из множеств, чьи элементы далее не могут быть разбиты, и в каждый момент времени этот файл может быть представлен как массив значений $M \times N$. Каждая запись, будучи набором из n элементов, может быть записана как строка массива.

Пример 1. Рассмотрим отношение FAM1, в котором мы собрали вместе родителей и детей. Каждая запись содержит в указанном порядке фамилию и имена отца, матери и детей. Итак, мы имеем записи

(Шевчук, Иван, Нина , (Ира, Борис)) $\in$ FAM l,

(Кравчук, Петр, Надежда, (Люда)) $\in$ FAM l.

Теперь, если мы обозначим через F и M множества родителей и матерей, то из определения следует, что

Иван (Шевчук) $\in F$, Нина (Шевчук) $\in M$,

Петр (Кравчук) $\in F$, Надежда (Кравчук) $\in M$,

Таким образом, Люда является членом семьи Кравчук, но Ира и Борис не являются детьми семьи Кравчук. Так как в этой семье больше одного ребенка, то соответствующая запись больше, и, следовательно, нарушены условия первой нормальной формы.

Из FAM1 мы можем получить отношение FAM2, построив его из S, F, M и C, где S — множество фамилий, а C — множество детей, путем конструирования записей:

(Шевчук, Иван, Нина , Ира),

(Шевчук, Иван, Нина, Борис),

(Кравчук, Петр, Надежда, Люда).

Отношение FAM2 находится в 1NF и может быть представлено с помощью табл. 1.

Таблица 1

Фамилия	Отец	Мать	Ребенок
Шевчук	Иван	Нина	Ира
Шевчук	Иван	Нина	Борис
Кравчук	Петр	Надежда	Люда

Однако не совсем ясно, что будет, если, например, мужчина и женщина Юрчук не имеют детей? Если мы хотим иметь в файле запись о них, то следует пересмотреть структуру файла. Это означает, что все следует начать сначала. Введем следующую терминологию.

Определение. При использовании таблицы для изображения отношения (файла с n-мерными наборами/записями, которые записываются в виде строк) столбцы называются *атрибутами.*

Слкдовательно, ФАМИЛИЯ, ОТЕЦ, МАТЬ и РЕБЕНОК есть атрибутами различных полей в FAM2. Для получения доступа к записям в файле используются так называемые ключи. Более точно это может быть определено в сроках атрибутов.

Определение. Атрибут или (упорядоченное) множество атрибутов, чьи значения однозначно определяют запись в файле, называются *ключом* этого файла. (Заметим, что в файле может быть много различных ключей).

Каждый ключ отношения/файла FAM2 должен включать атрибут РЕБЕНОК.

Перейдем к другим примерам.

Пример 2. Каждый владелец компьютера должен покупать к нему запасные части. Поэтому мы можем рассмотреть файл SUP1, структура которого показана в табл. 2.

Таблица 2

КОМПАНИЯ	ОТДЕЛЕНИЯ	МЕНЕДЖЕР
ACCE	КИЕВ	КОРЖ
IBL	КИЕВ	ИВАНОВ
DATAMETZ	ЖИТОМИР	ИВАНОВ
PRINTACO	ОДЕССА	КОЧАН
WOOLIES	ЖИТОМИР	КОЧАН
RTX	КИЕВ	КОРЖ
OXONDATA	ЛЬВОВ	КАЗАК

Атрибут КОМПАНИЯ является ключом в SUP1; вся другая информация в файле доступна с помощью ключа. Таким образом, например, можно извлечь атрибут ОТДЕЛЕНИЕ с помощью ключа WOOLIES или же МЕНЕДЖЕР из RTX.

Определение. Если запись локализована с помощью некоторого ключа, то поле, которое выделяется из этой записи, называется *проекцией.* В данном контексте проекцией является «из». Будем также говорить, что эти атрибуты *зависят от ключа.*

На рис. 1 представлен графический пример зависимостей в SUP1.

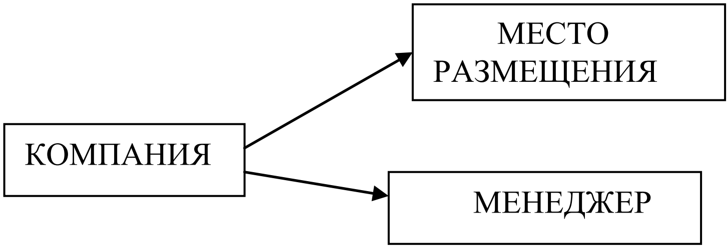


Рис. 1.

Пример 3. Модифицируем файл SUP1 с целью включения туда информации об имеющихся на складе запасных частях и об их количестве, которые отдельный поставщик желает продать в отдельной сделке. Включим в файл также код поставки, из которого мы можем выяснить скорость и частоту поставок. Во избежание лишней детализации введены номер компании и номер запасной части (табл. 3).

Таблица 3

КОМПАНИЯ	МЕСТО ПОЛОЖЕНИЯ	ПОСТАВЩИК	ЗАП. ЧАСТЬ	КОЛИЧЕСТВО
1	КИЕВ	2	1	10
1	КИЕВ	2	2	1
1	КИЕВ	2	3	10
2	КИЕВ	2	2	2
2	КИЕВ	2	4	5
3	РОССИЯ	4	2	2
3	РОССИЯ	4	3	10
3	РОССИЯ	4	4	4
5	РОССИЯ	4	1	10
5	РОССИЯ	4	3	10
6	КИЕВ	2	1	10

Из табл. 3 выясним, какие преобразования мы можем делать с SUP2, а какие нет:

а) *вставка*. Например, мы не можем прибавить в файл запись, которая указывает, что компания 4 (PRINTACO) находится в Одессе, без указания деталей, которые она может поставлять;

б) *удаление.* Если компания 6(RTX) прекратила поставки запчастей 1, тогда мы обязаны удалить все записи, которые относятся к этой компании и имеющие в поле ЗАП. ЧАСТЬ код 1;

в) *модификация*. Если код поставщика для Киева изменился, например, из-за транспорта, то соответствующее поле должно быть изменено в каждой записи, где есть код Киев в поле МЕСТО РАСПОЛОЖЕНИЕ. Что можно сделать для того, чтобы уменьшить или отодвинуть эти проблемы? С практической точки зрения мы должны выделить информацию в SUP2 так, чтобы по возможности избежать повторений. Таким образом, мы получаем возможность вставки/удаления части записи в SUP2. Возможнаое и разумное разделение дается в SUP3 (табл. 4). Тогда остаток информации в SUP2 может содержаться в поле ЗАП.ЧАСТЬ.

Таблица 4

SUP3

КОМПАНИЯ	МЕСТО РАСПОЛОЖЕНИЯ	ПОСТАВЩИК
1	КИЕВ	2
2	КИЕВ	2
3	РОССИЯ	4
5	РОССИЯ	4
6	КИЕВ	2

ЗАП. ЧАСТЬ

КОМПАНИЯ	ЗАП. ЧАСТЬ	КОЛИЧЕСТВО
1	1	10
1	2	1
1	3	10
2	2	2

Таблица 5

КОМПАНИЯ	МЕСТО РАСПОЛОЖЕНИЯ	ПОСТАВЩИК
1	КИЕВ	7
2	КИЕВ	7
3	РОССИЯ	4
4	НИВКИ	3
5	РОССИЯ	4
6	КИЕВ	7

ЗАП. ЧАСТЬ

КОМПАНІЯ	ЗАП. ЧАСТЬ	КОЛИЧЕСТВО
1	1	10
1	2	1
1	3	10
2	2	2

2	4	5
3	2	2
3	3	10
3	4	4
5	1	10
5	3	10
6	1	10

2	4	5
3	2	2
3	3	10
3	4	4
5	1	10
5	3	10

Используя эту конфигурацию, можно, например:

а) включить в SUP3 запись, которая означает, что компания 4 находится на Нивках (код поставщика 3);

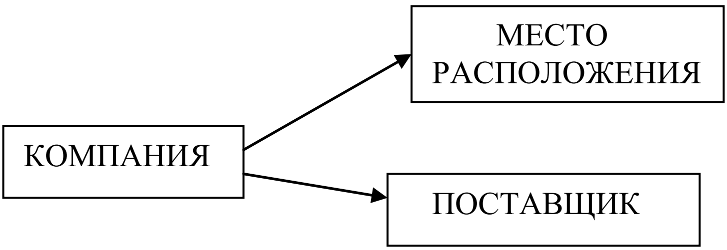
б) удалить ссылку на компанию 6 как на поставщика запчасти 1, но оставить соответствующий код в SUP3;

в) изменить код поставщика для КИЕВ на 7 путем замены только трех входов, которые отвечают компаниям с кодом 1, а не всех шести.

Результаты этих изменений приведены в табл. 5. Это уже значительно лучше, однако все же может быть еще усовершенствовано. Чтобы увидеть, в каком направлении продолжать исследование, отделим ключи и подчиненные части (рис. 2). Заметим, что ЗАП. ЧАСТЬ требует объединенного ключа.

Все не-ключи непосредственно связаны с ключом. Это дает нам следующее свойство нормальной формы.

Определение. Файл имеет *вторую нормальную форму* (2NF), если он имеет форму 1NF и неключевые атрибуты целиком независимы от ключа.



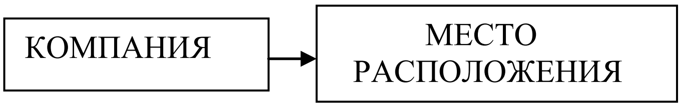
SUP 3

ЗАП. ЧАСТИ

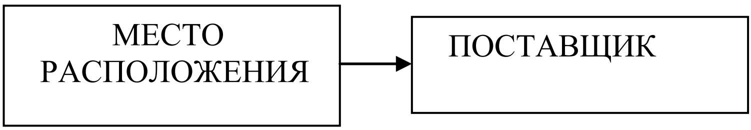
Рис. 2.

Пример 3 (продолжение). Файл SUP3 все еще является довольно сложным в том смысле, что для данной записи ПОСТАВЩИК может быть установлен с помощью исследования поля КОМПАНИЯ или же поля МЕСТО РАСПОЛОЖЕНИЯ. Это является причиной того, что в требовании а) код поставщика для Нивок должен быть вставлен перед записью кода 4 компании, а требование б), возможно, потребует изменить более одной записи для модификации единого поля данных, относящегося к коду поставщика. На практике мы можем убрать эту проблему путем проектирования SUP3 в SUP4 и DEL (табл. 6). (Заметим, что коды поставщика, изменяемые таким образом, препятствуют той возможности, что любые другие записи в файле вызывают противоречивую информацию. В SUP3 можно иметь запись вида «ПОСТАВЩИК КОМПАНИИ 6-2» и «ПОСТАВЩИК КОМПАНИИ 1-7» на некотором этапе модификации SUP2, несмотря на тот факт, что обе компании находятся в КИЕВЕ.)

Зависимость отношений в SUP4 и DEL изображена на рис. 3.



SUP4



DEL

Рис. 3.

Нетранзитивность отношения зависимости является внутренним свойством, из которого возникает понятие третьей нормальной формы.

Таблица 6.

КОМПАНИЯ	МЕСТО РАС-ПОЛОЖЕНИЯ
1	КИЕВ
2	КИЕВ
3	РОССИЯ
4	НИВКИ
5	РОССИЯ
6	КИЕВ

МЕСТО РАС-ПОЛОЖЕНИЯ	ПОСТАВЩИК
КИЕВ	7
РОССИЯ	4
НИВКИ	3

Определение. Файл находится в *третий нормальной форме* (3NF), если он является файлом 2NF, и каждый атрибут, который не является ключом, нетранзитивным образом зависит от ключа. Возможен и другой путь - каждый атрибут, который не является ключом, зависит только от ключа и ни от чего другого.

Как было отмечено ранее, существует много других «нормальных» форм, но мы не ставим изучения файлов своей целью в дальнейшем. Достаточно лишь иметь в виду, что информация в файлах является одной из реализаций математического понятия отношения. Практическое использование отношений SUP4 и DEL требует явной связи атрибута МЕСТО РАСПОЛОЖЕНИЯ файла SUP4 с атрибутом МЕСТО РАСПОЛОЖЕНИЯ файла DEL. Это - отношение эквивалентности (между компонентами различных файлов, имеющих одно и то же имя). Подобные отношения эквивалентности могут быть использованы для определения внутренних связей и других структурных данных. В качестве иллюстрации рассмотрим рис. 4.

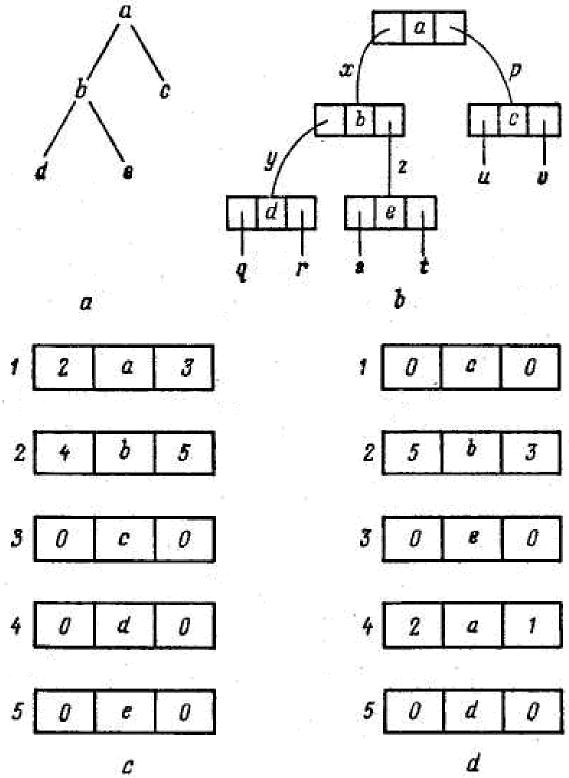


Рис. 4.

Диаграмма на рис. 4, *a* изображает дерево, диаграмма на рис. 4, *b* подобна диаграмме структурных данных, а диаграммы на рис. 4, *c* и *d* описывают их возможные применения. Отметим, что отношение эквивалентности различны, но результирующие структурные связи сохраняются.

С математической точки зрения это является разбиением на классы эквивалентности. Следовательно, мы можем определить произвольное представление этого дерева как элемент множества $T = D/E$, где

$$D=\{a=(x, a, p),\ b = (y, b, z),\ c = (u, c, v),\ d=(q, d, r),\ e=(s, e,t)\},$$

$$E = \{(x,b), (y,d), (p,c), (z, e)).$$

Эти вопросы обсуждались нами раньше.

11.4. Составные отношения

Подобно тому как мы устанавливали внутренние связи в файлах, для выделения некоторых данных из информации (например, ПОСТАВЩИК и МЕСТО РАСПОЛОЖЕНИЯ 3 посредством файлов DEL и SUP4 в примере 3) часто приходится связывать бинарные отношения друг с другом. Руководствуясь предыдущими рассуждениями, можно определить это понятие следующим образом.

Определение. Пусть заданы множества A, B и C и отношение σ между A и B и ρ между B и C. Определим отношение между A и C следующим образом: оно действует из A в B посредством σ, а потом из B в C посредством ρ. Такое отношение называют *составным* и обозначают $\rho \circ \sigma$, т.е.

$$(\rho \circ \sigma)(a)=\rho(\sigma(a)).$$

Следовательно, $(x, y) \in (\rho \circ \sigma)$, если существует $z \in B$ такое, что $(x, z) \in \sigma$ и $(z, y) \in \rho$. Отсюда следует, что $\mathcal{D}_{\rho\circ\sigma}=\sigma^{-1}\mathcal{D}_{\rho}$. Чтобы проиллюстрировать ситуацию, рассмотрим рис. 1.

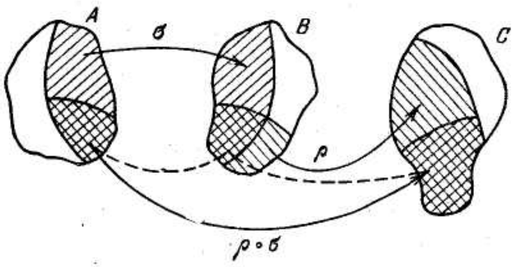


Рис. 1.

Области определения и значений σ и ρ заштрихованы в разных направлениях. Следовательно, сегменты с двойной штриховкой на A, B и C представляют собой $\mathcal{D}_{\rho\circ\sigma}$, $\mathcal{D}_{\rho} \square \mathcal{R}_{\sigma}$ и $\mathcal{R}_{\rho\circ\sigma}$, соответственно.

Замечание. Из записи отношений σ и ρ следует, что они применяются справа налево. Следовательно, $(\rho\circ\sigma)(a)$ означает, что вначале берется a и преобразуется посредством σ, а затем преобразуется посредством ρ. В алгебре это иногда записывают в виде $a\sigma\rho$. Следует обращать внимание при чтении других математических книг на то, какой порядок выполнения отношений приняты в той книге.

Пример 4. Пусть σ и ρ — отношения на N такие, что

$$\sigma = \{(x, x+1): x \in N\}, \quad \rho = \{(x^2, x): x \in N\}.$$

Тогда

$$\mathcal{D}_\rho = \{x^2: x \in N, \quad \mathcal{D}_\sigma = \{x: x, x + 1\ x \in N = N,$$

$$\mathcal{D}_{\rho\circ\sigma} = \sigma^{-1}\mathcal{D}_\rho = \{x: x \in N \text{ и } x + 1 = y^2\}, \text{ где } y \in N = \{3, 8, 15, 24, ...\}$$

(рис. 2).

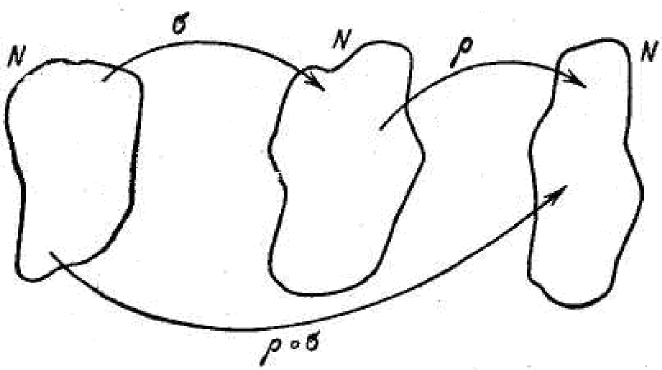


Рис. 2.

В случае, когда мы рассматриваем отношение на множестве, она может быть скомбинировано само с собой.

Например, используя отношение из примера 4, имеем

$$\sigma\circ\sigma = \{(x, x + 2): x \in N\} \text{ и } \rho\circ\rho = \{(x^4, x): x \in N\}.$$

Эти отношения можно также обозначать соответственно σ^2 и $\rho^{2.}$ В общем-то эти обозначения не совсем законы для множеств, однако их легко можно обосновать, поскольку если $(x, y) \in \sigma\circ\sigma$, то $((x, z), (z, y)) \in \sigma\times\sigma$ при некоторому z; никакого недоразумения при этом не возникает, поскольку известна структура получаемого результата.

Используя это обозначение, мы можем определить σ^n для любого $n \in N$, $n > 1$, следующим образом:

$$\Sigma^n = \{(x, y): x\sigma z \text{ и } z\sigma^{n-1}y \text{ для некоторого } z\}.$$

Если мы вновь возьмем отношение σ и ρ из примера 4, то получим

$$\Sigma^n = \{(x, x + n), x \in N\} \text{ и } \rho^n = \{\{x^{2^n}, x): x \in N\}.$$

Рассмотрим вопрос о том, насколько в этих случаях может быть применена аналогия с умножением. Пусть A — множество, а R — отношение на A. Тогда отношения $I_A\circ R$, R и $R\circ I_A$ эквивалентны; поэтому I_A является тождественным отношением на A, которое ведет себя подобно числу 1 по отношению к умножению чисел. Чтобы дополнить аналогию, желательно было бы иметь возможность писать $R^{-1}\circ R=I_A=R\circ R^{-1}$. Однако в общем случае этого делать нельзя. Для того чтобы иметь такую возможность, необходимо наложить дополнительные условия.

Замыкание отношений.

Понятие замыкания есть фундаментальным математическим понятием и используется в большинстве разделов математики. Чтобы проиллюстрировать это понятие, рассмотрим следующий пример.

Возьмем объект x_0 и процесс p, который порождает множество и определяет последовательность $x_1, x_2, \ldots \ldots, x_n, \ldots$ такую, что

$$x_1 \in p(x_0),$$

$$x_2 \in p(x_1),$$

$$\ldots\ldots\ldots\ldots$$

$$x_n \in p(x_{n-1}),$$

$$\ldots\ldots\ldots\ldots$$

Множество, которое содержит все элементы всех последовательностей, которые могут быть выведены при помощью p, и начинающиеся с x_0, называется *замыканием процесса p* относительно x_0. Поэтому «ответ» будет содержаться в $p^n(x_0)$ при каком-то n. Однако мы не знаем заранее значения n. Более того, если мы возьмем произвольный элемент y из этого замыкания и выполним процесс p, начиная с y, то не получим ничего нового. Результат уже содержится в замыкании. Множество не может быть расширено таким путем (оно замкнута).

Пример 5. Возьмем квадрат S, размеченный, как это показано на рис. 3, и определим процесс r следующим образом. Из заданного положения S процесс r порождает множество всех положений, получаемых в результате поворота по часовой стрелке на прямой угол.

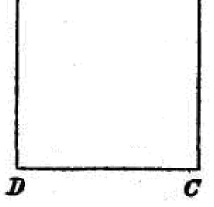


Рис. 3.

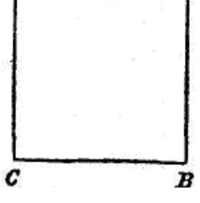
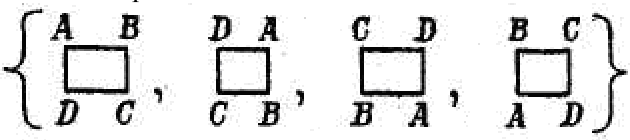


Рис. 4.

Таким образом, $r(S)$ дает конфигурацию, которая изображена на рис. 4. После применения r четыре раза, мы вернемся к положению, с которого начали, и, следовательно, замыкание в данном случае есть множество из четырех позиций.



Рассмотрим теперь, что произойдет, если процесс определить с помощью отношения. (В действительности это всегда возможно, потому что мы можем определить подходящее отношение при помощи множества $\{(x, \; y) : y \in p(x)$, где p — исследуемый процесс).) Для построения замыкания отношения A достаточно иметь составные отношения $A, A^2, \dots, A^n, \dots$, которые затем комбинируются обычным теоретико-множественным путем.

Определение. *Транзитивным замыканием* (или просто *замыканием)* отношения A на множестве называется бесконечное объединение

$$\bigcup_{n=1}^{\infty} A^n = A \cup A^2 \cup A^3 \cup \dots$$

Транзитивность замыкания отношения следует, очевидно, из его определение, однако слово «транзитивное» часто включают, чтобы подчеркнуть различие между этой и подобной ей операцией, которая в скором времени будет определена. Транзитивное замыкание отношения A обозначают A^+.

Пример 6.

1. Пусть R— отношение на N такое, что

$$R = \{(x, y) : \; y = x+1\}.$$

Тогда

$$R^+ = \{(x, y) : \; x < y\}.$$

2. Пусть σ — отношение на Q такое, что

$$\sigma = \{(x, y) : x < y\}.$$

Тогда

$$\sigma^+ = \sigma.$$

3. Пусть ρ — отношение на Q такое, что

$$\rho = \{(x, y): x * y = 1\}.$$

Тогда

$$\rho^{+}\{(x, x): x \neq 0\} \square \rho.$$

4. Пусть L — множество станций Киевского метро и a, b и c - последовательные станции. Если отношение N на L определено как N={(x, y): x есть следующей за y станцией), то (a, b) и $(b, c) \in N$ и (a, a), (b, b), (c, c) и $(a, c) \in N^2$. Следовательно $N^+ = U_L = L \times L$.

Из этих примеров легко видеть, что замыкание отношения в общем случае не является рефлексивным. Однако иногда удобно сделать его таким. Это можно легко сделать. Вначале мы примем допущение, что тождественное отношение на X, $I = \{(x, x): x \in X\}$ является нулевой степенью произвольного отношения на X. Таким образом, $A^0 = I$ для любого A.

Определение. *Рефлексивным замыканием* A^* отношения A называют множество

$$A^* = \bigcup_{n=0}^{\infty} A^n$$

Замыкания отношений связаны между собой очевидным соотношением

$$A^* = A^+ \cup И.$$

Пример 7. Используя отношения, которые определены в предыдущих примерах, получаем

$$R^* = \{(x, y): x \leq y\},\ \sigma^* = \{(x, y): x \leq y),$$

$$\rho^* = \rho^+ \cup \{(0, 0)\},\ N^* = N^+.$$

11.5. Реляционная модель данных

Основные понятия. В основе реляционной модели данных (РМД) лежит математическая теория отношений. Этим определяется и название модели (RELATION - отношение). Отношение служит средством структуризации данных.

Таким образом, представив n-местное отношение в виде таблицы, тем самым определенным образом структурируем данные. Поэтому, естественно, подобные преобразования называют реляционной структурой или реляционным типом. Массив данных, представленный набором реляционных структур, образует реляционную БД, и схема реляционной БД будет представлена набором схем отношений:

$$R_1(A^1_1, A^1_2,..., A^1_k);$$
$$R_2(A^2_1, A^2_2,..., A^2_l);$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$R_m(A^m_1, A^m_2,...,A^m_n),$$

где A_i^j — имя атрибута, R_j— имя отношения.

Ограничение модели. Число ограничений в реляционной модели невелико, что обеспечивает достаточную свободу в выборе представления типов связей и сущностей. Основным ограничением является невозможность представления в отношении дубликатов строк. Это ограничение позволяет уточнить понятие ключа отношения. В РМД ключ определяется как подмножество атрибутов, которые позволяют однозначно идентифицировать кортеж. Так как дубликаты строк в отношении запрещаются, то это означает, что каждое отношение имеет, по крайней мере, один ключ (состоящий из всех атрибутов). Отношение может иметь и несколько ключей, называемых возможными ключами. Один из возможных ключей выбирается в качестве **первичного.** Следует иметь в виду, что первичный ключ не разрешается обновлять и никакой из его компонентов не может принимать значения «неопределено».

Второе ограничение модели состоит в том, что порядок столбцов в таблице является значимым. Пренебрегать упорядочением столбцов можно только в том случае, если каждому столбцу присвоено уникальное имя.

На значения атрибутов в модели можно задавать ограничения в явном виде. Большинство явных ограничений, которые встречаются на практике, это ограничения на зависимости между атрибутами. Явное задание ограничений обеспечивает возможность самостоятельного исследования зависимостей между атрибутами.

Зависимость отбивает тот факт, что один объект зависит от другого. В РБД в качестве таких объектов рассматриваются либо аргументы, либо кортежи. Одним из основных типов зависимостей, рассматриваемых в реляционных БД, являются **функциональные зависимости.**

Рассмотрим формальное определение функциональной зависимости, принимая следующие обозначения. Большими буквами $A,B,C,\ldots$

обозначим одиночные атрибуты, X, Y, Z — множество атрибутов, а a, b, c,... и x, y, z,... — соответствующие им значение. Большие буквы R, S будем применять для обозначения отношений.

Предположим, что существует некоторое **универсальное отношение** U, в котором каждый атрибут имеет уникальное имя. При этом будем считать, что множество атрибутов любого другого отношения представляет собой некоторое подмножество атрибутов универсального отношения U.

Пусть A и B атрибуты отношения R. Тогда говорят, что атрибут B отношения R функционально зависит от атрибута A, если в каждый момент времени каждому значению a соответствует не более одного значения b. Функциональную зависимость f атрибута B от атрибута A обозначают: $f{:}A \to B$.

Эту зависимость f можно также представить множеством упорядоченных пар $\{<a, b> | a \in A, b \in B\}$, в которых каждому значению a соответствует только одно значение b. При этом говорят, что B **функционально зависит** (или просто зависит) от A, а A **функционально определяет** (или **определяет**) B.

Если существует единственная функциональная зависимость B от A, то ее обозначают просто $A \to B$. В случае отсутствия между ними функциональной зависимости вводят обозначения $A \not\to B$.

Если $A \to B$ и одновременно $B \to A$, то между A и B существует взаимно-однозначное соответствие, которое записывается как $A \leftrightarrow B$.

Пусть: $f{:}A_1, A_2,\dots ,A_n \to B$ и $g{:}A_1, A_2,\dots, A_m \to B$, где $m<n$. Так как атрибуты $A_1, A_2,\dots, A_m$ функционально определяют B, то $A_{m+1}, A_{m+2},\dots ,A_n$ называют посторонними в f. В этом случае B неполно зависит от $A_1, A_2,\dots ,A_n$. Если для данного f не существует g с вышеуказанными свойствами, т.е. левая часть f не содержит посторонних атрибутов, то говорят, что реализуется полная функциональная зависимость.

Пусть есть множество атрибутов $A_1, A_2,\dots ,A_n$ отношения R, а также множество F функциональных зависимостей $X \to Y$, где X и Y — подмножества атрибутов множества $A_1, A_2,\dots ,A_n$. Тогда из функциональных зависимостей, которые входят в F, могут быть выведены другие функциональные зависимости, присущие отношению R.

Обозначим через F^+ замыкание множества функциональных зависимостей F, т.е. полное множество зависимостей, которую можно получить из F. Множество зависимостей F^+ можно построить из F на основе следующих правил вывода функциональных зависимостей (ФЗ):

1. ФЗ1 (свойство рефлексивности).

2. ФЗ2 (свойство пополнения).
3. ФЗ3 (свойство транзитивности).

Это полный набор правил, т.е. он позволяет по заданному множеству F определить полное множество функциональных зависимостей F^+, присущий рассматриваемой схеме отношений $R(A_1, A_2, \ldots, A_n)$.

Рассмотрим эти правила более подробно.

Правило ФЗ1 (свойство рефлективности). Если $X \subseteq U$, $Y \subseteq U$ и $Y \subseteq X$, то имеет место функциональная зависимость $X \rightarrow Y$. Например, задано отношения $U(A_1, A_2, A_3, A_4, A_5)$.

Рассмотрим два множества атрибутов:

$$X = \{A_1, A_2, A_3, A_4\} \text{ и } Y = \{A_1, A_4\}.$$

Исходя из свойства транзитивности, можно сказать, что существует функциональная зависимость $X \rightarrow Y \in F^+$. Данное правило говорит о том, что, имея исходную функциональную зависимость $A_iA_jA_kA_n \rightarrow A_iA_j$, можно в состав множества атрибутов левой части выражения вводить любые атрибуты из множества U. При этом функциональная зависимость будет сохраняться. Можно также в правую часть включать атрибуты которые уже расположенны в левой части.

Правило ФЗ2 (свойство пополнения). Если $X \subseteq U$, $Y \subseteq U$, $Z \subseteq U$ и имеет место функциональная зависимость $X \rightarrow Y$, то $X \cup Z \rightarrow Y \cup Z$. В отличие от правила ФЗ1 данное говорит о том, что для его применения не существенно выполнение условий $Y \subseteq X$. Т.е. любые атрибуты из множества U можно одновременно подставлять в левую и правую части выражения функциональной зависимости F.

Например, имеется универсальное отношение $U(A_1, A_2, A_3, A_4, A_5)$ и заданы наборы атрибутов $X = \{A_1, A_3\}$, $Y = \{A_2, A_4\}$, $Z = \{A_5\}$. Тогда из условия, что существует функциональная зависимость $X \rightarrow Y$:

$$\{A_1, A_3\} \rightarrow \{A_2, A_4\}$$

следует, что имеет место зависимость

$$\{A_1, A_3, A_5\} \rightarrow \{A_2, A_4, A_5\}.$$

Правило ФЗ3 (свойство транзитивности). Если $X \subseteq U$, $Y \subseteq U$, $Z \subseteq U$ и имеют место зависимости $X \rightarrow Y$ и $Y \rightarrow Z$, то $X \rightarrow Z$. Например, имеются подмножества атрибутов $X = \{A_1, A_3\}$, $Y = \{A_2, A_4\}$, $Z = \{A_5\}$. Тогда из условия существования зависимостей $\{A_1, A_3\} \rightarrow \{A_2, A_4\}$, $\{A_2, A_4\} \rightarrow \{A_5\}$ следует, что имеет место зависимость $\{A_1, A_3\} \rightarrow \{A_5\}$.

Кроме этих правил часто используют дополнительные правила следствия ФЗ1, ФЗ2 и ФЗ3.

Правило ФЗ4 (свойство расширения). Если $X \subseteq U$, $Y \subseteq U$ и задана функциональная зависимость $X \rightarrow Y$, то тогда для любого $Z \subseteq U$ имеет место функциональная зависимость $X \cup Z \rightarrow Y$.

Правило Ф35 (свойство продолжения). Если $X \subseteq U$, $Y \subseteq U$, $W \subseteq U$, $Z \subseteq U$ и задана функциональная зависимость $X \to Y$, то для любых $W \subseteq Z$ имеет место зависимость $X \cup Z \to Y \cup W$.

Правило Ф36 (свойство псевдотранзитивности). Если $X \subseteq U$, $Y \subseteq U$, $W \subseteq U$, $Z \subseteq U$ и заданы функциональные зависимости $X \to Y$, $Y \cup W \to Z$, то имеет место функциональная зависимость $X \cup W \to Z$.

Правило Ф37 (свойство аддитивности). Если $X \subseteq U$, $Y \subseteq U$, $Z \subseteq U$ и заданы функциональные зависимости $X \to Y$, $X \to Z$, то имеет место функциональная зависимость $X \to Y \cup Z$.

Правило Ф38 (свойство декомпозиции). Если $X \subseteq U$, $Y \subseteq U$, $Z \subseteq U$ и при этом $Z \subseteq Y$ и задана функциональная зависимость $X \to Y$, то будет иметь место функциональная зависимость $X \to Z$.

Очевидно, что даже при небольшом числе зависимостей F число функциональных F^+ должно быть довольно большое.

Функциональная зависимость играет важную роль в моделировании данных. Вместе с тем можно отметить, что это отнюдь не общий вид зависимости. Важную роль играет также многозначная зависимость.

Многозначные зависимости. Рассмотрим отношение $R(X, Y, Z)$, где X, Y, Z — множество атрибутов. Кортеж отношения $R(X, Y, Z)$ обозначим через $<x, y, z>$. Если в отношении $R(X, Y, Z)$ присутствуют кортежи $<x, y, z>$; $<x, y', z>$;... $(x, y, z'>$; $<x, y', z'>$, то говорят, что существует многозначная зависимость атрибутов Y и Z от X. Эта зависимость обозначается $X \to\to Y$.

Пример. Рассмотрим таблицу, в которой указанные школьники-победители олимпиад по предметам:

№ п/п	Ф. И. О. школьника	Школа №	Предмет
1	Давыдов В. А.	154	Математика
2	Давыдов В. А.	154	Физика
3	Кулешов К. Г.	820	Биология
4	Жукова А. М.	154	Математика
5	Жукова А. М.	154	Физика

В нашем примере имеют место две многозначные зависимости:

ШКОЛА №→→ПРЕДМЕТ (например, № 154→ → (математика, физика), № 820→→ (биология)) и ШКОЛА № →→Ф. И. ШКОЛЬНИКА

Другими словами, X многозначно определяет Y, если и только если множество $Y=\{y|(x, y, z) \in R\}$ определяется только X.

Многозначную зависимость определяют путем следующей проверки.

Если для двух кортежей t и s отношения *R(X, Y, Z)* справедливо первое условие:

$$t[X]= s\ [X]$$

т.е. t и s совпадают по значениям атрибутов X и существует третий кортеж *u,* такой, что выполняется второе условие:

$$u[X, Y]=t[X, Y]; \quad u[Z]=s[Z],$$

то существует многозначная зависимость.

Пример. Рассмотрим отношение ПОБЕДИТЕЛИ ОЛИМПИАДЫ (Ф. И. О. ШКОЛЬНИКА, ШКОЛА №, ПРЕДМЕТ).

Проверим на основании вышеприведенного условия многозначную зависимость ШКОЛА №→→ПРЕДМЕТ

1) t [№154]= s [№154].

Можно отметить, что существует четыре кортежа, у которых значение атрибута ШКОЛА И (№ 154) совпадает, т.е. первое условие выполняется. В качестве t и s возьмем, например, кортежи

t [Давыдов В. А. № 154 математика],

s [Жукова А. М. № 154 физика].

Рассмотрим, существует ли такой кортеж *u,* для которого будет выполняться и второе условие, определяющнн многозначную зависимость.

Можно выделить кортежи, для которых выполняется соотношение *u*[№ 154, математика] =t [№ 154, математика]; это следующие кортежи:

t [Давыдов В. А., № 154, математика],

u [Жукова А. М., № 154, математика],

т.е. первому уравнению во втором условии - кортежи удовлетворяют. Анализируя таблицу, заметим, что для того, чтобы существовала многозначная зависимость ШКОЛА №→→ПРЕДМЕТ, должно выполняться условие

u[Жукова А. М.]=s [Жукова А. М.],

что действительно имеет место.

Другой способ проверки многозначной зависимости $X\to\to Y$ на отношение *R(X, Y, Z)* может быть осуществлен на основе проверки существования кортежа $<x, y, z>$ при условии, что существуют кортежи $<x, y, z'>$ и $<x, y', z>$.

Следует иметь в виду, который в общем случае формальная проверка должна выполняться на множестве всех возможных экземпляров кортежей отношения.

Правила вывода многозначных зависимостей сходны с правилами вывода функциональных зависимостей:

МЗ1 (дополнение). Если $X \subseteq U$, $Y \subseteq U$, $X \to\to Y$, то имеет место ногозначительная зависимость $X \to\to U - X - Y$.

МЗ2 (присоединение). Если $X \subseteq U$, $Y \subseteq U$, $Z \subseteq W$, $X \to\to Y$, то имеет место зависимость $X W \to\to YZ$.

МЗ3 (транзитивность). Если $X \subseteq U$, $Y \subseteq U$, $X \to\to Y$, $Y \to\to Z$, то $X \to\to Z — Y$.

Существуют также правила вывода для совокупности ФЗ и МЗ:

ФМ1. Если $X \subseteq U$, $Y \subseteq U$, $X \to Y$, то $X \to\to Y$.

ФМ2. Если $X \subseteq U$, $Y \subseteq U$, $Z \subseteq U$, $W \subseteq U$ и W не пересекается с Y (т.е. W | Y=0), и $X \to\to Y$, $W \to Z$, то имеет место зависимость $X \to\to$Z.

ФМ3. Если $X \to\to Y$, $XY \to Z$, то $X \to Z — Y$.

Функциональная и многозначная зависимости являются свойствами, существующими между двумя атрибутами (множествами). С их помощью осуществляют декомпозиции (разбиение) отношения на два или восстанавливают исходное отношение, соединяя два отношения (рис. 1).

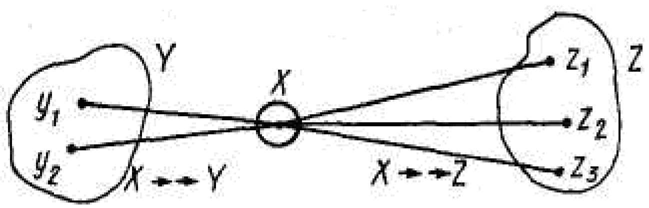


Рис. 1. Пример изображения многозначных зависимостей

Одним из основных вопросов любой организации данных есть вопросы сохранения жизнеспособности прикладных программ при изменениях в базе данных для того, чтобы избежать трудностей поддержания базы данных и назначения ключей для каждого отношения. Рассмотрим отношение $R\{A_1, A_2, ..., A_n\}$. Возможный ключ k отношения R — это комбинация атрибутов (возможно, состоящих из одного атрибута), обладающих следующими свойствами.

1. В каждом кортеже отношения R величина k единственным образом определяет этот кортеж.

2. Не существует атрибута в возможном ключе k, который мог бы быть удален без нарушения свойства 1.

Всегда существует, по крайней мере, один возможный ключ, т.е. комбинация всех атрибутов R удовлетворяет свойству 1.

Если в отношении R имеется несколько возможных ключей, то один из них выбирается в качестве первичного.

Атрибут A_i отношения R называется также первичным, если он входит в состав любого ключа (возможного или первичного) отношения.

Нормальные формы схем отношений. Рассмотрим четыре уровня нормализации схем отношений и соответственно четыре нормальные формы отношений: 1НФ, 2НФ, 3НФ, 4НФ. Их взаимное отношение можно представить в виде, представленном на рис. 2.

Из рис. 2 следует, что отношения являются как бы вложенными друг в друга по возрастанию номеров.

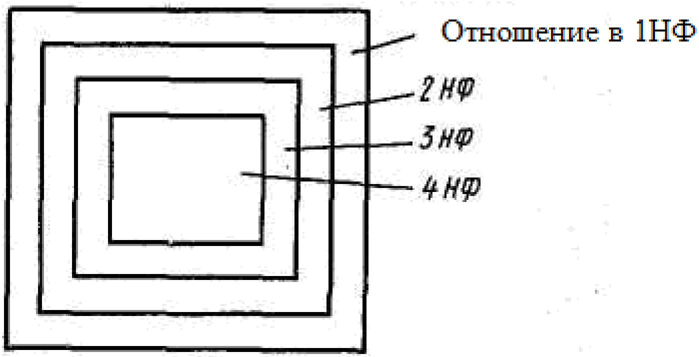


Рис. 2. Взаимное отношение нормальных форм

Так, например, если отношение находится в 4НФ, то оно будет удовлетворять и условиям 3НФ, 2НФ, 1НФ.

Отношение находится в **первой нормальной форме** (1НФ), если каждый атрибут отношения есть простым (атомарным) атрибутом, т.е. отсутствуют составные. Рассмотрим схему отношения:

АВТОМОБИЛЬ (МОДЕЛЬ, МАРКА, МОЩНОСТЬ, СТОИМОСТЬ, ИЗГОТОВИТЕЛЬ (НАЗВАНИЕ ЗАВОДА, ГОРОД)).

В этом случае атрибут ИЗГОТОВИТЕЛЬ составной. Для приведения к 1НФ отношения необходимо избавиться от составного отношения - ИЗГОТОВИТЕЛЬ. Этого можно добиться, рассматривая вместо составного атрибута его составляющие:

АВТОМОБИЛЬ (МОДЕЛЬ, МАРКА МОЩНОСТЬ, СТОИМОСТЬ, НАЗВАНИЕ ЗАВОДА ИЗГОТОВИТЕЛЯ, ГОРОД).

Приведение отношения к 1НФ достаточно для реализации языков запросов.

Чтобы рассмотреть 2НФ, введем понятие **полной зависимости**. Пусть X и Y — подмножества атрибутов отношения R и $X \to Y$. Если Y функционально не зависит от любого подмножества A множества X (причем A не совпадает с X), то Y называется полностью зависимым от X в R. Тогда говорят, что отношение R находится **во второй нормальной форме** (2НФ), если оно нормализовано, т.е. находится в первой нормальной форме, и каждый непервичный атрибут полностью зависит от первичного ключа.

Отношение R находится **в третий нормальной форме** (3НФ), если оно находится во второй нормальной форме и каждый непервичный атрибут в отношении R не содержит транзитивных зависимостей от первичного ключа. Транзитивная зависимость наблюдается в R, если существует такой атрибут A, что $X \to Y$ и $Y \to A$, $Y \leftrightarrow X$, где X — ключ.

Можно заметить, что 2НФ и 3НФ накладывают ограничения на зависимости только в связи с непервичными атрибутами. Поэтому в рассмотрение вводят **усиленную третью нормальную форму** (или нормальную форму Бойса-Кодда).

Отношение R находится в усиленной третий нормальной форме, если для всех зависимостей $X \to A$, когда A не принадлежит X, X является возможным ключом отношения R. Обычно атрибут, от которого функционально полно зависит другой, называют **детерминантой**. Поэтому говорят, что отношение R находится в усиленной третий нормализованной форме, если все детерминанты являются ключами.

Можно отметить следующее отличие этих нормальных форм по информационному содержанию. Вторая нормальная форма более информативна, чем первая, а третья более информативна, чем вторая.

Поэтому 3НФ и усиленная 3НФ больше понятны пользователю и приспособленны к реализации.

Существует также нормальная форма, которая учитывает многозначные зависимости, ее называют **четвертой нормальной формой** (4НФ). Отношение R находится в четвертой нормальной форме, если каждый раз, когда существует многозначная зависимость $X \to \to Y$ (где $Y \neq \varnothing$, $Y \not\subset X$ и XY состоит не из всех атрибутов R), также существует зависимость $X \to A$ для любого атрибута A в R.

Получение отношений нормальной формы достигается декомпозицией их схем. Под декомпозицией схемы отношения $R=\{A_1, A_2,..., A_n\}$ понимается замена схемы совокупностью отдельных схем $\rho = \{ R_1, R_2,...,R_k\}$ таких, что $R_1 \cup R_2 \cup ... \cup R_n = R = \{A_1, A_2,..., A_n\}$. При этом не накладывается никаких ограничений на пересечение любых двух элементов совокупности ρ.

Для получения нужных данных из базы разрабатываются специальные языки манипулирования данными, обеспечивающие выполнение необходимых операций. Для разработки и исследования языков манипулирования данными используют математический аппарат, основанный на следующих трех подходах:

реляцинная алгебра;

реляцинное исчисление с переменными-кортежами;

реляцинное исчисление с переменными -атрибутами.

Реляционная алгебра. В ней определяются основные операции над данными реляцинного типа. Все операции, вводимые в реляцинной алгебре, можно разделить на традиционные над множествами и специализированные, вводимые для удобства поиска в БД.

К операциям первой группы относятся: объединение, пересечение, разность, декартово произведение. К операциям второй группы следует отнести: проекцию, ограничение, соединение, деление.

Объединение. В результате применения этой операции получается отношение, объединяющее кортежи, содержащиеся в исходных отношениях. Пусть имеем два исходных отношения R_1 и R_2. Операция объединения этих отношений будет обозначаться $R_1 \cup R_2$:

$$R_1 \cup R_2 = \{r \mid r \in R_1 \ \text{або} \ r \in R_2\}.$$

Объединяемые отношения должны иметь одинаковые атрибуты (должны быть объединимыми):

Пример.

R_1:

A	B	C
a_1	b_1	c_1
a_2	b_2	c_2
a_3	b_3	c_3

R_2:

A	B	C
a_3	b_4	c_2
a_4	b_5	c_3
a_2	b_2	c_2

$R_1 \cup R_2$

A	B	C
a_1	b_1	c_1
a_2	b_2	c_2
a_3	b_3	c_3
a_3	b_4	c_2
a_4	b_5	c_3
a_2	b_2	c_2

Пересечение. В данной операции (которая обозначается I) получают отношение, включающее кортежи, общие для R_1 и R_2:

$$R_1 \cap R_2 = \{r \mid r \in R_1 \;\; i \;\; r \in R_2\}.$$

Пример. Для отношения R_1 и R_2 из предыдущего примера получим

$R_1 \cap R_2$:

A	B	C
a_2	b_2	c_2

Разность. В результате применения этой операции $(R_1 \backslash R_2)$ получается отношение, которое содержит кортежи, являющиеся кортежами отношения R_1 и не являющиеся кортежами отношения R_2:

$$R_1 \setminus R_2 = \{r \mid r \in R_1 \;\; \text{и} \;\; r \notin R_2\}.$$

Пример.

$R_1 \setminus R_2$:

A	B	C
a_1	b_1	c_1
a_3	b_3	c_3

Декартово (прямое) произведение. В этой операции ($R_1 \times R_2$) из m-местного отношения R_1 и n-местного отношения R_2 получают $(m+n)$-местное отношение. Причем первые m элементов представляют кортежи из отношения R_1, а последние n элементов - кортежи из отношения R_2:

$$R_1 \times R_2 = \{\langle r_1, \; r_2, \rangle \mid r_1 \notin R_1 \;\; \text{и} \;\; r_2 \in R_2\}.$$

Пример.

$R_1 \times R_2$:

A	B	C	A	B	C
a_1	b_1	c_1	a_3	b_4	c_2
a_2	b_2	c_2	a_3	b_4	c_2
a_3	b_3	c_3	a_3	b_4	c_2
a_1	b_1	c_1	a_4	b_5	c_3
a_2	b_2	c_2	a_4	b_5	c_3
a_3	b_3	c_3	a_4	b_5	c_3
a_1	b_1	c_1	a_2	b_2	c_2
a_2	b_2	c_2	a_2	b_2	c_2
a_3	b_3	c_3	a_2	b_2	c_2

Проекция. Операция проекции предназначена для изменения числа столбцов в отношении, т.е. в том случае, когда из строк-кортежей требуется исключить какие-либо атрибуты.

Обозначим через j_1, j_2, ..., j_n — номера столбцов n-местного отношения R. Операцию определения проекции отношения R обозначим через $\pi_{j_1, j_2, \ldots, j_n}(R)$, а сама операция заключается в том, что из отношения R выбираются столбцы и компонуются в указанном порядке j_1, j_2, ..., j_n.

Пример. Рассмотрим отношение ТЕЛЕВИЗОР.

Наименование	*Индекс*	*Диагональ кинескопа, см*	*Цена, руб*
Рекорд	ВЦ-311	47	640
Рубин	Ц-266Д	67	1040
Темп	Ц-280Д	61	755
Электроника	Ц-283	61	755
Горизонт	Ц-355	57	610

Тогда можно получить проекцию $\pi_{1,4}$ (ТЕЛЕВИЗОР):

Наименование	Цена, руб.
Рекорд	640
Рубин	1040
Темп	755
Электроника	755
Горизонт	610

$\pi_{4,1,3}$ (ТЕЛЕВИЗОР):

Цена, руб.	Наименование	Диагональ кинескопа, см
640	Рекорд	47
1040	Рубин	67
755	Темп	61
755	Электроника	61
610	Горизонт	51

π_4 (ТЕЛЕВИЗОР):

Цена, руб.
640
1040
755
755
610

Можно отметить, что в последней проекции оказались одинаковые строки.

Ограничение. Ограничением называют такую операцию, в которой отношение исследуют по строкам и выделяют множество строк, которые удовлетворяют заданным условиям.

Обозначим через r строку (кортеж) отношения R, а через θ определим одно из отношений: $=$, $\neq$, $<$, $\leq$, $>$, $\geq$. Тогда θ-ограничение между избранным атрибутом A в отношении R и некоторой величиной C определяют следующим образом: $R[A\theta C]=\{r|r\in R$ и $r[A]\theta C\}$, где $r[A]$ соответствует значению атрибута A в строке r.

Таким образом, θ-ограничение обеспечивает получение среди строк отношения R только тех строк, в которых значение атрибута A и значение величины C удовлетворяют условию сравнения θ.

Например, нужно из отношения ТЕЛЕВИЗОР выделить те марки телевизоров, которые имеют стоимость меньше 700 руб.

ТЕЛЕВИЗОР [цена < 700]:	РЕКОРД	ВЦ-311	47	640
	ГОРИЗОНТ	Ц-355	51	610

Следует отметить, что в частном случае в операции сравнения вместо величины *C* можно использовать другой домен *B*. Тогда операция ограничения определяется как

$$R(A\theta B)=\{r\,|\,r\in R \text{ и } r(A)\theta r(B)\}$$

Пример. Пусть отношение *R* имеет вид:

$$R:\quad \begin{array}{cccc} (A, & B, & D, & E) \end{array}$$
$$\begin{bmatrix} a & 15 & 17 & p \\ k & 22 & 28 & q \\ e & 39 & 11 & e \\ m & 16 & 43 & m \end{bmatrix}$$

Так как операция θ-ограничение предусматривает выбор среди строк отношения *R* только тех, в которых значениЯ атрибутов *A* и *B* удовлетворяют условию сравнения θ, то для условий (*B*<*D*) и *(A= E)* получим следующие отношения:

$$R\,[B<D]:\quad \begin{array}{cccc} A & B & D & E \end{array}$$
$$\begin{bmatrix} a & 15 & 17 & p \\ k & 22 & 28 & q \\ m & 16 & 43 & m \end{bmatrix}$$

$$R\,[A=E]:\quad \begin{array}{cccc} A & B & D & E \end{array}$$
$$\begin{bmatrix} e & 39 & 11 & e \\ m & 16 & 43 & m \end{bmatrix}$$

Соединение. Операция соединения обратна операции проекции. Рассмотрим для простоты два бинарных отношения $R_1(A, B)$ и $R_2(B, C)$. Соединением (обозначается $><$) отношений R_1 и R_2 $(R_1 >< R_2)$ называют операцию, при которой соединяют два отношения, используя в качестве признака соединения общий атрибут *Y:*

$$R_1 \bowtie R_2=\{\langle A,\ \mathrm{B},\ C\rangle\,|\,\langle A,\ B,\rangle\in R_1 \text{ и } \langle B,\ C,\rangle\in R_2\}.$$

Отношение $R_1 >< R_2$ является отношением с атрибутами <*A, B, C*>.

Например:

R_1: Наименование	Индекс	R_2: Индекс	Цена
Рекорд	ВЦ-311	ВЦ-311	640
Темп	Ц-280Д	Ц-280Д	755
Электроника	Ц-283	Ц-283	755
Горизонт	Ц-355	Ц-355	610

Отношение $R_1 >< R_2$ будет иметь вид:

Наименование	Индекс	Цена
Рекорд	ВЦ-311	640
Темп	Ц-280	755
Электроника	Ц-283	755
Горизонт	Ц-355	610

Операция соединения соответствует случаю, когда просто стыкуются таблицы отношений. Но поскольку в получающейся таблице атрибуты с одинаковым содержанием присутствуют дважды, один из одинаковых столбцов исключают. Такую операцию называют естественным соединением. Операцию соединения можно использовать не только для бинарных, но и для n-местных отношений:

R^1 >< R_2 >< ,..., >< $R_n=\{M|M$ — кортежи, которые образованы соединениям атрибутов отношений $R_1, R_2,..., R_n\}$.

Отметим, что рассматривались в основному операции соединения фактически по условию равенства двух атрибутов в двух отношениях. В общем случае соединение можно осуществлять не только по условию равенства, но и по любой другой θ: =, ≠, <, ≤, >, ≥:

$$R_1 \underset{A\theta B}{\bowtie} R_2 = \{(r,\ s) \mid r \in R_1 \text{ и } s \in R_2 \text{и } (r(A)\theta r(B))\},$$

когда θ - оператор равенства, то операцию называют эквисоединением.

Если сравнение θ возможно, то для атрибутов A и B одинаковые имена необязательные.

Пример. Пусть имеются отношения $R_1(A, B, C)$ и $R_2(D, E, F)$

R_1: $A\ B\ C$ $\qquad$ R_2: $D\ E\ F$

$$\begin{bmatrix} 8 & 4 \\ 2 & 7 \\ 5 & 5 \\ 8 & 3 \end{bmatrix} \qquad \begin{bmatrix} d_1 & 8 & 3 \\ d_2 & 8 & 2 \\ d_3 & 8 & 10 \\ d_4 & 5 & 12 \end{bmatrix}$$

Для условий (C=E), (B>F) получим

$$R_1 \underset{C-E}{\bowtie} R_2: \quad \begin{array}{c} \begin{matrix} A & B & C & D & E & F \end{matrix} \\ \begin{bmatrix} a_2 & 5 & 8 & d_1 & 8 & 3 \\ a_2 & 5 & 8 & d_2 & 8 & 2 \\ a_2 & 7 & 8 & d_1 & 8 & 3 \\ a_2 & 7 & 8 & d_2 & 8 & 2 \\ a_3 & 2 & 5 & d_4 & 5 & 12 \end{bmatrix} \end{array}$$

$$R_1 \underset{B>F}{\bowtie} R_2: \quad \begin{array}{c} \begin{matrix} A & B & C & D & E & F \end{matrix} \\ \begin{bmatrix} a_1 & 3 & 4 & d_2 & 8 & 2 \\ a_2 & 5 & 8 & d_1 & 8 & 3 \\ a_2 & 5 & 8 & d_2 & 8 & 2 \\ a_2 & 7 & 8 & d_1 & 8 & 3 \\ a_2 & 7 & 8 & d_2 & 8 & 2 \end{bmatrix} \end{array}$$

Таким образом, с помощью операции соединения можно объединять строки разных отношений по критерию сравнения значений каких-нибудь двух атрибутов.

Деление. Рассмотрим деление m-местного отношение R_1 на n-местное отношение R_2.

Пусть из общего количества m атрибутов отношения R_1 выделим несколько атрибутов: $A, B,...,F$ и из них составим список, обозначив его через M. Набор значений атрибутов из M столбцов можно рассматривать как проекцию отношения R на список атрибутов M, т.е. $\pi_M(R_1)$. Тогда через $\overline{M}$ будут обозначаться атрибуты дополнительного к M, т.е. атрибуты отношения R_1, не вошедшие в список M, и соответственно значения атрибутов из списка $\overline{M}$ определяются $\pi_{\overline{M}}(R_1)$. Будем полагать, что делимое (отношение R_1) может быть представлено таким образом, что его атрибуты сгруппированы в порядке $R_1(\overline{M}\ , M)$. В отношении R_2 также выделим несколько атрибутов $G, K,...,P$ и составим из них список, обозначив его через N (формально этот список представляет собой проекцию $\pi_N(R_2)$). Если проекции $\pi_M(R_1)$ и $\pi_N(R_2)$ объединимы , т.е. имеют одинаковое количество атрибутов, то можно рассматривать операцию деления R_1 по M на R_2 по N (что обозначается как $R_1[M$ ч $N]R_2$). Операция деления $R_1[M$ ч $N]R_2$ представляет собой операцию, которая определяет такое наибольшее множество значений атрибутов с $\pi_{\overline{M}}$ (R_1) (обозначим его через $r_1(\overline{M})$), что прямое произведение этого множества $r_1(\overline{M})$ с $\pi_N(R_2)$ содержится в R_1.

Операции деления можно определить с помощью уже ранее введенных операций таким образом:

$$R_1[M \div N]R_2 = \pi_{\overline{M}}(R_1) \setminus \pi_{\overline{M}}((\pi_{\overline{M}}(R_1) \times \pi_N(R_2)) \setminus R_1),$$

где $\pi_{\overline{M}}$ - это проекция отношения на атрибуты списка $\overline{M}$.

Пример.

R_1:

A	B	C	D
a_1	b_1	c_3	d_1
a_2	b_1	c_1	d_1
a_3	b_2	c_2	d_2
a_1	b_3	c_1	d_3
a_3	b_2	c_1	d_1
a_3	b_1	c_2	d_2

R_2:

C	D
c_1	d_1
c_2	d_2

Включим в список матрицы *М* атрибуты *С* и *D* (из отношения R_1), тогда в список $\overline{M}$ войдут атрибуты *А* и *В,* а в список *N* — атрибуты *С* и *D* из отношения R_2:

$R_1[(C,\ D) \div (C,\ D)]R_2$

A	B
a_3	b_2

Действительно,

$\pi_{\overline{M}}(R_1)$:

A	B
a_1	b_1
a_2	b_1
a_3	b_2
a_1	b_3
a_3	b_1

$\pi_{\overline{M}}(R_1) \times \pi_N(R_2)$:

A	B	C	D
a_1	b_1	c_1	d_1
a_1	b_1	c_2	d_2
a_2	b_1	c_1	d_1
a_2	b_1	c_2	d_2
a_3	b_2	c_1	d_1
a_3	b_2	c_2	d_2
a_1	b_3	c_1	d_1
a_1	b_3	c_2	d_2
a_3	b_1	c_1	d_1
a_3	b_1	c_2	d_2

$\pi_{\overline{M}}(R_1) \times \pi_N(R_2) \setminus R_1$:

A	B	C	D
a_1	b_1	c_1	d_1
a_1	b_1	c_2	d_2
a_2	b_1	c_2	d_2
a_1	b_3	c_1	d_1
a_1	b_3	c_2	d_2
a_3	b_1	c_1	d_1

$$\pi_{\overline{M}}((\pi_{\overline{M}}(R_1) \times \pi_N(R_2)) \setminus R_1):$$

A	B
a_1	b_1
a_2	b_1
a_1	b_3
a_3	b_1

$$R_1[(C,\ D) \div (C,\ D)]\, R_2 = \pi_{\overline{M}}(R_1) \setminus \pi_{\overline{M}}((\pi_{\overline{M}}(R_1) \times \pi_N(R_2)) \setminus R_1):$$

A	B
a_3	b_2

При построении языка манипулирования данными на основе реляционной алгебры каждый оператор языка реализует некоторый набор алгебраических операций, в результате выполнения которых получается желательное исходное отношение.

Другой подход к построению языка манипулирования данными основан на использовании моделей реляционного исчисления.

Суть этого подхода заключается в том, что желательный результат определяется не заданием набора операций над отношениями, а путем описания требований, которым должно удовлетворять результирующее отношение. СУБД представляется самой подобрать последовательность операций, которые ведут к поставленной цели. Естественно, что языки, которые основаны на использовании моделей реляционного исчисления, представляют собой языка очень высокого уровня.

Реляционное исчисление. В реляционном исчислении, равно как и в реляционной алгебре, имеется набор понятий и операций, которые позволяют записывать любое отношение в виде некоторой формулы или формального выражения (α-выражения).

Формулы в реляционном исчислении кроме арифметических операций (= , ≠, <, ≤, >, ≥) включают дополнительные **логические** операции. К ним относят операции квантификации:

$\forall$ - квантор общности и $\exists$ - квантор существования.

Квантор общности $\forall$ читается как: «все для всех», «для каждого», «который бы не был».

Например, запись $\forall x$ читается «для любого *x*» или «для всех *x*».

Квантор существования $\exists$ читается «для некоторого», «существует хотя бы один». Например, $\exists\ y$ читается «для некоторого *y*» или «существует *y* такой, что». Кроме кванторов в реляционном исчислении используются логические операции: $\vee$, $\wedge$, $\urcorner$.

Операция $\vee$ носит название **логическое сложение,** или **дизъюнкция,** и по смыслу соответствует слову «ИЛИ». Операция $\wedge$ **логическое умножение, или конъюнкция,** и отвечает слову «И». Операция $\urcorner$ — операция отрицания, отвечает «НЕ», т.е. запись $R_1 \wedge R_2$ будет читаться как «отношение R_1 И отношение R_2», запись $R_1 \vee R_2$ — «отношение R_1 ИЛИ отношение R_2».

При записи выражения в реляционном исчислении используется понятие **свободных или связных переменных.**

Вхождение переменных *x* в формулу реляционнго исчисления $\psi(x)$ связано, если в ψ она находится в части формулы, которая начинается квантором $\forall$ или $\exists$, за которым непосредственно следует переменная *x*. В таких случаях говорят, что квантор связывает переменную *x*. В остальных случаях вхождение переменной *x* в формулу ψ свободно.

Например, в формуле

$$\forall \boldsymbol{x}(R_1(x,\ y)\bigvee(\exists y)R_2(x,\ y,\ z))$$

переменная *x* связана, переменная *y* в отношении R_1 свободна, а в R_2 — связана, переменная *z* свободна.

Формулы в реляционном исчислении строятся из атомов и совокупности арифметических и логических операторов. Атомы в реляционном исчислении представляются по-разному. В зависимости от того, что используется в качестве переменной - кортеж (строка) или атрибут (столбец). Поэтому различают реляционное исчисление с переменными-кортежами и переменными -доменами.

Реляционное исчисление с переменными-кортежами. Выражение такого исчисления может иметь следующий вид:

$$\{r\,|\psi(r)\},$$

где *r*-кортеж: ψ — некоторая формула исчисления. Например, выражение $\{r|R_1(r) \wedge R_2(r)\}$, где в качестве формулы $\psi(r)$ используется выражение $R_1(r) \wedge R_2(r)$, означает, что необходимо получить множество всех кортежей *r*, таких, что они принадлежат одновременно как отношению R_1, так и отношению R_2. В том случае, когда в качестве переменных в реляционном исчислении используются кортежи, атомы, из которых конструируются формулы, они могут быть следующих типов:

1. Атом — отношение $R(r)$, где *r* — кортеж в отношении *R*.

2. Атом — конструкция типа $s[i]\theta v[j]$ или $s[i]\theta c$, где *c* — некоторая константа; θ — арифметический оператор $(=, \neq, <, \leq, >, \geq)$;

s, v — кортежи; i, j — номера (или имена) атрибутов (столбцов) в соответствующих кортежах. Например, атом (s[l] =q[7]) означает, что первая компонента кортежа s равна седьмому кортежу q, а атом s[5]<10 означает, что пятая компонента меньше 10. Два вышеперечисленных типа вида атомов являются единственными в реляционном исчислении.

Сами формулы рекурсивно конструируются из атомов по следующим правилам.

1. Любой атом - это формула. Все вхождения переменных-кортежей, которые упомянуты в атоме, являются свободными.

2. Если ψ формула, то $\neg\psi$ — тоже формула ($\neg$ - символ логического отрицания).

3. Если ψ_1 и ψ_2 формулы, то и выражения $\psi_1 \wedge \psi_2$ и $\psi_1 \vee \psi_2$ также являются формулами. Причем свободными (связными) являются те и только те вхождения переменных, которые происходят от свободных (связных) вхождений переменных ψ_1 и ψ_2.

4. Если ψ формула и r - свободная переменная-кортеж этой формулы, то $\forall r(\psi))$ и $\exists r(\psi)$ также формулы, переменная в этом случае становится связанной.

5. Формулы могут при необходимости заключаться в скобки.

6. Ничто другое не является формулой.

Реляционное исчисление с переменными-доменами. В этом случае в качестве переменных вместо кортежей используются домены. Реляционное исчисление с переменными на доменах строится с использованием тех же операторов, которые и с переменными кортежами. Но в качестве атомов формул исчисления используются следующие типы:

1. $R(x_1, x_2, ..., x_n)$, где R — n-арное отношение, x_i — константа или переменная на некотором домене.

Атом $R(x_1, x_2, ..., x_n)$ указывает, что значения x_i, которые являются переменными, должны быть выбраны так, чтобы $(x_1, x_2, ..., x_n)$ было кортежем отношения.

2. $x\theta в$, где x и y — константы или переменные на некотором домене, θ - арифметический оператор сравнения.

В остальном формулы реляционного исчисления с переменными-доменами строятся аналогично формулам исчисления с переменными-кортежами.

Выражение реляционного исчисления с переменными на доменах имеет вид $\{x_1, x_2.....x_n|\psi(x_1, x_2, ..., x_n)\}$, где ψ — формула, обладающая тем свойством, что только ее свободные переменные на доменах являются различными переменными $x_1, x_2, ..., x_n$.

Выражение реляционного исчисления $\{r|\psi(r)\}$ или $\{x_2, x_2, ..., x_n|\psi(x_1, x_2, ..., x_n)\}$ служит основой реальных языков манипулирования данными.

В реляционном исчислении доказано, что для любого простого выражения исчисления существует эквивалентное ему выражение реляционной алгебры. Поэтому может быть построена универсальная процедура для перевода выражений реляционного исчисления в эквивалентное по смыслу алгебраическое выражение.

11.6. Сетевая модель данных

В основе разработки сетевых моделей данных лежит возможность представления связей между данными в графической форме.

Наиболее развитой сетевой моделью является модель, которая предложена Рабочей группой по базам данных (РГБД) Ассоциации по языкам обработки данных (КОДАСИЛ). Нужно отметить, что в основе модели КОДАСИЛ лежат понятия «сущность» и «связь», а к основным типам структур модели относят: элемент данных, агрегат, запись, набор.

Сущность — это собирательное понятие, некоторая абстракция реально существующего объекта предметной области процесса или явления. Набор однородных объектов или явлений определяет **тип** сущности, а каждый конкретный объект в наборе представляет **экземпляр** сущности. Связи между сущностями фиксируются заданием множества отношений. При анализе связей между сущностями наиболее часто используются бинарные связи, т.е. связи между двумя сущностями. По характеру бинарные связи между типами сущностей различают:

один к одному (1:1);

один ко многим $(1:M)$;

много к одному $(M : 1)$;

много ко многим $(M:M)$.

Элемент данных - это наименьшая единица данных, которой можно оперировать в БД и выполнять построение всех других структур. Можно отметить, что элемент данных представляет собой аналог поля в файловых системах. Элемент данных имеет имя, которое хранится в БД как часть описания базы. Именами элементов данных могут быть, например, ИНДЕКС ИЗДЕЛИЯ, ДАТА ВЫПУСКА, СТОИМОСТЬ и т.д. В сетевых моделях элементы данных используются для представления атрибутов сущности.

Агрегат данных — совокупность элементов данных, которые имеют общее имя, которую можно рассматривать как единое целое. Например, агрегат данных ДАТА состоит из элементов данных: ЧИСЛО, МЕСЯЦ, ГОД.

Запись - совокупность элементов данных, которые описывают конкретный экземпляр объекта (сущности). Предположим, что сущность ТЕЛЕВИЗОР описывается элементами данных: МАРКА; ИНДЕКС, ЦЕНА. Тогда запись в этом объекте для конкретного изделия может быть: РЕКОРД, ВЦ-311, 640. Можно отметить, что запись эквивалентна кортежу в реляционных моделях данных.

Сетевая модель РГБД КОДАСИЛ в качестве базовых использует понятие «экземпляр» и «набор».

Тип — это общее понятие, которое представляет собой собрание экземпляров записи. Каждый тип записи состоит из некоторого числа элементов данных, значения которых размещаются в экземплярах записи данного типа. В качестве связей между типами записей используются наборы. Каждый набор представляет собой отношение (связь) между двумя или несколькими типами записей. Он отображает множество связей между экземплярами записей типа «владелец» и «член». Для каждого типа набора один тип записи может быть объявлен «владельцем», а остальные — его «члены». При этом любой экземпляр записи типа «член» может быть связан не больше чем с одним экземпляром типа «владелец».

Графическая интерпретация сетевой модели данных представляет собой ориентированный граф без петель. Причем, вершинам графа соответстуют типы записей, а дугам - наборы, отражающие связи между соответствующими типами записей. Направленные стрелки на дуге ориентированы от записи типа «владелец» к записи типа «член».

Подмножество дуг, соединяющих один запись - владельца с несколькими записями- членами, называется **экземпляром набора**.

Рассмотрим, например, граф, отражающий упрощенную БД комплектующих деталей телевизора (рис. 1).

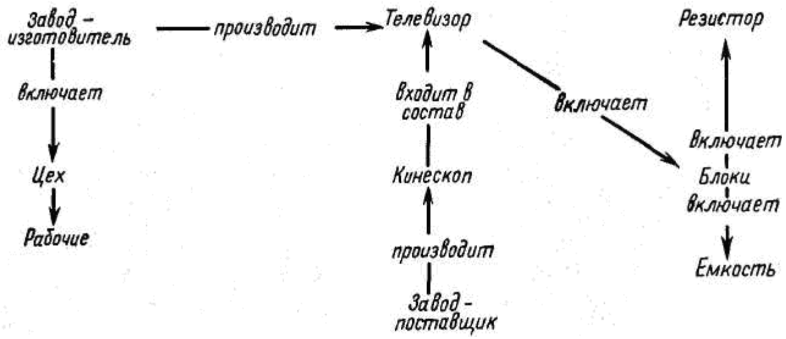


Рис. 1. Пример построения базы данных комплектующих телевизора

Стрелки между вершинами отвечают наборам данных, отражающих связи между записями, а надписи над стрелками - именам наборов.

Как тип записи, так и набор данных в общем случае могут быть представленны таблицами. Но в отличие от таблиц реляционных моделей, в сетевых моделях данных они могут допускать дубликаты строк или записей.

В модели КОДАСИЛ вводится особый тип набора данных, называемый **сингулярным** и имеющий только один экземпляр набора этого типа. В нем запись «владелец» отсутствует (владельцем является система управления базой данных). Этот тип набора обычно используется для создания традиционного файла, состоящего из однотипных записей.

В сетевой модели данных КОДАСИЛ есть несколько ограничений, которые необходимо учитывать при построении модели. Основным внутренним ограничением является функциональность связей, так как нельзя реализовать связи типа *М:М.* В модели это ограничение соответствует положению: в конкретном экземпляре набора экземпляр записи числа может иметь не больше одного экземпляра записи владельца.

Для того чтобы отобразить принятую схему данных в памяти ЭВМ, требуется описать все таблицы, соответствующие записям и наборам. Для этого группой КОДАСИЛ был предложен язык описания данных, позволяющий задать схему данных с помощью четырех типов статей.

Статья схемы задает имя схемы БД. Статья запишется как:

SCHEMA NAME IS имя схемы.

Статья области характеризует область памяти, в которой размещаются экземпляры записей БД. С помощью этой статьи можно в случае необходимости распределять БД по разным ЗУ. Поскольку в общем случае можно выделить под БД несколько разных областей, каждая из них должна иметь собственное уникальное имя и описываться следующим образом:

AREA NAME IS имя области;

Статья записи содержит описание типа записи, которая включает имя записи и характеризуючі все элементы данных, которые входят в ее состав. Каждому типа записи отвечает своя статья. Статья записи начинается предложением:

RECORD NAME IS имя записи;

Следующий за этим предложением текст зависит от варианта реализованного языка КОДАСИЛ. Наиболее распространенными есть ЯОД КОДАСИЛ. Поскольку ряд действующих СУБД реализуют ЯОД, дальше будем рассматривать случай, когда статья записи формируется на его основе. Тогда вторым предложением в схеме записи будет идти предложение:

$$\text{LOCATION MODE IS} \begin{cases} \text{DIRECT} \\ \text{CALC (имя процедуры) USING (имя calc-элемента)} \\ \text{DUPLICATES ARE (NOT) ALLOWED} \\ \text{VIA имя набора SET} \\ \text{SYSTEM} \end{cases}$$

где в фигурных скобках указано одно из возможных описаний.

DIRECT используется в том случае, когда предполагается, что номер страницы области для размещения записи будет определяться программой, которая выполняет включение записи в БД.

CALC применяется, когда предполагается, что специальная программа будет использовать значение ключа БД.

Ключ базы данных — это идентификатор, который уникально определяет запись, помещенную в БД.

Для того чтобы можно было различать отдельные экземпляры записей, которые сохраняются в БД, каждому экземпляру записи присваивается значение ключа, который играет роль внутритсистемного идентификатора

Знавая значение ключа, можно быстро отыскать соответствующую запись. В формате CALC USING имя calc-элемента, в качестве имен Calc-Элемента используются имена элементов данных записи.

В том случае если ключи не имеют дубликатов значений, то в варианте CALC следует указать DUPLICATES ARE NOT ALLOWED. Если ключ имеет дубликаты значений, например в качестве ключа задан элемент данных ФАМИЛИЯ в записи типа СТУДЕНТ, то следует указать DUPLICATES ARE ALLOWED.

VIA SET используется в том случае, когда нужно запись разместить физически как можно ближе к соответствующему экземпляру набора, в который он будет включен.

SYSTEM применяется в случае, если размещение записей возлагается на саму СУБД (в соответствии с заложенными в нее алгоритмами).

Для того чтобы приписать рассматриваемый тип тип записи к некоторой области, используется предложение

WITHIN имя области AREA

Для описания внутренней структуры записи в статью записи включается подсистема данных, которая имеет вид

$$\left\{\begin{array}{l} \left\{\begin{array}{l}\text{BINARY}\\ \text{DECIMAL}\end{array}\right\} \\ \left\{\begin{array}{l}\text{FIXED}\\ \text{FLOAT}\end{array}\right\} \\ \text{REAL} \\ \text{CHARACTER} \\ \text{DATA-BASE-KEY} \end{array}\right\}$$

С помощью этой подсистемы каждому элемента данных приписывается тип значений данных: двоичное, десятичное, фиксированное, плавающее, натуральное, символьное, ключ БД.

Статья набора позволяет описать наборы БД. Формат задачи набора имеет вид

SET NAME IS имя набора;

$$\text{OWNER IS} \left\{\begin{array}{l}\text{имя записи}\\ \text{SYSTEM}\end{array}\right\}$$

$$\text{ORDER IS PERMANENT INSERTION IS} \left\{\begin{array}{l}\text{SORTED}\\ \text{PRIOR}\\ \text{NEXT}\\ \text{LAST}\\ \text{FIRST}\end{array}\right\}$$

Первое предложение статьи задает имя описываемого набора. Второе указывает имя типа записи, являющейся записью владельца набора. И последнее - на способ включения экземпляров записей-членов в экземпляры описываемого типа набора.

Обычно предполагается, что на внутреннем уровне экземпляры набора, которые включают несколько членов записей, организованы в виде цепочки.

Цепочка записей реализуется с помощью специального служебного элемента — **указателя,** который содержит (указывает на) адрес записи, логически связанного с рассматриваемой. Посредством указателей можно организовать обращение не только к следующим записям, но и к предыдущим.

Команды, которые входят в фигурные скобки третьего предложения схемы, как раз и позволяют организовать цепочку необходимым образом и означают:

FIRST - запись включается первым в цепочку перебора записей, т.е. сразу же после записи владельца (рис. 2).

LAST - запись включается последней (рис. 2).

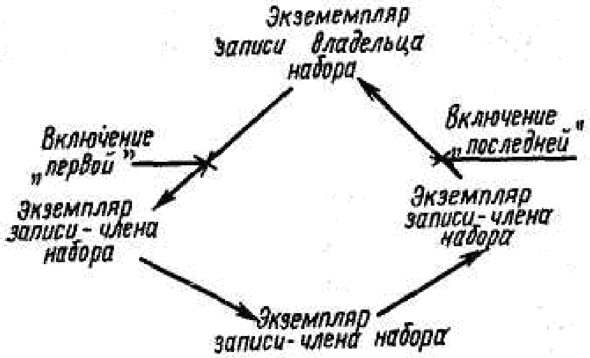


Рис. 2. Реализация вариантов FIRST и LAST

NEXT - включение следующей записи, т.е. записи, которая следует за текущей записью набора (рис. 3). Под текущей записью набора понимается тот экземпляр записи конкретного экземпляра набора, на который указывает «текущая» запись, в типе набора.

PRIOR - включение предыдущей записи, т.е. перед текущей записью набора (рис. 3).

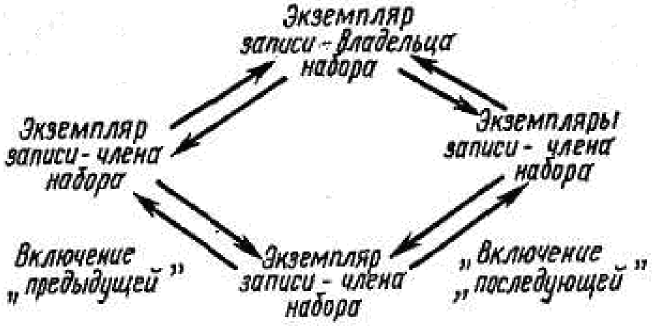


Рис. 3. Реализация вариантов NEXT и PRIOR

SORTED - СУБД поддерживает упорядоченность экземпляров однотипных записей, которые входят в экземпляр набора, в соответствии с возрастанием и убыванием ключевого данного.

Описание числа набора проводится с помощью подстатьи члена набора, имеющей вид

$$\text{MEMBER IS имя записи}\left\{\begin{array}{ll}\text{AUTOMATIC} & \text{MANDATORE}\\ \text{MANUAL} & \text{OPTIONAL}\end{array}\right\}$$

В зависимости от способа включения экземпляра записи-члена в экземпляр набора различают два типа членства в наборе:

AUTOMATIC - автоматический;

MANUAL - ручной.

Чтобы поместить в БД новую запись, необходимо сформировать ее в рабочей области программ и только затем с помощью операторов языка манипулирования данными записать в БД. Если необходимо экземпляр записи-члена поместить в БД, но при этом не соединять ни с каким экземпляром набора, то указывают оператор MANUAL. Последующее включение этой записи в набор должно быть осуществлено программистом явным образом с помощью соответствующих средств языка манипулирования данными. Тип членства AUTOMATIC предусматривает, что в момент помещения в БД запись должна быть автоматически включена в набор. Операторы MANDATORY (обязательный) и OPTIONAL (необходимый) характеризуют вид исключения экземпляра записи из экземпляра набора описываемого типа. Так тип OPTIONAL означает, что запись может исключаться из набора данных в произвольный момент времени.

Поскольку БД предполагает соответствующее манипулирование данными, рабочей группой КОДАСИЛ был также предложен соответствующий язык, позволяющий осуществлять запись в базе данных: STORE (запомнить), ERASE (стереть), MODIFY (изменить), CONNECT (присоединить), DISCONNECT (исключить) и др.

Оператор STORE (запомнить) - помещает данные, подготовленные в рабочей области программы, в базу.

Оператор ERASE (стереть) - удаляет из БД экземпляр записи.

Оператор MODIFY (изменить) - обновляет текущую запись банка данных.

Оператор CONNECT (присоединить) - включает текущую запись в набор.

Оператор DISCONNECT (исключить) - исключает текущую запись из набора, но при этом она остается в БД.

Можно отметить, что в языке манипулирования данных существуют и другие операторы, но приведенный набор операторов ЯМД является функционально полным, т.е. позволяет образовывать в БД произвольное допустимое схемой базы данных состояние.

11.7. Иерархическая модель данных

Иерархическая модель данных, так же как и сетевая, основаны на возможности представления структур данных в виде графов. Но в отличие от сетевой на иерархическую модель накладываются более жесткие ограничения. Граф иерархической базы данных имеет древовидную структуру связей.

Древовидная структура или дерево - это граф, не содержащий циклов. Дерево представляет собой связанный граф, так как каждая вершина в нем завершает по крайней мере одно ребро (рис. 1).

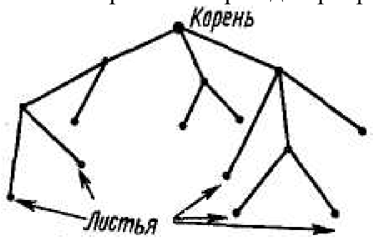


Рис. 1. Граф дерева

При работе с иерархической моделью данных граф, описывающий структуру данных, является направленным или ориентированным, т.е. дерево в этом случае будет ориентированным.

В зависимости от направления дуг в ориентированном графе выделяют такие типы вершин, как корень и лист.

Корень — это вершина, которая имеет одну или несколько исходящих дуг и ни одной входящей.

Лист - это вершина, которая имеет одну или несколько входящих дуг и ни одной исходящей.

Корень можно рассматривать как источник направленной древовидной структуры. От корня до любого листа легко проследить элементарную последовательность дуг. Число дуг между корнем и листом называется уровнем диста. Вершину графа, который не является ни корнем, ни листом, называют **узел ветвления.**

Примером иерархической структуры в виде дерева служит схема управления большинства организаций. Корнем схемы является директор. От него исходят одна или несколько дуг к его заместителям. От них отходят дуги к более низким уровням управления и т.д. к непосредственных исполнителей, которые на графе соответствуют листьям.

Основными понятиями в иерархической модели данных есть **тип записи** и **иерархические отношения.** Вершины в дереве соответствуют типу сущности и называются типом записи. Тип записи состоит из одного или более **элементов данных.** Во многих иерархических моделях вместо понятия «тип записи» используют эквивалентное понятие «тип сегмента».

Иерархическое отношение (ветка дерева) соединяет два типа записей и представляет собой множество связей между экземплярами записей этих двух типов.

Дуги (ветки) дерева соответствуют функциональному типу связи, т.е. типам 1:1, 1:*M*, *M*:1, и их называют **связью исходной порожденной.** Дуга выходит из **типа родительской записи** и заходит в **тип порожденной записи.** Таким образом, рассматривая последовательность связей — исходной-порожденной, можно выделить типы родительских и порожденных записей. Каждый экземпляр родительского типа записей может иметь связь с несколькими (в том числе и с нулем) экземплярами порожденных записей. В свою очередь, каждый экземпляр записи порожденного типа подчинен ровно одному экземпляру записи родительского типа. Другими словами, иерархическое отношение можно рассматривать как функцию, признаком которой служит экземпляр порожденного типа записи, а ее значением является экземпляр родительского типа.

Иерархическая модель накладывает жесткие ограничения на иерархические отношения между записями. Поскольку любые два типа записей могут быть связаны не более чем одним иерархическим отношениям, то иерархическим связям не требуются собственные имена. Каждая из иерархических связей может быть однозначно идентифицирована указанием родительской и порожденной записи.

Модели данных « сущность-связь». Они появились как обобщение и развитие иерархических и сетевых моделей. Модель « сущность-связь» была задумана как средство представления предметной области, не зависимого от особенностей среды хранения и не связанного соображениями физической эффективности. Базовыми структурами в этих моделях являются типы сущностей и связей. Тип сущности в модели носит название множество сущностей. **Каждая сущность при этом идентифицирует объект предметной области**. Связь между ними фиксируется заданием множества отношений. Различают два вида отношений - **слабые** и **стандартные**. Слабые связи идентифицируют иерархические отношения.

Множество связей (МС) в данной модели можно представить как математическое отношение n типов сущностей. Если МС есть множество, то его можно определить следующим образом:

$$МС=\{<m_1, m_2,...,m_n>|m_1 \in M_1,\ \ m_2 \in M_2,..., m_n \in M_n\},$$

где m_1 — сущность, которая принадлежит множеству сущностей M_i, а кортеж $<m_1, m_2,...,m_n>$- связь, которая принадлежит множеству МС.

Домен в модели «сущность-связь» называют **множеством значений.** В этом случае значения представляет собой конкретный экземпляр множества значений. Для изображений множеств сущностей и отношений могут быть использованы диаграммы, аналогичные графам. МС изображается прямоугольником, множество отношений - ромбом. Множествам обоих типов присваиваются имена; множества сущностей соединяются с множествами отношений, в котором они участвуют с помощью ненаправленных линий. Множество значений представляется овалом (рис. 2).

Рис. 2. Модель данных «сущность-связь»

Использование моделей «сущность-связь» является удобным средством для представления концептуальной информации.

Бинарные модели. Отношения, используемые для описания предметной области, в общем случае могут быть *n*-мерными. Интенсивно идут исследование по представлению информации с помощью, как правило, бинарных отношений, поскольку они позволяют лаконично представить сложные связи.

Бинарные модели, основанные на использовании бинарных отношений, имеют простые базовые структуры, способные обеспечить эффективное представление предметной области. Поскольку графы позволяют наглядно задавать отношения, их использование в бинарных моделях также позволяет лучше уяснить особенности модели. Вершины графа в бинарных моделях соответствуют классификационному обобщению экземпляров данных в типы и называются *категориями,* а дуги — *бинарным отношением* категорий. Граф, который удовлетворяет этим структурным представление, носит название *графа типов.* Каждому бинарному отношению ставится в соответствие отношение, имеющее противоположное направление. Например:

СТУДЕНТ -УЧИТСЯ У - ПРЕПОДАВАТЕЛЬ - ОБУЧАЕТ -

Обеим направлениям бинарного отношения присваиваются уникальные имена, которые называются *функциями доступа.*

В бинарном отношении категорий СТУДЕНТ и ПРЕПОДАВАТЕЛЬ направление от категории СТУДЕНТ к категории ПРЕПОДАВАТЕЛЬ есть функция доступа УЧИТ У. Представленное отношение может быть охарактеризовано следующим образом:

СТУДЕНТ УЧИТСЯ У ПРЕПОДАВАТЕЛЯ

Функция доступа для противоположного направления будет

ПРЕПОДАВАТЕЛЬ УЧИТ СТУДЕНТА

Расширение бинарного графа типов позволяет рассматривать понятие двух родов-объектов и поименованных бинарных отношений (связей). Взаимосвязи, в которых принимает участие более двух объектов, в свою очередь, интерпретируются как объекты.

Объект - это реализация категории. Например, объект - это конкретные студенты и преподаватели. Объекты подразделяются на **абстрактные и конкретные**. Причем имеется в виду, что абстрактные всегда существуют, в то время как конкретные появляются и исчезают в описании реального объекта. Так, абстрактные объекты используются для представления чисел, дат и др.

Объекты соединены связями, которые отображают реализацию бинарного отношения. Каждой связи в обоих направлениях присваиваются имена, соответствующие функциям доступа.

Для создания категорий в бинарной модели данных используют оператор CATEGORY. Например, условие

СТУДЕНТ = CATEGORY

говорит о создании новой категории с именем СТУДЕНТ.

Для создания бинарного отношения следует задать его имя, функции доступа и категории соответствующему данному отношению объектов. Например:

СТУДЕНТ - ПРЕПОДАВАТЕЛЬ = RELATION (СТУДЕНТ, ПРЕПОДАВАТЕЛЬ, УЧИТСЯ У, УЧИТ) определяет бинарное отношение СТУДЕНТ - ПРЕПОДАВАТЕЛЬ категорий СТУДЕНТ и ПРЕПОДАВАТЕЛЬ с функциями доступа УЧИТЬСЯ У, ОБУЧАЕТ.

В бинарных моделях отсутствует явно выраженное понятие «свойство объекта». Свойства объектов могут быть определены с помощью бинарного отношения, заданного на множестве объектов и других элементарных объектов, с помощью которых задаются значения свойств.

Как уже отмечалрсь, отношение характеризуется двумя функциями доступа, каждая из которых определяет минимальное и максимальное число объектов в присоединяемой категории. Ограничение на число объектов задается оператором AFN, позволяющим задать краничные величины подмножеств значений функций доступа. Так, например, если для функции доступа ОБУЧАЕТ запишем:

$$ОБУЧАЕТ = AFN (0, \infty),$$

то это будет означать, что преподаватель может вообще никого не обучать, или обучать любое число студентов.

Логический доступ к данным бинарной модели обеспечивается с помощью программ, которые реализуют элементарные операции доступа.

Семантические сети. Для представления семантических (смысловых) текстов, задаваемых на естественном языком, разработаны семантические сетевые модели.

Семантическая сеть представляет собой ориентированный граф с намеченными вершинами и дугами. При этом если вершины обозначаются только в целях ссылок к ним, то метки дуг содержат сведения о некоторых их семантических свойствах и значениях.

Для представления данных используются четыре типа вершин: **концепты** (или понятия), **события, характеристики** (свойства) и **значения.**

Концепты - константы или параметры, которые специфицируют физические или абстрактные объекты.

События - отвечают действиям, наблюдаемым в представлякмой области.

Характеристики - вершины, соответствующие свойствам концепты.

Значения - вершины, соотносящиеся с областями значений, которые могут принимать характеристики.

Поскольку имеется четыре типа вершин, необходима соответствующая зависимая от этих типов интерпретация дуг, которые соединяют различные вершины. Модели семантической сети предусматривают возможность распределения вершин по типам. В этом случае следует различать вершины-концепты и вершины-классы, которые собственно и представляют определенные типы вершин. Например, КУЛЕШОВ - концепт, СТУДЕНТ- класс.

Различие между классом и концептом весьма близко к тому же, что между типом и экземпляром в других моделях. Отличие заключается в том, что граф семантической сети включает как классы, так и концепты. Кроме того, концепт может быть соотнесен с несколькими классами.

Поскольку семантические сети предусматривают задавать на графе в явном виде различие между вершинами-концептами и вершинами-классами, то в рассмотрение вводятся три вида дуг: утверждение; порождение экземпляра; бинарное отношение.

Утверждение — дуга, которая соединяет два концепта.

Порождение экземпляра — дуга, между классом и концептом.

Бинарное отношение — дуга, которая связывает два класса.

Рассмотрим пример семантической сети (рис. 3), иллюстрирующей сказанное.

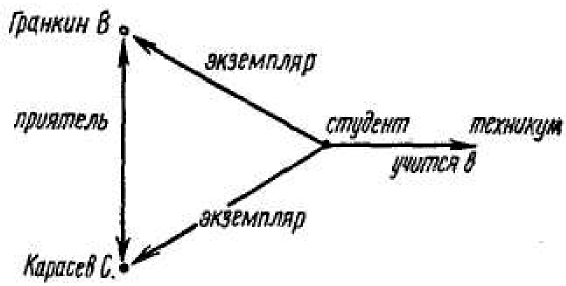


Рис. 3. Пример семантической сети

В данной семантической сети, вершины ГРАНКИН В. и КАРАСЕВ С. - концепты, СТУДЕНТ и ТЕХНИКУМ - классы. Дуга ПРИЯТЕЛЬ - утверждение. Дуги, связывающие вершину СТУДЕНТ с вершинами КАРАСЕВ С. и ГРАНКИН В., отображают связь экземпляров с классами. Дуга УЧИТСЯ В отображает бинарное отношение между классами СТУДЕНТ и ТЕХНИКУМ.

Классы могут быть связаны в иерархию в соответствии со связями ЕСТЬ НЕК и ЕСТЬ - ЧАСТЬ. Эти связи позволяют из отдельных понятий и классов строить более общие понятия и классы.

Для представления в семантической сети некоторых событий и действий вводится набор простых отношений, которые характеризуют основные компоненты события. Для построения с помощью семантической сети структуры события в первую очередь выделяют из него само действие, которое описывается обычно глаголом. После этого выделяют лиц, совершающих действие и объекты, над которыми оно осуществляется. Лицо, которое осуществляет действие, называется **агентом.** Вещи, над которыми действие осуществляется, называют **объектами**. Лицо, получающееся результатом действия или испытывающее его, называется **адресат**.

Рассмотрим предложение: «Мастер починил телевизора». В этом предложении выделим действие: ПОЧИНИЛ. Очевидно, что объектом является МАСТЕР (рис. 4).

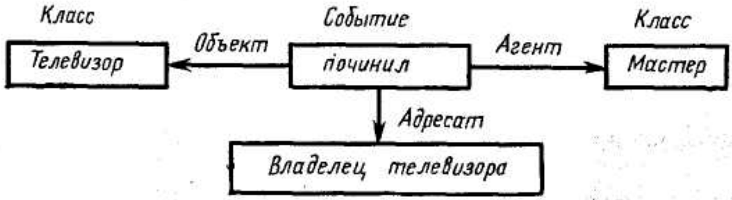


Рис. 4. Сеть предложения «Мастер починил телевизора»

В данном предложении адресат явно не указан, но его наличие можно предположить. Рассмотрим более сложное предложение:

«Вчера мастер Босин П. А. починил магнитофон «Юность», принадлежащий студенту Афонину К. А.».

Семантическая сеть для данного предложения представлена на рис. 5.

Рис. 5. Пример семантической сети

В этом предложении появилась новая дуга ВРЕМЯ, которое указывает на то, когда происходит событие. В общем случае в семантических сетях кроме простых отношений: агент, объект, адресат выделяют и другие, позволяющие описать широкий класс событий. К ним относят: время, место, инструмент, цель и др.

Операции, совершаемые над данными, задаваемые семантической сетью, разбиваются на два подмножества: операции над классами и над бинарными отношениями.

Над классами могут быть совершены четыре операции:

- создание экземпляра некоторого класса или установление принадлежности существующего экземпляра некоторого класса к еще одному;
- устранение принадлежности экземпляра к некоторому классу или полное его исключение;
- выборка экземпляров, которые принадлежат к одного класса;
- определение принадлежности экземпляра указанному классу.

Над бинарными отношениями могут быть совершены три операции:

- установление связи между классами;
- выборка всех экземпляров, связанных в данном бинарном отношении с указанным экземпляром;
- установление наличия связей между двумя экземплярами.

В общем случае реализация всех вышеперечисленных операций требует создания специальных программ, которые учитывают рассматриваемую предметную область.

Разработка семантических сетевых моделей явилась следствием повышения требований к интегрированному представлению данных, которое включает не только данные, но и их категории, свойства категорий и операции над данными. Важную роль в развитии моделей данных этого класса сыграли проблемы алгоритмизации процессов естественного языка. Модели семантических сетей широко используются при разработке систем искусственного интеллекта.

Инфологические модели данных. Ифологическое проедставленне полностью независимо от физических параметров среды хранения. Эта модель в качестве базовой использует понятие: объект, свойства и связи.

Под объектом понимается нечто представляющее интерес для решаемой задачи. Предполагается, что существование объекта связано с такими событиями, как появление (возникновение), изменение и исчезновение.

Объекты подразделяются на атомарные и составные.

Атомарный объект, — это любой объект, дальнейшее разложение которого на другие объекты невозможно.

Составной объект включает в себя множество объектов.

С каждым объектом связывается определенный набор свойств. Важным свойством существования объекта является время (время его возникновения и исчезновения).

С помощью базовых концепций объектов, свойств связей и времени формируется элементарный факт. Элементарный факт задается формально как тройка $<x, y, z>$ или $(<x_1, x_2, ..., x_n>rz)$, где $<x_1, x_2, ..., x_n>$ - кортежи объектов, y — свойство, r — отношение, z — время. Такую базовую структуру называют элементарной плеядой.

Так как в рамках инфологической модели все может быть объявлено объектом, то свойства и отношения, в свою очередь, также могут рассматриваться как объект.

В данной модели типы объектов вводятся путем группирования объектов и их свойств. Объектная группа $O(p)$, соотнесенная со свойством p, определяется как совокупность объектов, потенциально (вне связи со временем) имеющих свойство p. Объекты, обладающие свойством p, в определенный момент времени образуют подмножество $O_t(p)$ множества $O(p)$, называемое временным срезом объектной группы.

В инфологической модели данных понятия атрибута вводится с помощью понятий объекта, свойства, связи и времени. Атрибут определяется как множество свойств $A= \{p_i\}$ объектной группы $O(p)$, такое, что в каждый момент времени каждый объект x, принадлежащий $O_t(p)$, содержится, по крайней мере, в одной $O_i(p)$. Свойства p — это значение атрибута A объектной группы $O(p)$.

Базовые операции включения, обновления и удаления в инфологической модели связаны с вводом новых элементарных сообщений и предполагают наличие механизмов интерпретации элементарного сообщения.

Процедуры обращения к БД **(транзакции)** могут быть представлеы парой (оператор, параметр). К основным операциям, которые изменяют БД, относят <ДОБАВИТЬ, сообщение> <УБРАТЬ сообщение> <ЗАМЕНИТЬ сообщение 1, сообщение 2>.

На основе базовых процедур возможно конструирование более универсальных запросных средств.

Следует иметь в виду, что возможности инфологической модели существенно меньше возможностей естественного языка и их удобно использовать лишь в тех случаях, когда модель предметной области позволяет описать только очевидные факты.

11.8. Матрицы и бинарные отношения на конечных множествах

Формально *матрицей* над множеством S называется отображение

$$M: N_P \times N_q \to S, \qquad p, q \in N.$$

Обычно образ (i, j) обозначают через M_{ij} и изображают всю функцию массивом элементов из S, т.е.

$$M = \begin{bmatrix} M_{11} & M_{12} & \dots & M_{1q} \\ M_{21} & M_{22} & \dots & M_{2q} \\ \cdot & \cdot & \cdot & \cdot \\ M_{p1} & M_{p2} & \dots & M_{pq} \end{bmatrix}.$$

Говорят, что эта матрица имеет *p строк* и *q столбцов* и имеет *размер* $p \times q$. Матрица размера $p \times q$ имеет $p * q$ элементов. Когда $p=q$, матрицу называют *квадратной.* Множество всех матриц $p \times q$ над S обозначают через $M (p, q, S)$. . Множество $M ((p, p, S)$ будем обозначать через $M (p, S)$.

Рассмотрим бинарное отношение ρ между множествами A и B, где

$$A = \{a_1, a_2, ..., a_p\}, \qquad B = \{b_1, b_2, ..., b_q\},$$

т.е. $|A|=p$, $|B|=q$.

Упорядочение элементов в этих множествах выбрано произвольно, однако, однажды выбранное, оно далее остается фиксированным. Пусть это отношение ρ определено посредством выбора пар (a, b), где $a \in A$, $b \in B$.

Рассмотрим матрицу M над $\{0, 1\}$, т.е. $M: N_P \times N_q \to \{0, 1\}$, и свяжем элементы M с отношением ρ биекцией

$$\varphi\colon \mathrm{P}\,(A\times B) \to M\,(p,\, q,\, \mathrm{Z}_2)$$

(φ отображает *произвольное* отношение между A и B в матрицу $p \times q$ над $\{0, 1\}$);

$$\varphi\colon \rho \to M,$$

причем

$$(\varphi(\rho))_{ij} = M_{ij} = \begin{cases} 1, & \text{если} \quad (a_i,\, b_j) \in \rho, \\ 0, & \text{если} \quad (a_i,\, b_j) \notin \rho. \end{cases}$$

В случае, когда полезно подчеркнуть, что матрица M была получена из отношения ρ, мы будем обозначать ее через $M(\rho)$.

Пример 1. Возьмем случай $|A|=4$, $|B|=3$ и

$$\rho = \{(a_1,\, b_2),\, (a_1,\, b_3),\, (a_2,\, b_1),\, (a_3,\, b_1),\, (a_4,\, b_2)\}.$$

Тогда соответствующая матрица M имеет вид

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Таким образом, мы имеем способ табулирования или кодирования отношения и можем закодировать отношение посредством φ или декодировать посредством φ^{-1}. Этот процесс является отображением $(i,\, j)$ в $A\times B$ или M соответственно. Такое представление более удобно, чем теоретико-множественный способ определения отношений, поскольку с ним можно обращаться формальным образом. Оно становится даже более пригодным для вычислений, если наложить некоторую структуру на множество, из которого получается матрица. Возьмем опять $\{0, 1\}$ и определим на этом множестве логическое сложение (**или**) и умножение (**и**). Тогда, если M и N — матрицы $p\times q$, соответствующие отношениям ρ и σ, то матрица Q, представляющая отношение τ, где

$$\tau = \{(a,\, b)\colon\ (a,\, b) \in \rho \text{ или } (a,\, b) \in \sigma\},$$

определяется следующим образом: $Q_{ij}=(M_{ij}\text{или } N_{ij})=M_{ij}+N_{ij}$ (логическое сложение). Следовательно, имеет смысл называть *Q суммой* матриц M и N и писать

$$Q=M + N,$$

подразумевая, что Q, M и N имеют один и тот же размер и Q вычисляется по правилу покомпонентного сложения

$$Q_{ij}= M_{ij}+N_{ij}.$$

Это - пример использования коммутативной диаграммы, которая изображена на рис. 1, где производится операция на одном множестве с использованием операции на другом множестве посредством подходящего отображения φ.

$$\begin{tikzcd}
(\rho,\sigma) \arrow[r] \arrow[d, "\varphi"'] & \rho\cup\sigma \arrow[d, "\varphi"] \\
(M,N) \arrow[r, "+"'] & M+N
\end{tikzcd}$$

Рис. 1.

С этой диаграммой обычно связывается тождество

$$\varphi(\rho \cup \sigma) = \varphi(\rho) + \varphi(\sigma).$$

С помощью этого тождества можно дать более точное определение сложения матриц:

$$M + N = \varphi(\rho) + \varphi(\sigma) = \varphi(\rho \cup \sigma) = \varphi(\varphi^{-1}(M) \cup \varphi^{-1}(N)).$$

Пример 1. (продолжение). Пусть A и B те же, что и раньше, и пусть

$$\sigma = \{(a_1, b_1), (a_1, b_2), (a_2, b_1), (a_3, b_2)\}.$$

Тогда

$$N = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Дальше будет видно, что

$$\rho \cup \sigma = \{(a_1, b_1), (a_1, b_2), (a_1, b_3), (a_2, b_1), (a_3, b_1), (a_3, b_2), (a_4, b_2)\}$$

и, что эквивалентно,

$$M + N = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Более того, если мы возьмем множество $C = \{c_1, c_2, c_3, c_4, c_5)$ и рассмотрим отображение π между B и C, определенное следующим образом:

$$\pi = \{(b_1, c_1), (b_1, c_5), (b_2, c_2), (b_3, c_4), (b_3, c_5)\},$$

это оно может быть представлено в виде матрицы P, где

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

Очевидно, что отношение $\pi\circ\rho$ между A и C корректно определено и, следовательно, будет отвечать матрице 4 × 5. Обозначим эту матрицу через S. Как ее можно вычислить? Для этого надо вычислить S_{ij} для всех i, j, где $1\leq i\leq 4$, $1 \leq j \leq 5$. В силу биекции $S_{ij} = 1$ тогда и только тогда, когда $(a_i,\ c_j)\in\pi\circ\rho$. Однако это так, если только существует некоторое $b\in B$ такое, что $(a_i, b)\in \rho$ и $(b, c_j) \in\pi$, т.е.

$$\begin{aligned}(a_i,\ c_j)\in\pi\circ\rho \equiv\ &(a_i,\ b_1)\in\rho \ \text{ и } \ (b_1,\ c_j)\in\pi\\ \text{или } \ &(a_i,\ b_2)\in\rho \ \text{ и } \ (b_2,\ c_j)\in\pi,\\ \text{или } \ &(a_i,\ b_3)\in\rho \ \text{ и } \ (b_3,\ c_j)\in\pi;\end{aligned}$$

или же, что эквивалентно,

$$S_{ij} = M_{i1} * P_{1j} + M_{i2} * P_{2j} + M_{i3} * P_{3j} = \sum_{k=1}^{3} M_{ik} * P_{kj}.$$

Матрица S, вычисленная по такому правилу, называется *произведением* M и P и обозначается через $M * P$ или просто MP.

Рассмотрим опять естественное (коммутативное) отношение между двумя рассматриваемыми операторами (рис. 2).

$$\begin{tikzcd}(\rho,\pi) \arrow[r, "\circ"] \arrow[d, "\varphi"'] & \pi\circ\rho \arrow[d, "\varphi"] \\ (M,P) \arrow[r, "*"'] & M*P\end{tikzcd}$$

Рис. 2

Тогда

$$M * P = \varphi(\varphi^{-1}(P)\circ\varphi^{-1}(M)).$$

Замечание. Изменение порядка φ зависят от способа определения матрицы отношения; если (вместо этого) мы определим матрицу отношения следующим образом:

M_{ij}= 1 тогда и только тогда, когда $(a_j, b_i) \in\rho$,

то изменения порядка не будет. Хотя с математической точки зрения было бы более желательно иметь один и тот же порядок, это нарушило бы сложившуюся практику.

Пример 1 (продолжение). Выполним вычисления, соответствующие определенным выше отношениям. Получаем

$$MP = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

Если матрицы M и N имеют одинаковый размер, то их сумма существует и определяется формулой

$$(M+N)_{ij}=M_{ij}+N_{ij}$$

а если матрицы P и M согласованы (M имеет размерность $p\times q$, а P — размерность $q\times r$), то умножение матрицы M на P возможно и определяется следующим образом:

$$(MP)_{ij} = \sum_{k=1}^{q} M_{ik} * P_{kj}.$$

Хотя матрицы рассматриваются над (Z_2, $\wedge$, $\vee$), мы используем символы * и + для того, чтобы иметь возможность обобщения введенных выше операций. С этого момента обозначения $\wedge$ и $\vee$ будут использоваться лишь в тех случаях, когда общие операции им неадекватны.

В заключительной части этого пункта ограничимся рассмотрением матриц, которые представляются отношениями на конечном множестве A, где $|A| = n$. Тогда все матрицы согласованы и их сумма и произведение всегда определены.

Из покомпонентного определение сложения следует, что сложение матриц коммутативно и существует *нулевая* $(n\times n)$*-матрица* 0: $0_{ij}=0$ для всех i, j; $1\le i, j\le n$. С другой стороны, умножение матриц, вообще говоря, некоммутативно, однако существует единица, которая называется *единичной* $(n\times n)$*-матрицей* и определяется следующим образом: I: $I_{ij}=1$, если $i = j$, и $I_{ij} = 0$, если $i\ne j$. Так, если X — матрица $n \times n$ и $Y = XI$, то

$$Y_{ij} = \sum_{p=1}^{n} X_{ip} * I_{pj} = \sum_{\substack{p=1 \\ (p\ne j)}}^{n} X_{ip} * I_{pj} + X_{ij} * I_{jj}.$$

Так как все $I_{pj} = 0$, за исключением случая $p = j$, то в сумме все члены, исключая те, где $p= j$, равны нулю. Кроме того, $I_{jj} = 1$. Поэтому

$$Y_{ij} = X_{ij}, \quad \text{т.е.} \quad Y = X.$$

Итак, $X = XI$. Аналогично $IX = X$; поэтому

$$IX = X = XI.$$

К сожалению, обратная по умножению матрица может не существовать; однако если она существует, то она едиственная. Если матрица имеет обратную, то она называется обратной.

Пример 2. Не существует матрицы X такой, что

$$X\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix} = I.$$

Доказательство. Вычисление произведения дает

$$\begin{bmatrix}0 & X_{12}\\ 0 & X_{22}\end{bmatrix}.$$

Следовательно, какие бы значения компонент матрицы X не рассматривались, элемент (1, 1) произведения некогда не будет равен 1, откуда и следует требуемый результат. Таким образом, множество квадратных матриц заданного размера с определенными на нем операциями умножения и сложения образует кольцо.

Используя далее связь между бинарными отношениями на множестве и матрицами над (Z_2, $\wedge$, $\vee$), дадим следующие определения.

Транспонированной матрицей M называется матрица $M^{т}$ такая, что

$$M^{T}_{ij} = M_{ji}$$

(поэтому, если M получена из отношения σ, то $M^{т}$ может быть получена из отношения σ^{-1}); *транзитивное замыкание* M^{+} и *рефлексивное замыкание* M^{*} (изоморфны соответственно σ^{+} и σ^{*}) определяются следующим образом:

$$M^{+} = \sum_{n=1}^{\infty} M^{n}, \quad M^{*} = \sum_{n=0}^{\infty} M^{n},$$

где $M^{0} = I$, $M^{1} = M$ і $M^{n+1} = MM^{n}$ ($n \in$ N). (В некоторых случаях эти замыкания нельзя определить корректно (чтобы соответствующие ряды сходились), однако над (Z_2, $\wedge$, $\vee$) определение корректно, поскольку это — замкнутое полукольцо.)

В заключение заметим, что матрицы могут быть частично упорядосены путем поэлементного сравнения, а именно

$M \leq N$ тогда и только тогда, когда $M_{ij} \leq N_{ij}$ для всех i, j.

Из данного определения следует, что

$M \leq N$ тогда и только тогда, когда $M + N = N$,

при условии что + является операцией «максимум», подобной **или**.

12. Примеры из математической лингвистики

12.1. Синтаксические структуры

Между словами, образующими правильную русскую фразу, существуют различные лингвистические отношения. Указать явным образом эти отношения — это и значит описать синтаксическую структуру фразы.

Формальное описание свойств таких отношений и методов их выделения во фразе является одной из важных задач математической лингвистики. Поскольку здесь мы не имеем в виду обсуждать связь математической лингвистики с общей лингвистикой, то мы не будем углубляться в анализ лингвистического характера вводимых отношений, а будем апеллировать к тому знанию русского языка и его грамматики, которым наверняка обладает любой читатель данной книги.

Пусть дана некоторая русская фраза, и пусть M — множество вхождений слов в эту фразу. Мы фиксируем на M несколько важнейших грамматических отношений, определяемых ролью в этой фразе данного вхождения слова.

В дальнейшем мы говорим о *вхождении* слов, а не о словах, поскольку одно и то же слово может во фразе повторяться несколько раз, причем разные вхождения одного и того же слова могут играть разную роль и иметь разные связи.

Например, «Лазурь да глина, глина да лазурь, чего ж тебе еще? Скорей глаза сощурь, как близорукий шах над перстнем бирюзовым, над книгой звонких глин, над книжною землей, над гнойной книгою, над глиной дорогой, которой мучимся, как музыкой и словом». В этом стихотворении Осипа Мандельштама несколько раз повторяются слова: «как», «над», «книга», «глина».

В математической лингвистике изучаются также отношения между словами, не зависящие от их вхождений в тексты (скажем, «принадлежность одному роду» или «принадлежность одной части речи»). Но эти отношения (так называемые *парадигматические отношения)* являются отношениями на другом множестве—на множестве слов русского языка (Отметим, что термин «множество» здесь не вполне уместен, поскольку не существует общепринятого соглашения, что считать словами русского языка.). Мы же в этом параграфе изучаем только отношения между вхождениями слов в некоторую фразу (так называемые *синтагматические отношения).*

Начнем с того, что перечислим основные отношения между словами во фразе.

Простейшее из возможных отношений на M — это отношение следования: *x левее y.* Далее мы будем обозначать следование знаком неравенства. Таким образом, запись $x < y$ означает, что y расположен во фразе правей, чем x. Легко видеть, что отношение «быть левее» задает совершенный) строгий порядок на множестве M. (Но, как во все лингвистических формальных моделях, здесь возможны исключения: сноски нарушают линейность расположения слов в тексте и тем самым совершенность порядка.)

Редукцию отношения следования < мы обозначим символом A. Соотношение xAy означает, что y является непосредственным правым соседом слова x. Легко видеть, что xA^ny выполняется в том и только том случае, когда слово y отстоит от x ровно на n позиций вправо.

Казалось бы, такое чисто геометрическое отношение не имеет особого лингвистического смысла. Так естественно думать ввиду того, что в русском языке порядок слов сравнительно свободен. (Кстати, это связано с тем, что в русском языке имеется богатая система окончаний, достаточно полно показывающих связи слов в предложении. Поэтому русский язык может позволить себе роскошь меньше заботиться о порядке слов, чем, скажем, английский.) Однако и в русском языке порядок слов не вполне произволен с точки зрения грамматики и смысла текста. Например, отрывок «возьмем такое четное число, что...» не переделывается в отрывок «возьмем четное такое число, что...». В качестве второго примера укажем, что в русском языке существуют такие слова (предлоги), которые в тексте всегда предшествуют соответствующим существительным.

Второй важный тип отношений — это грамматическое управление. Управление—это отношение, обобщающее такие привычные типы отношений, как «определяемое — определение», «сказуемое — подлежащее», «сказуемое—дополнение» и т. п. Например, известные из грамматики утверждения «предлог „к" требует дательного падежа», «предлог „о" управляет предложным падежом» и т. д. означают именно то, что тот или иной предлог управляет существительным в таком-то падеже. В последнем случае мы имеем пример так называемого обязательного управления — предлог не может «повиснуть» в правильной русской фразе без управляемого существительного. Фразу «мой коллега был в кино с» можно даже понять, счесть осмысленной, но уже никак нельзя полагать грамматически правильной.

И. А. Мельчук выделяет в русском языке 33 типа грамматических управлений (или отношений непосредственной доминации). Но мы здесь будем называть *управлением* объединение всех этих отношений. Управление, обозначаемое далее символом →, является асимметричным отношением. Чтобы были понятны дальнейшие примеры, мы условимся, согласно сложившейся традиции, считать, что управление идет от определяемого к определению, от сказуемого к подлежащему, от предлога к существительному, от глагола к прямому дополнению, от глагола к предлогу. На основе этих соглашений и их естественных аналогов можно усвоить, как расставляются управления в конкретных текстах. (Признаемся, что есть сложные ситуации, когда разные лингвисты по-разному расставят управления в одной и той же фразе.) Ниже дается фраза с изображенным графом управления. Сначала мы приведем саму фразу с нумерацией слов, а затем уже на рис. 12.1 дадим граф управления:

«И	сатана,	привстав,	с	веселием	на	лике	лобзанием	своим
1	2	3	4	5	6	7	8	9

насквозь	прожег	уста,	в	предательскую	ночь	лобзавшие	Христа».
10	11	12	13	14	15	16	17

(А. С. Пушкин)

Союз «и», стоящий впереди, является тем самым сомнительным случаем. Его можно было бы рассматривать здесь просто как ритмическую вставку в текст. Обратите внимание на то, что стрелки на рис. 12.1 не пересекаются друг с другом. Как мы увидим ниже, это обстоятельство отнюдь не случайно.

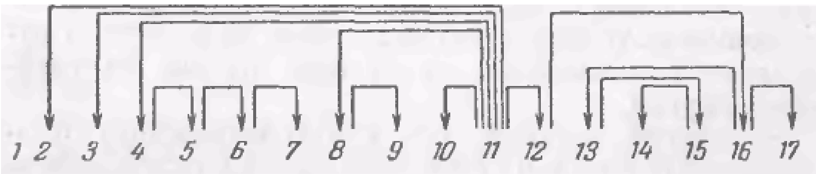


Рис. 12.1.

Транзитивное замыкание отношения управления называется *руководством* и обозначается символом $\Rightarrow$. Так, в графе на рис. 12.1 выполнено соотношение $x_{11} \Rightarrow x_7$. В силу леммы 7.7, если в графе управления нет контуров, то отношение руководства является строгим

порядком. Отношение руководства — это косвенное управление, управление через промежуточные инстанции.
Третий тип отношений между вхождениями слов во фразу — это согласование. Вообще говоря, под *согласованием* понимается наличие явно выраженных общих грамматических признаков, связывающих данную пару слов в коллектив. Например, согласование прилагательного и существительного по роду, числу и падежу. Отношение согласования мы будем обозначать символом σ. В предыдущей фразе мы имеем, например, соотношения $x_{14}\sigma x_{15}$ и $x_{15}\sigma x_{14}$, в то время как управление выполняется лишь в одну сторону: $x_{15} \rightarrow x_{14}$. Здесь видно уже одно отличие согласования от управления. Первое симметрично: определяемое и определение имеют, вообще говоря, согласованные грамматические показатели. Управление имеет направленность, оно асимметрично. Но согласование вовсе не является только «симметризацией» отношения управления. Во-первых, возможно управление без всякого согласования. Например, управление от глагола к обстоятельству типа «Вчера он уехал». Здесь согласованы (по роду и числу) только глагол и местоимение, но не наречие и глагол. Во-вторых, согласование может быть не связано ни с каким управлением.
Это обстоятельство проще пояснить не на русском тексте, а на алгебраических выражениях (Алгебраические формулы вполне естественно рассматривать как тексты некоторого искусственного языка.). В этих выражениях «согласованы» соответствующие друг другу левая и правая скобки.
В русском языке роль скобок выполняют конструкции вида: «если..., то...»; «или..., или...» и т. п. Соответствующие друг другу парные союзы находятся в отношении согласования, но ни один из них не управляет другим.
Четвертый важный тип синтагматического отношения — это отношение однородности («быть однородными членами предложения»). Мы будем обозначать это отношение символом *ν*. Типичный пример предложения с однородными членами:

«Швед, русский — колет, рубит, режет».
(А. С. Пушкин).

Здесь — два однородных подлежащих и три однородных сказуемых. Пример на однородные дополнения можно найти в других отрывках из А. С. Пушкина: «Сват Иван, как пить мы станем, непременно уж помянем трех Матрен. Луку с Петром, да Пахомовну потом». Или еще:

«И твое воспоминанье заменит душе моей силу, гордость, córованье и отвагу юных дней».

Пятый тип отношений носит несколько отличный от предыдущих характер. Дело в том, что всякая фраза русского языка довольно естественно членится на коллективы. Мелкие коллективы сами входят в более крупные. Такое членение фразы на коллективы (или, как принято говорить в лингвистике, *составляющие)* обеспечивает, в частности, возможность понимания фразы. Наша языковая интуиция позволяет нам довольно однозначно выделять составляющие в русской фразе. В следующем примере составляющие выделены скобками:

«(Все[1] это[2]) (сильно[3] (поколебало[4] (мою[5] (авторскую[6] уверенность[7]))))».

В сущности кавычки здесь тоже играют роль скобок, выделяющих максимальную составляющую. В число составляющих мы будем включать и отдельные слова. Итак, множество $\mathcal{M}$ составляющих состоит из некоторых непустых множеств вхождений слов в данную фразу. Отношение вхождения в составляющую на множестве $\mathcal{M}$ мы будем обозначать далее рис. 12.2.

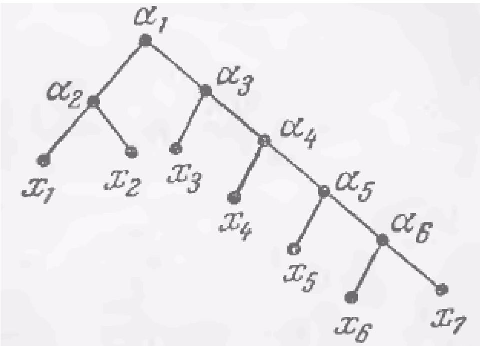


Рис. 12.2. Дерево составляющих.

Дерево составляющих, обычным знаком включения $\subset$. Тогда запись $y \subset \alpha_i$ обозначает: $x_j \in \alpha_i$, если $y = x_j$ — слово из фразы, и $\alpha_j \subset \alpha_i$, если $y = \alpha_i$ составляющая этой фразы, отличная от α_i. Из определения явствует, что *вхождение в составляющую* является строгим порядком. В рассмотренном примере мы выделим следующие составляющие:

$$\alpha_1 = \{x_1,\ x_2,\ x_3,\ x_4,\ x_5,\ x_6,\ x_7\},$$
$$\alpha_2 = \{x_1,\ x_2\},\quad \alpha_3 = \{x_3,\ x_4,\ x_5,\ x_6,\ x_7\},$$
$$\alpha_4 = \{x_4,\ x_5,\ x_6,\ x_7\},$$
$$\alpha_5 = \{x_5,\ x_6,\ x_7\},\quad \alpha_6 = \{x_6,\ x_7\}.$$

Граф для редукции данного отношения изображен на рис. 12.2. Обратите внимание на то, что этот граф является несимметричным деревом, которое наиболее сильно ветвится вправо. Это обстоятельство есть проявление языкового правила, а не случайное свойство примера.

* * *

Система отношений между элементами фразы (словами и составляющими) описывает формальны мобразом синтаксическую структуру фразы. Выбирая различные наборы содержательно определяемых отношений и описывая их формальные свойства, мы получаем ту или иную формальную модель синтаксической структуры фразы. До того, как говорить дальше о конкретных отношениях и их свойствах, полезно уточнить, что мы можем ожидать от формального описания лингвистических объектов и отношений между ними. Типичная ситуация в математической лингвистике может быть описана следующим образом. Мы исходим из некоторого класса однотипных лингвистических объектов (например, из совокупности фраз русского языка). Обычно каждый из этих объектов естественным образом расчленяется на элементы, т. е. его можно рассматривать как множество элементов определенного типа. Так, фразу можно рассматривать как множество вхождений слов (или слов и составляющих). Слово можно, в свою очередь, рассматривать как состоящее из морфем: корней, суффиксов, приставок, окончаний. Обычно мы, используя знание языка и его грамматики, умеем не только выделить элементы, составляющие данный лингвистический объект, но и установить между этими элементами некоторые отношения. Например, мы умеем указать управления во фразе, выделить согласования, однородные члены и составляющие.

Мы можем выделить некоторые, инвариантные относительно замены объекта, свойства этих отношений, т. е. свойства, которыми одноименные отношения обладают для любого объекта из выбранного класса. Например, отношение следования является для некоторого достаточно широкого класса фраз русского языка совершенным порядком.

Итак, нас интересуют отношения, которые могут быть более или менее однозначно определены на любом из лингвистических объектов, принадлежащих некоторому классу, и те свойства этих отношений, которые выполнены (вообще говоря) для любого объекта из этого класса. Иными словами, когда мы употребляем словосочетание «Отношение управления», то мы имеем в виду класс отношений (Это обстоятельство мы подчеркнули, написав «Отношение» с большой буквы.), каждое из которых на основе принятых соглашений определено на некоторой фразе русского языка. При этом для любой фразы русского языка определено некоторое отношение управления. Когда же мы в рамках математической лингвистики говорим о свойствах Отношения управления, то мы имеем в виду такие свойства, которые выполнены (опять-таки, вообще говоря!) для любого отношения управления в любой фразе. Например, свойство «каждое слово управляет не более, чем одним словом» выполняется в следующей фразе:

«Идёт, гудёт зелёный шум».

Однако это свойство не выполняется в очень многих других фразах, и оно, в нашем понимании, является не свойством Отношения управления, а только свойством данной фразы (или же — отношения управления в данной фразе).
Нас будут интересовать только инвариантные свойства Отношений. Однако дело обстоит не столь просто и с инвариантными свойствами. Некоторые свойства Отношений вытекают логически из их определений. Например, асимметричность управления означает просто, что мы условились считать, что управление может идти только от одного слова к другому (от сказуемого к подлежащему, но не наоборот). Свойство Отношения «входить в составляющую» быть строгим порядком проистекает из того, что это Отношение определено через включение множеств. Свойства лингвистических Отношений принципиально не могут быть незыблемыми просто потому, что носитель языка — человек — обладает свободой воли. Следовательно, он волен нарушать любое формальное правило или сознательно следовать ему. Когда мы пытаемся установить систему формальных правил, описывающих структуру языка, у нас часто возникает иллюзия, что дальнейшее развитие и уточнение этой системы когда-то в светлом будущем даст полностью адекватное описание языка. Но любая самая подробная система общих правил непрерывно нарушается живым развитием языка. Даже такое простое правило, что длина фразы не может быть слишком большой, может нарушаться. В книге польского писателя Ежи Анджеевского «Врата Рая» всего две фразы.

Вторая из них такая: «И шли целую ночь». Но, разумеется, текст первой фразы очень четко членится на коллективы. В частности, когда мы описываем свойства лингвистических Отношений и обнаруживаем, что эти свойства не столь просты, как казалось, то перед нами возникают два очевидных пути. Первый состоит в уточнении и поиске более сложной формулировки этих свойств (Сюда относятся поиски обобщенных определений проективности, введение разрывных составляющих и т. п.). Второй путь — попытаться по-иному определить сами эти отношения во фразе (Так, существуют различные соглашения о расстановке управлений в случае однородных членов, придаточных предложений и т. п.).

Попробуем выразить эту же мысль несколько строже. Переход от лингвистики к математической лингвистике состоит в том, что классу лингвистических (наблюдаемых или мыслимых) объектов мы соотносим список Отношений и их свойств (аксиом). Этот список будем, в соответствии с имеющейся в математике терминологией, называть *Теорией.* В этой Теории Отношения являются всего лишь названиями классов наблюдаемых в лингвистике отношений. Свойства Отношений должны быть сформулированы так, чтобы они имели смысл для настоящих отношений.

Множество с заданными на нем отношениями $A_1, A_2, ..., A_n$ называется *моделью Теории,* если установлено биективное соответствие между списком Отношений Теории и совокупностью отношений $\{A_1, A_2, \ldots, A_n\}$ и соответствующие отношения обладают всеми свойствами, предусмотренными данной Теорией.

Теория считается состоятельной для класса лингвистических объектов, если эти объекты, как множества элементов с соответствующими отношениями, в подавляющем большинстве являются моделями этой Теории. В нашем основном примере наблюдаемый лингвистический объект — это фраза русского языка.

Теория содержит пять перечисленных выше Отношений (в качестве вариантов можно рассматривать Теории, содержащие только часть этих Отношений). Свойства Отношений в этой Теории постулируются таким образом, чтобы они удовлетворялись для одноименных отношении в основной массе фраз русского языка (Подчеркнем, что моделью языка (в том смысле, как это понимают лингвисты) является Теория, а моделями Теории (в математическом смысле) служат, в частности, лингвистические объекты, моделируемые этой Теорией!). (Можно строить Теорию и с таким расчетом, чтобы она обслуживала все языки мира; аксиомы такой Теории являются лингвистическими универсалиями.)

Первый путь уточнения Теории состоит в более сложной формулировке свойств Отношений с тем. чтобы они удовлетворялись на большем количестве фраз. Второй путь — в уточнении соглашений об определении отношении во фразах.

Оба эти пути относительно полезны, но настоящего решения проблемы не дают. Остается третий путь — признать, что все наши формальные Теории (формальные модели языка) являются не самодовлеющими, а только отражающими глубинные объективные свойства живого языка. Эти модели отражают какую-то языковую норму, но язык может ее нарушить ради сохранения чего-то в данной ситуации более существенного. Для языка важно не перейти некоторый допустимый предел сложности, после которого речь перестает быть понятной. Поэтому нарушения формальных законов в живой речи возникают в сущности из-за стремления языка к сохранению глубинных законов. В силу этого стремления наблюдаемые свойства лингвистических Отношений приобретают гораздо больший смысл. Они перестают быть умозрительной или эмпирической схемой, а становятся характеристикой языковой нормы, отражающей глубинные свойства языка. Закон не теряет своей объективности, но оказывается глубже, чем выражающая его Теория. Однако вне формальных Теорий мы никогда и не подойдем к пониманию лингвистических законов. Более того, чем яснее Теория, чем она более явно выражена, тем легче уяснить ее связь с глубинными законами. Когда мы понимаем истинную цену формальной Теории (модели языка), мы яснее видим, что при всех кажущихся нарушениях языковых норм глубинные законы языка чрезвычайно устойчивы, а попытки их нарушить приводят к невосполнимым потерям.

Здесь напрашивается аналогия с глубинными нравственными законами. Ввиду очевидной условности любой формальной системы морали может показаться, что здесь вообще нет априорных законов, а существуют только созданные людьми соглашения. Однако в сфере нравственных законов действует эффект компенсации, о котором Вл. Соловьев выразился так: «Человек может не исполнить своей нравственной обязанности, но тогда он неизбежно теряет свое нравственное достоинство».

*** * ***

После этих небольших общих рассуждений перейдем к описанию формальных свойств введенных выше классов отношений.

1. Следование. Об этом отношении нельзя сказать ничего другого, кроме того, что оно есть совершенный строгий порядок.

Ясно, что случаи типа подстрочных примечаний к середине фразы, подстрочных или надстрочных пометок к отдельным словам нарушают совершенность (линейность) порядка, но являются теми самыми исключениями, которыми следует пренебречь в формальной модели.

2. Управление. В нормальном случае отношение управления обладает следующими свойствами:

1. Если выполнены соотношения $x_1 \to x_2 \to \ldots \to x_n$ (n>2), то невозможно $x_1 \to x_n$ (антитранзитивность).

2. Существует единственный элемент x, для которого соотношение $y \to x$ не выполнено ни при каком y.

3. Для всякого x существует не более одного такого y, что $y \to x$.

Из свойства 1 вытекает, что отношение управления асимметрично и его граф не содержит контуров. Лемма 7.7 позволяет тогда заключить, что транзитивное замыкание отношения управления (отношение руководства) является строгим порядком. Из условий 2, 3 можно вывести, что руководство является древесным порядком.

Нарушений свойства 1, по-видимому, не отмечалось в реальных фразах, т. е. руководство всегда является строгим порядком. Однако нарушение древесного порядка для руководства может возникнуть по причинам нарушений свойств 2 и 3. По существующим соглашениям вершиной графа управления может быть только сказуемое, т. е. лишь сказуемое во фразе может никем не управляться. Остальные члены предложения всегда имеют старшего (управляющего) в этой фразе. Но, когда во фразе имеются два однородных сказуемых, условие 2 автоматически нарушается. Так как, с другой стороны, подлежащее управляется всеми сказуемыми, при этом нарушится и условие 3. Это можно увидеть на следующем отрывке из стихотворения А. С. Пушкина:

«Всех чаще мне она приходит на уста и падшего
1 2 3 4 5 6 7 8 9
крепит неведомою силой».
10 11 12

На рис. 12.3 показан граф управления для этой фразы; отклонение от древесности проявляется в том, что подлежащее «она» имеет два управляющих слова. (Стрелка управления 10→8 поставлена условно, чтобы избежать изолированных вершин.)

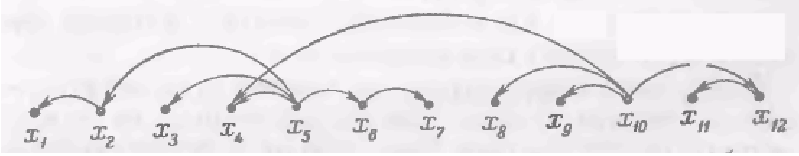


Рис. 12.3. Недревееная структура руководства.

Наблюдаемое пересечение стрелок (которое ниже мы назовем отклонением от проективности) связано не только с нарушением древесной структуры руководства, но и с нарушением естественного порядка слов ради поэтического ритма. При нормальном порядке: «Она приходит мне на уста всех чаще...» пересечения стрелок исчезнут. Отметим, что и при отклонении от древесности граф обычно остается связным.

3. **Согласование.** Это отношение симметрично и антирефлексивно (Поскольку кажется неестественным полагать слово согласующимся с самим собой. Возможны, разумеется, и иные соглашения.). В общем случае оно не транзитивно. Хороший пример нетранзитивности согласований дает фраза из следующего за этим абзаца: «...во фразе могут быть выделены группы, каждая из которых содержит...». Здесь согласованы пары «которых» — «группы» (по числу) и «группы» — «каждая» (по роду), но не согласована пара «которых»— «каждая».

4**. Однородность**. Это отношение симметрично и транзитивно. Будем считать, что слово, не входящее в группу однородных, не однородно и к самому себе, т. е. что однородность есть свойство группы, а не отдельного слова. Тогда во фразе могут быть выделены группы, каждая из которых содержит по несколько однородных членов, а остальные не входят ни в одну из групп.

5. **Вхождение в составляющие**. Уже из определения следовало, что вся фраза есть (максимальная) составляющая. Это дает нам условия:

1. Для всякого $x \in \tilde{M}$ существует такое *y*, что либо $y \subset x$, либо $x \subset y$ (если фраза сосюит больше чем из одного слова).

2. Существует единственный элемент, который не входит ни в какую составляющую.

Следующее содержательное лингвистическое утверждение состоит в том, что составляющие не могут частично перекрываться. Они либо не содержат общих элементов, либо одна содержит другую. В формальных терминах это означает:

3. Если $x \subset y$ и $x \subset z$, то либо $y \subset z$, либо $z \subset y$, либо $y = z$.

К этим свойствам можно добавить

4. Антирефлексивность.
5. Транзитивность,

вытекающие из определения строгого порядка.

Эти пять условий означают, что отношение «входить в составляющую» является древесным порядком. Этот лингвистический факт— возможность представления любой фразы в виде дерева составляющих — явился основой для создания серии формальных моделей (начиная с наиболее известной порождающей модели Н. Хомского), описывающих процесс «порождения» фразы языка путем последовательной подстановки вместо каждой составляющей содержащихся в ней составляющих или слов русского языка.

Подчеркнем важное обстоятельство, которое иной раз забывается. **Свойство текста расчленяться в дерево составляющих есть первичный лингвистический факт, полученный в результате осмысления конкретных лингвистических наблюдений, а не следствие принятой модели порождения. Наоборот, создание моделей порождения текстов стало возможным только после осознания того, что текст естественно членится на составляющие и это членение обладает древесным порядком.** После этого можно уже искать различные формальные интерпретации этого факта и строить всевозможные модели порождения (**кроме модели Н. Хомского можно указать на реляционные грамматики Ирены Беллертовой, матричные грамматики С. Абрахама, диспозиционные грамматики В. Б. Борщева и Ю. А. Шрейдера, грамматики с управлением Э. Д. Стойкого).** В. Б. Борщев обратил внимание на то, что и в формальных грамматиках, не описывающих процесса порождения (имеются в виду так называемые окрестностные грамматики В. Б. Борщева), возникает естественная структура составляющих. Мы подчеркиваем данное обстоятельство именно потому, что в результате изучения математической лингвистики возникает часто впечатление о том, что возможность членения на составляющие есть исключительно свойство языков, описываемых порождающими подстановочными грамматиками. На самом же деле ситуация в точности противоположна — **возможность описать естественный язык порождающей грамматикой есть следствие существования составляющих в языке и некоторых гипотез о составляющих,** в которые мы здесь не имеем возможности вникать.

* * *

До сих пор мы рассматривали только свойства, присущие каждому из отношений отдельно, но более содержательны свойства, связывающие различные отношения. К изучению таких свойств мы и приступим.

Следование и управление

Отношения следования и управления во фразе нормально связаны так называемым условием проективности. Фраза называется *проективной,* если дважды упорядоченное множество $\langle M, <, \Rightarrow \rangle$ удовлетворяет условию $П_1$ (здесь M — множество вхождений слов во фразу, < — отношение следования, $\Rightarrow$— отношение руководства.
На рис. 12.4 изображены два примера проективных фраз.

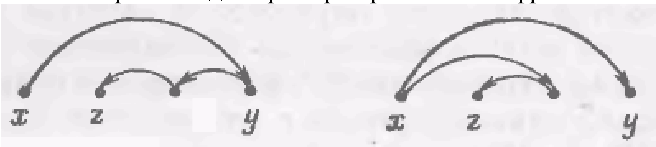


Рис. 12.4. Свойство проективности.

Условимся рисовать граф управления так, чтобы слова во фразе располагались на прямой в их естественном порядке, задаваемом отношением следования, а все стрелки, изображающие управления, проводились бы с одной стороны этой прямой (над ней). При таком соглашении часто используют иное определение проективности. Фраза называется *проективной,* если дважды упорядоченное множество $\langle M, <, \Rightarrow \rangle$ удовлетворяет условию $П_2$. Так как, согласно теореме 7.18, условие $П_2$ влечет условие $П_1$, то проверка непересечения стрелок и непокрытия максимальных элементов гарантирует проективность в обоих смыслах.
Из теоремы 7.19 следует, что в случае, когда руководство образует древесный порядок, оба определения проективности равносильны. В случае недревесной структуры, изображенной на рис. 12.3, мы имеем пример непроективной фразы (в обоих указанных смыслах).
Фраза называется *квазипроективной,* если стрелки управления можно провести так, чтобы они не пересекались. На рис. 12.5 изображена квазипроективная, но не проективная фраза. В этой фразе выполнены соотношения $x \to y$ и $x < y < z$, но не выполнено $z \Rightarrow y$.

Рис. 12. 5. Квазипроективная структура

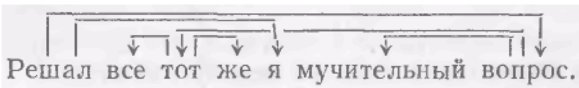
Удобную формулировку условия проективности можно получить еще следующим образом. Условимся проводить дополнительную стрелку управления от знака препинания, отмечающего конец фразы, к сказуемому. Фразу (Отношение руководства которой является деревом) можно назвать *проективной,* если пополненный указанным способом граф управления можно нарисовать без пересечения стрелок. В самом деле, последнее условие равносильно тому, что основные стрелки управления не дают пересечений и путь к корню дерева не перекрыт накрывающей стрелкой.

Существует и четвертый вариант определения проективности. Пусть отношение руководства является деревом. Фразу можно назвать *проективной,* если выполняется условие $П_3$. Из теоремы 12.20 вытекает, что при сделанном допущении это определение проективности равносильно предыдущим.

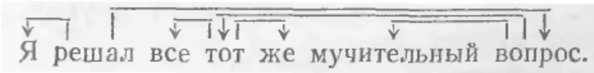
В этом варианте определения виден содержательный смысл слова «проективность»: отмеченные точки должны беспрепятственно проектироваться на горизонтальную прямую, лежащую выше всех точек, а отрезки должны не перепутываться при таком проектировании.

К сожалению, в некоторых лингвистических работах это определение приводится неточно. Так, в книге Ю. Д. Апресяна «Идеи и методы современной структурной лингвистики» (Москва, «Просвещение», 1966, стр. 248) опущено условие непересечения отрезков. Но тогда, как показывает пример на рис. 7.12, фраза может быть непроективной в смысле первых двух определений. В частности, фраза «Читал человек раскрытую веселый книгу» имеет как раз ту структуру управлений, что дана на рис. 7.12. Однако по определению Ю. Д. Апресяна ее пришлось бы счесть за проективную.

Приведем еще пример непроективной структуры из А. Блока:



Ясно, что такой порядок слов возник из-за внутренней ритмики стиха. В нормальном для прозы порядке слов и нормальной прозаической ритмике все вполне проективно:



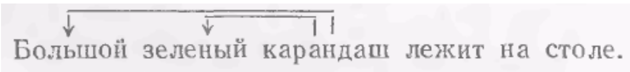
К счастью для поэтов, русский язык дает широкие возможности образования непроективных структур, но не создает их без особой на то надобности. Впрочем, в современной русской литературной прозе попадаются неслыханно непроективные структуры.

Однородность, следование и управление

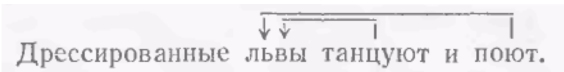
Можно сформулировать несколько простых свойств, связывающих порядок слов во фразе, отношение управления и отношение однородности. Предположим, что выполнено соотношение xvy. Тогда, вообще говоря, справедливы следующие утверждения:

1) Если $z \to x$, то $z \to y$ и z не находится между x и y.
2) Если $x \to z$ и $y \to z$, то z не находится между x и y.
3) Если $x < y$, $x \to z$ и $y \to w$, то $z < w$, $x < w$, $z < y$.

Первое свойство означает, что однородными членами управляют одни и те же слова и что управляющее находится по одну сторону от обоих управляемых:



Второе свойство означает, что общее управляемое находится по одну сторону от однородных управляющих:



Третье свойство состоит в том, что области управления однородных слов не перепутываются:

Красный мак и белый ландыш стоят в вазе.

Можно показать, что, при указанных условиях на однородности, во фразе можно ввести разумную скобочную структуру (К. И. Бабицкий, О дистрибутивной теории предложений с сочиненными частями, НТИ, 1967, № 6.)

Составляющие и следование

Основное условие, которое лежит в основе всех подстановочных грамматик, состоит в неразрывности составляющих. Составляющая α называется *неразрывной,* если из того, что $x \in \alpha,\ y \in \alpha$ и *z* находится между *x* и *y,* следует $z \in \alpha.$

Неразрывная составляющая занимает целый отрезок во фразе. Представление о том, что все составляющие неразрывны, лежит в основе упоминавшихся выше подстановочных порождающих грамматик. На самом деле в русском языке (а также английском, немецком и др.) существуют и разрывные составляющие. Например, сложное глагольное время легко приводит к разрывным составляющим: «Он *будет* завтра *читать* лекцию». Такой порядок слов возможен в русском языке, а для немецкого языка вынесение инфинитива в конец фразы является нормой. Можно такие случаи рассматривать как трансформации нормальных предложений или же по-иному выделять составляющие во фразе, не считая обязательным включение основного и вспомогательного глагола в одну составляющую.

Гипотеза о том, что все составляющие неразрывны, равносильна следующему. Возьмем каждую составляющую в скобки. В силу неразрывности составляющих на каждую из них уйдет одна пара скобок. В силу древесности структуры составляющих и неразрывности для двух пар скобок возможны лишь такие варианты расположения $[(\ \ )],\ [\ \ ](\ \ )$ и невозможно расположение вида $([\ \ )]$. Такая расстановка скобок называется *правильной скобочной структурой.*

Пусть $\mathfrak{M}$ — множество составляющих некоторой фразы. Отношение следования < во фразе индуцирует на $\mathfrak{M}$ строгий порядок, определяемый следующим образом. Будем полагать $\alpha_i < \alpha_j,$ если для любых представнтелей .$x_i \in \alpha_i$ и $x_j \in \alpha_j$ выполнено $x_i < x_j.$ Аналогично устанавливается отношение следования между словом x_i и составляющей $\alpha_j : x_i < \alpha_j,$ если x_i лежит левее любого представителя из этой составляющей, и $\alpha_j < x_i$, если x_i лежит правее составляющей. Ясно, что отношение следования на $\tilde{M}$ уже не является совершенным порядком, т. к. при $\alpha_i \subset \alpha_j$ не выполняется ни $\alpha_i < \alpha_j$, ни $\alpha_j < \alpha_i$. Более того, отношение следования выполняется на тех и только тех парах, для которых не выполняется отношение включения. Нетрудно видеть, что множество $\tilde{M}$ с отношениями $\subset$ и < является упорядоченным деревом (в смысле § 4 гл. 7). Глубина этого дерева является важной лингвистической характеристикой фразы. На рис. 12.6 приведено дерево составляющих для простой фразы.

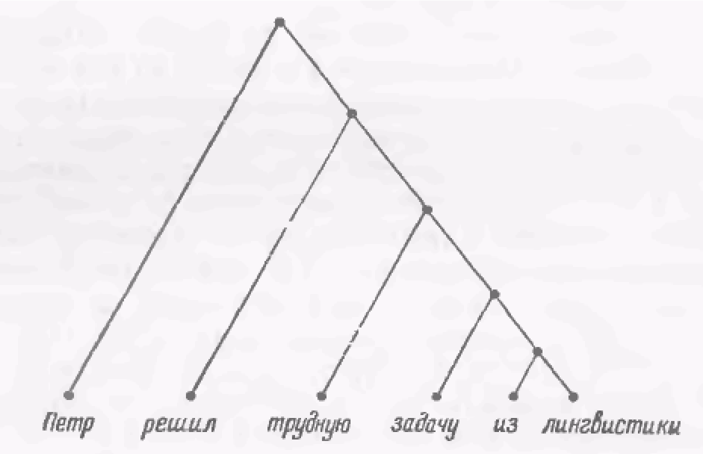


Рис. 12.6.

Глубина этого дерева или, как часто говорят в математической лингвистике, *глубина* этой *фразы* равна единице.

Эту характеристику фраз впервые ввел В. Ингве, обративший внимание на тот факт, что глубина реальных фраз в языке ограничена. Он же высказал гипотезу, что ограниченность глубины связана с ограниченностью человеческой памяти, сказывающейся в процессе порождения речи или ее восприятия.

Количественная формулировка гипотезы Ингве состоит в том, что для любой фразы естественного языка глубина *у* дерева составляющих ограничена:

$$\gamma \leqslant 9. \tag{12.1}$$

(Ортодоксальные лингвисты пишут вместо (12.1) неравенство $\gamma \leqslant 7 \pm 2.$)

Эта гипотеза эмпирически оправдывается. Средняя глубина, посчитанная по фразам русского языка, оказывается заметно меньшей, чем 9.

Подчеркнем два важных обстоятельства. Первое из них состоит в том, что гипотеза Ингве в сущности прямо не связана с какими-то предположениями о процессе порождения речи. Она говорит только об асимметрии дерева составляющих, построенного над фразой русского или другого естественного языка. При этом не имеет никакого значения, применима ли к данному языку та или иная формальная модель порождения. Второе обстоятельство состоит в том, что в оценке (12.1) существенны не какие-либо мистические свойства числа). Этой оценке можно придать следующую формулировку: **глубина фразы естественного языка не выходит за пределы «средних по А. Н. Колмогорову» чисел.** Напомним, что число *п* называется «средним по Колмогорову», если человек практически способен перебрать все подмножества множества из *п* элементов.

В отличие от числа 7 это обстоятельство не кажется ни мистическим, ни случайным.

Если мы все же будем рассматривать схему порождения фразы в какой-либо подстановочной грамматике, то оказывается, что глубина дает оценку для минимально необходимой памяти, используемой в процессе порождения. Именно, если σ есть минимальное количество символов, которое мы обязаны хранить на каждом шаге порождающей процедуры, то можно доказать, что

$$\gamma + 1 \leqslant \sigma \tag{7.2}$$

и для контекстно-свободной грамматики Хомского это неравенство обращается в равенство.

Равенство $\gamma + 1 = \sigma$ для контекстно-свободных грамматик Хомского принадлежит Ингве.

Составляющие и управление

Связь между структурой составляющих и управлениями во фразе можно выразить (в нормальном случае) в виде следующих свойств. Пусть *S*(α) — совокупность всех слов, входящих в составляющую α. Тогда

1) Каждое *S*(α) есть дерево по отношению руководства.

2) Если α и α_1 — составляющие, то отношение управления может выполняться только для корней деревьев $S(\alpha)$ и $S(\alpha_1)$.

Иначе говоря, управление от одной составляющей к другой может передаваться только через главные элементы этих составляющих. При некоторых дополнительных условиях свойства 1) и 2) гарантируют проективность управлений. Будем говорить, что α_1 и α_2 суть *соседние составляющие,* если $\alpha_1 < \alpha_2$ (или $\alpha_2 < \alpha_1$) и не существует никакого элемента *z,* лежащего между ними:

$$\alpha_1 < z < \alpha_2 \quad (\alpha_2 < z < \alpha_1).$$

Система составляющих называется *полной,* если для любых двух несовпадающих и несоседних элементов фразы (слов) *x* и *y* существуют такие соседние составляющие α_1 и α_2, что

$$x \in S(\alpha_1) \text{ и } y \in S(\alpha_2).$$

(Легко убедиться, что полнота системы составляющих равносильна тому, что дерево составляющих *бинарно,* т. е. нз каждой вершины выходит не более двух отростков.)

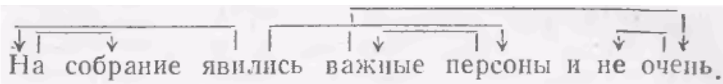
Оказывается, если система составляющих полна и выполняются условия 1), 2), то фраза проектнвпа. Это следует непосредственно из теоремы 7.21.

* * *

Остановимся еще раз на причинах, по которым в реальных фразах нарушаются описанные свойства синтаксической структуры.

Первая из них: говорящий сознательно нарушает нормальную структуру предложения, чтобы добиться выполнения какого-то иного свойства. Мы уже видели, что ради сохранения поэтического ритма часто привлекают непроектнвные структуры. Ингве убедительно показал, что непроективиость может возникнуть и тогда, когда порядок слов, обеспечивающий проективность, ведет к нежелательному росту глубины фразы.

Другая важная причина возникновения отклонений от нормы и, в частности, от проективности состоит в наличии эллипсисов. Рассмотрим следующий пример непроективной фразы:



Ясно, что это предложение является эллипсисом от следующего проективного предложения:

На собрание явились важные персоны и ие очень важные персоны.

Итак, мы имеем исходное предложение с отношениями следования и управления и его гомоморфизмы в другое предложение, на котором эти отношения индуцируются как α-образы. Но мы уже видели, что при переходе к α-образам свойства отношений могут портиться. То же происходит и здесь. В некотором смысле эллипсис можно также рассматривать как своего рода компенсацию: экономится число слов в предложении за счет ухудшения синтаксической структуры.

Третья причина в сущности двойственна предыдущей. Появление однородных членов можно рассматривать как «расклеивание» некоторого исходного предложения без однородностей. В данном случае мы имеем дело с корреспонденцией исходных отношений следования и управления, при которой также портятся свойства отношений.

Проведенный анализ синтаксических структур около 11000 английских предложений (большей частью сложных) показал, что около 500 из них являются пепроективными. Подавляющее большинство этих непроективностей было связано с эллипсисами и однородными членами.

12.2. Общее понятие текста

Как мы увидели из предыдущего параграфа, фраза естественного языка есть не просто цепочка слов, а множество с системой отношений. С другой стороны, *текст* можно представлять составленным из слов, букв, слогов, словосочетаний и т. д. Поэтому оказывается удобным сформулировать общее понятие текста, которое годилось бы для весьма разнообразных лингвистических ситуаций.

Мы попробуем изложить здесь некоторое достаточно общее представление о понятии текста, возникшее у Ю.А. Шрейдера совместно с М. В. Араповым и В. Б. Борщевым из попыток единообразного подхода к различным лингвистическим объектам. Интуитивно, **текст — это первичный материал лингвистического исследования.** Поэтому естественно требовать, чтобы слово, фраза или последовательность фраз русского языка могли быть интерпретированы как текст с формальной точки зрения. Однако не менее **естественно требовать, чтобы таблица, набор дескрипторов (ключевых слов), химическая или математическая формула также могли рассматриваться как частный случай общего понятия текста**. Такое требование во всяком случае оправдано захватническими устремлениями современной лингвистики.

Представим себе теперь, что текст подвергнут предварительной формальной обработке; стоит ли считать, что теперь это не текст, но некоторый иной объект высшей (или низшей) природы? **Нам представляется, что фразу с расставленными управлениями (или**

преобразованную каким-то другим способом) стоит рассматривать как некоторую разновидность текста. Ведь и самый классический лингвист редко имеет дело с непосредственной речью. Сама запись речи через формальные значки-буквы есть уже некоторая обработка исходного материала. Филолог, интересующийся древнерусским языком, имеет дело не столько с рукописями, сколько с печатными их публикациями, где слова расчленены, буквы стандартизованы.

Попробуем, сначала — не формально, разобраться, из каких существенных компонент слагается текст. Разумеется, **текст составлен из знаков. Но еще до расстановки конкретных знаков нужно определить позиции (места), где разрешается ставить знаки, и отношения между этими местами. Следующий шаг состоит в том, чтобы осознать первоочередную роль отношений между местами. Так, структура обычного текста определяется прежде всего тем, что знаковые позиции образуют линейную последовательность, т. е. между местами определено отношение совершенного порядка. Структура таблицы определяется тем, что между местами в таблице существуютдва отношения порядка: «горизонтальное» и «вертикальное».**

Поэтому целесообразно «места» рассматривать как элементы абстрактного множества *М,* на котором определена система отношений. Отсюда естественно возникает

Определение 12.1. *Синтаксической схемой* $S = \langle M; A_1, \ldots, A_n \rangle$ называется множество *М* с заданными на нем отношениями $A_1, \ldots, A_n$.

Это понятие по существу совпадает с понятием *модели* по А. Тарскому. Важность математической теории моделей для описания лингвистических ситуаций, по-видимому, впервые четко сформулировали В. Б. Борщев и М. В. Хомяков в работе «Окрестностные грамматики и модели перевода» (НТИ, сер. 2, 1970, №3 и № 4).

Множество *М* мы будем называть *несущим множеством.*

Пусть теперь зафиксировано некоторое множество $\mathfrak{A}$, которое мы будем называть *алфавитом.* Тогда отображение $\varphi\colon M \to \mathfrak{A}$ можно интерпретировать как расстановку знаков алфавита на местах: каждому месту (элементу множества *М)* сопоставляется некоторый знак (элемент алфавита $\mathfrak{A}$).

Мы получаем

Определение 12.2. *Текстом* $T = \langle S, \varphi \rangle$ называется синтаксическая схема *S* с заданным отображением φ несущего множества *М* в алфавит $\mathfrak{A}$.

Хотя это определение может показаться чересчур абстрактным для такого простого и, казалось бы, первоначального понятия как текст, оно в сущности только выражает в точных терминах все то, что мы обычно вкладываем в понятие **текста: выбор исходного алфавита, т. е. набора простейших символов, выбор синтаксической схемы, помещение символов алфавита в различные места синтаксической схемы, отношения между различными вхождениями символов в данную синтаксическую схему**. Следующий ряд примеров показывает, насколько данное определение текста является общим.

Пример 1. Алфавит $\mathfrak{A}$ есть список словоформ русского языка, S — конечное множество M с единственным отношением < совершенного строгого порядка. Тогда **текст — это отрезок натурального ряда с приписанными каждому номеру словоформами**. Говоря менее формально, текст — это любая линейная последовательность русских словоформ, может быть, с повторениями. Иначе говоря, такой текст — это просто цепочка вида

$$x_1, x_2, \quad . . . x_n,$$

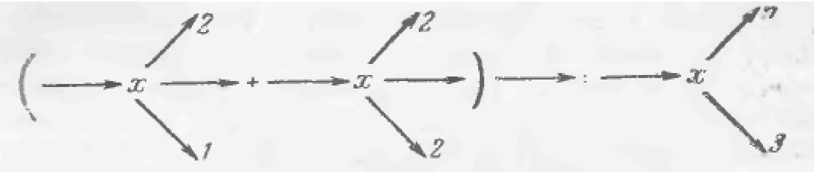
где все x_i — русские словоформы. В частности, любая фраза русского языка может рассматриваться как текст такого вида. Можно было бы также расширить алфавит $\mathfrak{A}$, внеся туда все знаки препинания и цифры.

II р и м е р 2. Алфавит $\mathfrak{A}$ — тот же, что и раньше, по M — конечное множество с четырьмя отношениями: следования, управления, согласования и однородности — обладающими свойствами, описанными в предыдущем параграфе. Тогда текст — это последовательность русских словоформ с основными синтаксическими отношениями.

П р и м е р 3. Пусть $\mathfrak{A}$ - кириллический алфавит, а на множестве M задан совершенный строгий порядок. Тогда текст — это конечная последовательность знаков Кириллицы. В частности, каждое русское слово можно рассматривать как текст такого вида, т. е. последовательность букв обычного русского алфавита (одного из основных современных вариантов Кириллицы).

Удобно рассматривать синтаксические схемы вида: $S = \langle M, R_1, R_2, R_3 \rangle$, где M — конечное множество, каждое из отношений R_1, R_2, R_3 есть редукция отношения строгого порядка и между любыми двумя элементами множества M может выполняться только одно из этих отношении. Эти отношения имеют следующую содержательную интерпретацию: R_1 — «непосредственно следовать», R_2 — «стоять над, быть верхним индексом», R_3 — «стоять под, быть нижним индексом». С помощью таких синтаксических схем можно ввести классы текстов в двух следующих примерах:

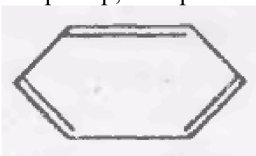
П р и м е р 4. Пусть $\mathfrak{A}$— алфавит, составленный из латинских и греческих букв, цифр и алгебраических знаков (скобки и знаки операций). Тогда любая алгебраическая формула может рассматриваться как текст с описанной выше синтаксической схемой. Например, формула $(x_1^0 + x_2^2) : x_3^n$ имеет спитаксическую структуру вида



где стрелками указано выполнение отношений R_1, R_2, R_3.

Пример 5. Пусть $\mathfrak{A}$—множество цифр и символов химических элементов. Текстами здесь являются обычные линейные химические формулы типа H_2O.

П р и м е р 6. Пусть теперь алфавит $\mathfrak{A}$ состоит из текстов предыдущего вида. Синтаксическая схема имеет вид $S = \langle M, R_1, R_2, \ldots \rangle$, где R_1, R_2, R_3, ... — отношения, интерпретируемые как типы химических связей. При этом для любой пары элементов из M может быть выполнено только одно из отношений R_1, R_2, R_3,… Например, изображение бензольного кольца



есть синтаксическая схема с двумя типами отношений валентности. Тем самым задается класс текстов, имеющих вид структурных формул органической химии.

Здесь мы столкнулись с важной ситуацией, когда тексты одного уровня образуют алфавит для текстов следующего уровня. Впрочем, мы уже видели, что словоформы русского языка суть тексты в обычном русском (кириллическом) алфавите. Сами же словоформы могут рассматриваться как элементы алфавита, в котором записаны русские предложения. (Кстати, и сами буквы можно рассматривать как тексты в алфавите Морзе из точки и тире.)

Пример 7. Пусть алфавит $\mathfrak{A}$ состоит из совокупности дескрипторов для некоторой области науки или техники (грубо говоря,

дескрипторы **— это основные термины данной области, с помощью которых можно охарактеризовать содержание некоторогодокумента — статьи, реферата и т. п**.). Множество *M* не имеет никаких отношений. Тогда текст — это просто набор дескрипторов без всяких связей между ними (Тексты с тривиальной синтаксической схемой — без отношений — называются иногда *мешками*). Такие тексты используются в так называемых *информационно-поисковых системах без грамматики* в качестве *индексов* (или *поисковых образов)* документов, позволяющих автоматически отыскивать нужный потребителю документ.

Обсудим несколько подробней пример 2 с точки зрения традиционной лингвистической терминологии. В обычном русском тексте явно задается только одно отношение — линейный порядок слов во фразе. Итак, **синтаксическая схема для обычного текста *T* — это конечное множество *M* с совершенным порядком**. **Текст над этой синтаксической схемой — это цепочка словоформ русского языка, т. е. текст в обычном смысле.** В процессе понимания текста мы явно или неявно устанавливаем дополнительные грамматические и смысловые отношения между словами и, в частности, можем вносить в текст новые элементы (например, символы составляющих). Тем самым в процессе понимания (или в процессе автоматического анализа) образуется новый текст T' над несущим множеством $M' \supseteq M$ с заданной системой отношений (управление, согласование, однородность, вхождение в составляющую и, быть может, многие другие). Формально текст T' является также текстом в смысле нашего определения. Но лингвистический смысл текста T' отличен от смысла исходного текста *T*. Для лингвиста естественно было бы тексту T' присвоить специальное название (например, *проанализированный текст* или *ультратекст* или что-нибудь более красивое). Мы не будем здесь нарушать привилегии лингвистов и не будем вводить нового термина. Нам важно только отметить формальное сходство *T* и T' (и тот и другой суть тексты над некоторым множеством с отношениями) и различие по существу: первый есть текст, данный в непосредственном наблюдении, а второй — некоторая конструкция, описывающая (скорее всего неполно) процесс понимания (а, может быть, и порождения). Синтаксическая схема текста T' определяет структуру синтаксических отношений исходного текста *T,* которые в исходном тексте *T* не выражены явно. Итак, **синтаксическая структура — это текст, очищенный от конкретных слов, но с явно указанными контекстными отношениями.** Иногда *синтаксической структурой* называют то, что мы обозначили T', но это не естественно.

Синтаксическая структура — это не конкретный текст T', а то общее, что есть у одинаково синтаксически устроенных текстов. Например, если есть два исходных текста T = «Маша ест кашу» и T_1 = «Петя читает книгу», то проанализированные тексты T' и T будут различны, хотя синтаксические схемы здесь, очевидно, одинаковы.

В действительности, интересно рассматривать не отдельные тексты, а классы однотипных текстов — текстов с аналогичной синтаксической схемой и общим алфавитом. Что такое «общий алфавит», понять легко, но что такое «аналогичная синтаксическая схема» — это еще требует разъяснения. Заметим, что в каждом из рассмотренных примеров мы имели дело именно с классами текстов.

Так, в примере 1 синтаксической схемой являлось любое конечное множество с совершенным порядком. В этом примере мы фактически имели дело с некоторой знаковой системой, определяемой выбором алфавита и условием, что «места» в текстах упорядочены.

В примере 2 мы задали класс текстов тем условием, что на несущем множестве обязаны быть определены четыре отношения с фиксированными свойствами.

Теперь мы попробуем несколько точнее определить понятие знаковой системы. Напомним, что в 1.12.1 мы условились называть *Теорией* список символов отношений и свойств этих отношений («аксиом»). Подразумевается, что свойства разрешается формулировать в таком виде, чтобы они обретали смысл, если символы отношений интерпретированы как отношения на некотором непустом множестве. Например, Теория может состоять из одного символа < и «аксиом»:

1) ни для какого x невозможно $x < x$;
2) если $x < y$ и $y < z$, то $x < z$;
3) для любых несовпадающих x и y выполнено либо $x < y$, либо $y < x$.

Эти аксиомы являются бессмысленными (но синтаксически правильными) фразами, пока не указана интерпретация, т. е. конкретное множество с отношением. Но как только переменные x, y, z, . .. мы станем интерпретировать как элементы некоторого множества M, эти аксиомы превратятся в осмысленные утверждения, говорящие, что отношение < есть совершенный строгий порядок на M.

Более точно (с точным определением понятия синтаксически правильной фразы) понятие Теории определяется в математической теории моделей.

Пусть теперь выбраны некоторая Теория и алфавит $\mathfrak{A}$.

Определение 12.3. *Знаковой системой* называется множество текстов $T = \langle S, \varphi\rangle$ с синтаксическими схемами $S = \langle M,\ A_1,\ \ldots,\ A_n\rangle$, у которых отношения A_1, A_2, ..., A_n взаимно-однозначно соответствуют символам отношений данной Теории и удовлетворяют аксиомам этой Теории, а φ есть отображение несущего множества *М* в фиксированный алфавит $\mathfrak{A}$.

Подчеркнем, что в знаковой системе зафиксированы только алфавит и Теория, а множества *М* могут быть разными.

Например, знаковая система может состоять из всех линейных последовательностей русских словоформ. Здесь фиксирован алфавит (множество русских словоформ), Теория (указано, что в синтаксических схемах есть единственное отношение — совершенный строгий порядок), но несущее множество, задающее длину цепочки, может быть произвольным.

<u>Языком</u> **в математической лингвистике обычно называется некоторое множество текстов в фиксированной знаковой, системе**. Примером языка может служить множество таких цепочек, составленных из знаков алфавита, которые удовлетворяют определенным условиям или порождаются некоторой процедурой (т. е. описываются некоторой «грамматикой»). Если же класс синтаксических схем состоит из конечных множеств с совершенным порядком, то **<u>язык</u> — это некоторое множество конечных цепочек, составленных из элементов алфавита** $\mathfrak{A}$.

В рамках математической теории моделей знаковая система — это множество моделей некоторой теории, для которых заданы отображения в фиксированный алфавит.

Следует подчеркнуть одно очень важное обстоятельство. Когда мы рассматриваем знаковую систему естественного языка, то, как бы мы ни выбирали допустимый класс синтаксических схем, или, что равносильно, Теорию, множество реально встречающихся текстов представляет всегда очень малую долю от всех возможных текстов данной знаковой системы.

По-видимому, здесь мы сталкиваемся с принципиальным отличием лингвистических структур от привычных физических моделей. В физике мы привыкли, что **все фазовые пространства, т. е. совокупности возможных состояний физической системы, устроены как гладкие многообразия в эвклидовом (или ином) пространстве**. Множество всех осмысленных текстов естественного языка имеет какую-то принципиально иную геометрическую структуру, для понимания которой у нас еще не выработалась соответствующая математическая интуиция. В этом, по-видимому, коренятся многие существенные трудности описания естественных языков. Весьма вероятно, что это обстоятельство есть общая трудность для математического моделирования биологических систем.

Рассмотрим теперь множество M с заданными отношениями A_1, A_2, ... , A_n. Возникает естественная проблема экономного описания этих отношений. С такой проблемой мы уже столкнулись при описании строгих порядков (на конечных множествах): оказалось, что отношение порядка можно задать с помощью редукции этого отношения.

Следующая постановка этой проблемы принадлежит К. И. Бабицкому. Пусть отношения A_1, A_2, ... , A_n обладают следующими свойствами:

1) $A_i \cap A_j = \varnothing$ при $i \neq j$;

2) $A_1 \cup A_2 \cup \ldots \cup A_n$ есть полное отношение.

Эти свойства означают, что для каждой пары $\langle x, y \rangle$ выполнено ровно одно соотношение xA_1y. Проблема в простейшем варианте состоит в том, чтобы определить на множестве M такие отношения B_1, B_2,...,B_m, чтобы 1) каждое отношение B_j выполнялось ровно для одной пары элементов и 2) для любой пары x и y было однозначно определено произведение $B = B_{i_1} B_{i_2} \ldots B_{i_k}$ так, что xA_1y равносильно либо xBy, либо yBx.

Простейшее решение этой проблемы состоит в том, что на множестве M любым способом устанавливается совершенный порядок, а следовательно, нумерация: $\{x_1, x_2, \ldots, x_p\}$. Тогда мы полагаем $x_i B_i x_{i+1}$.

Недостаток этого решения в том, что оно определяется не самой Теорией, т. е. свойствами отношений A_i, а проводится для конкретной реализации Теории на множестве M.

Ясно, что более общее решение может быть осуществлено только в предположении каких-то существенных алгебраических свойств синтаксической схемы.

12.3. Модели сочетаемости

В этом параграфе мы рассмотрим сравнительно частную модель, иллюстрирующую полезность рассмотрения в математической лингвистике отношений толерантности. Начнем с нескольких замечаний общего порядка.

Разработанные в математической лингвистике методы имеют известный предел применимости, обусловленный, по-видимому, ограниченностью концепций, на которых эти методы основываются. Как только нам хочется учесть при описании языка сравнительно тонкие индивидуальные свойства языковых единиц (слов, морфем, предложений), для описания которых требуется учитывать десятки и даже сотни признаков, мы вынуждены констатировать отсутствие адекватного математического аппарата. Не хватает способа описания «размытых» моделей. Так, например, существует значительная разница между математическим описанием семантических и синтаксических структур. В задачах синтаксиса всегда выделяется важное понятие *правильной* (или, как часто говорят, *отмеченной) структуры.* В силу этого основные задачи синтаксиса сводятся к нахождению способа удобного перечисления (порождения, распознавания) текстов с отмеченной структурой из данной знаковой системы. Аналогичные задачи могут возникать и при переходе к семантике: задание множества осмысленных текстов данного языка, задание множества фраз (текстов), имеющих тот же смысл, что и наперед указанная фраза, и т. п. Решая эти задачи с помощью готового аппарата порождающих грамматик, мы наталкиваемся на следующую принципиальную трудность: при решении синтаксических проблем часто можно с полным правом огрубить ситуацию, считая, что существует четкое разбиение всех текстов на множество отмеченных и дополнительное к нему множество неотмеченных. Однако в более тонких проблемах, в частности, в семантике, появляется «размытая» картина — наряду с текстами, безусловно осмысленными (семантически отмеченными), есть еще больше текстов, об осмысленности которых можно спорить. Причем, уменьшая от текста к тексту совсем немного степень осмысленности, мы за несколько шагов можем прийти к текстам, весьма далеким от правильно составленных. Точно так же, допуская перифразы с небольшим отклонением смысла, мы приходим за серию шагов к тексту, имеющему существенно иной смысл.

Аналогично, устанавливая смысловую близость слов или оборотов, интересно рассматривать не столько случаи полного тождества (одинаковости) смыслов (такие ситуации сравнительно бедны), сколько случаи сходства смыслов или, что то же самое, наличие достаточно большого множества общих значений.

Таким образом, при переходе к изучению семантики речь идет не просто о новой интерпретации синтаксических моделей (например, отмеченные тексты интерпретируются не как синтаксически правильные, а как осмысленные), а о новом классе «размытых» математических моделей.

Эти модели должны задавать не просто множества отмеченных текстов, а «облако» из таких множеств, так что при переходе от множества к множеству отмеченность «почти» сохраняется.

Суть дела состоит, конечно, не в том, чтобы перейти от точных синтаксических моделей к неточным семантическим. Это было бы просто отходом от принципов математической лингвистики. Речь идет о более трудной вещи: о переходе к точно задаваемым моделям, описывающим в точных терминах расплывчатость семантических явлений, без навязывания самим явлениям не присущей им излишней определенности и однозначности.

Для пояснения этого принципиального тезиса приведем аналогию из физики. **В классической механике <u>движения</u> характеризуются точно определяемыми <u>координатой и импульсом</u>. Эксперименты над микрочастицами показали, что координату и импульс нельзя одновременно задавать с произвольной точностью. Можно было бы на этой основе вообще отказаться от точного математического аппарата при описании динамики микрочастиц. Но квантовая механика пошла по иному пути: был создан точный аппарат, позволяющий говорить на точном языке о возникающих неопределенностях. Этот аппарат основан на принципиально новом способе описания состояний микромира: <u>вместо координаты и импульса вводится так называемая волновая функция, описывающая «размазанность» частицы в фазовом пространстве.</u> Заметим, что аппарат квантовой физики сам по себе не менее точно сформулирован, чем классический.**

Перейдем теперь к формальному описанию модели сочетаемости. Рассмотрим два множества M и L и соответствие φ между ними. Обозначим через $\mathfrak{M}$ график соответствия φ, т. е. множество пар $\langle x,\ \xi\rangle$, где $x \in M,\ \xi \in L,$ а x и ξ находятся в указанном соответствии.

На множестве пар $\mathfrak{M}$ мы будем считать заданным отношение «подобнозначности», обозначаемое через τ. Запись

$$\langle x,\ \xi\rangle\,\tau\,\langle y,\ \eta\rangle$$

читается: ξ имеет относительно x значение, подобное тому, что η имеет относительно y. Относительно τ будем предполагать, что оно симметрично и рефлексивно, т. е. является отношением толерантности. Мы будем обозначать через $\langle\mathfrak{M},\tau\rangle$ соответствующее пространство толерантности.

Рассмотрим следующий пример. Множество M состоит из основ русских существительных, а миожество L из падежных окончаний. Пару $\langle x,\ \xi\rangle$ мы включим в график соответствия, если основа x сочетаетсяс окончанием ξ, т. е. если в русском языке существует словоформа, полученная прибавлением к основе x окончания ξ. Грубо говоря, пара $\langle x,\ \xi\rangle$— это и есть словоформа, образованная из основы x с помощью окончания ξ. Отношение $\langle x,\ \xi\rangle\,\tau\,\langle y,\ \eta\rangle$ в данном случае, по определению, означает, что словоформы $\langle x,\ \xi\rangle$ и $\langle y,\ \eta\rangle$ могут выражать один и тот же падеж. Например,

$$\text{ран-а} \ \tau \ \text{стол}$$

и

$$\text{стол} \ \tau \ \text{книг-у},$$

поскольку первая пара словоформ может выражать именительный падеж, а вторая — винительный.

Однако словоформы «ран-а» и «книг-у» не могут выражать общего падежа; таким образом, отношение x в данном случае не транзитивно.

Ясно, что этот пример можно развить для других типов основ и для других интерпретаций отношения τ (совпадения рода, числа и падежа или времени, лица и числа, или еще каких-либо комбинаций грамматических признаков).

Во всяком случае, как показывает анализ ранее приведенного примера, отношение подобиозначности τ, вообще говоря, не является транзитивным.

Вопрос о том, между какими парами фактически имеет место отношение подобнозначности, стоит вне рамок математической модели и решается информантами по соглашению.

Следующий пример состоит в том, что для слов «голос», «ветер», «игла», «течение», образующих множество *M,* можно образовывать пары путем приписывания эпитетов из множества L = {большой, громкий, сильный, острая, бурное}. В русском языке заведомо допустимо образование осмысленных пар: бурное течение, сильный голос, острая игла, большой ветер, но сомнительны выражения вроде: бурная игла, громкий ветер. Возможны различные точки зрения на то, какие из этих пар имеют подобные (сходные) значения. Можно считать, что все пары выражают значение усиления и поэтому равнозначны. Можно считать пары типа «острая игла» и «большая игла» или «большое течение» и «борное течение» неподобнозначными.

Можно было бы пойти с самого начала по другому пути: выделить заранее некоторые (смысловые) признаки и объявить подобнозначными те пары, в которых эти признаки можно найти. Тогда отношение x автоматически оказалось бы транзитивным, поскольку любое отношение, определяемое как совпадение некоторой фиксированной группы признаков (попадание в общий класс), транзитивно.

Мы же принимаем противоположную точку зрения: сначала определяется подобнозначность конкретных пар (в пределах точности, принимаемой информантом), а лишь затем выясняется, можно ли подобнозначные пары («синонимы») классифицировать на группы.

Определение 12.4. Назовем *предсемьей* пару вида $\langle\varphi, \tau\rangle$, где $\varphi = \langle\mathfrak{M}, M, L\rangle$ — соответствие, а τ — отношение толерантности на $\mathfrak{M}$.

Понятие предсемьи определяет важный тип структуры, который можно наглядно изобразить следующим образом. Сопоставим элементам множеств M и L вершины некоторого графа $\mathfrak{M}$. Элемент $x \in M$ будем соединять с элементом $\xi \in L$ ребром, если x и ξ находятся в соответствии φ, т. е. если $\langle x, \xi\rangle \in \mathfrak{M}$. На множестве ребер $\mathfrak{M}$ задается толерантность τ.

Например, имеется множество клиентов M и множество обслуживающих мастеров L. Некоторые мастера обслуживают некоторых клиентов. Про некоторые пары таких обслуживании утверждается, что они сходны. В частности, может оказаться, что эти обслуживания можно разбить на непересекающиеся классы сходных: починка обуви, химчистка, ремонт часов и т. п. Это соответствует случаю транзитивности отношения τ.

На рис. 12.7 толерантные ребра помечены одинаковым способом.

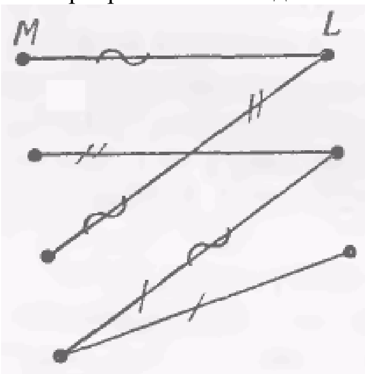


Рис. 12.7. Предсемья.

В этом примере отношение τ не транзитивно. В транзитивном случае ребра каждого типа можно выкрасить в особый цвет.

Определение 12.5. Предсемья $\langle\varphi, \tau\rangle$ называется *связной,* если

а) соответствие φ всюду определено;

б) соответствие φ сюръективно;

в) множество M непусто.

Очевидно, в случае связной предсемьи множество L также непусто и соответствующий граф не имеет изолированных вершин.

Иными словами, в связной предсемье всякий клиент обслуживается хотя бы одним мастером и, наоборот, каждый мастер обслуживает хотя бы одного клиента.

Определение 12.6. Связная предсемья $\langle\varphi, \tau\rangle$ называется *семьей,* если

а) для любых $x \in M$ и $y \in M$ и любого $\xi \in L$ такого, что $\langle x, \xi\rangle \in \mathfrak{M}$, существует $\eta \in L$ такое, что $\langle y, \eta\rangle \in \mathfrak{M}$ и $\langle x, \xi\rangle \tau \langle y, \eta\rangle$.

б) для любых $\xi \in L$ и $\eta \in L$ и любого $x \in M$ такого, что $\langle x, \xi\rangle \in \mathfrak{M}$, можно найти $y \in M$ такое, что $\langle y, \eta\rangle \in \mathfrak{M}$ и $\langle x, \xi\rangle \tau \langle y, \eta\rangle$.

Свойство а) можно назвать *полнотой:* если в семье может быть выражен некоторый смысл относительно слова *х,* то тот же смысл можно выразить и относительно любого другого слова *у.*

Свойство б) можно назвать *однородностью:* если ξ выражает некоторый смысл относительно слова *х,* то любой другой элемент $\eta \in L$ для каких-то слов выражает тот же смысл.

Иначе говоря, все типы обслуживании, которые имеет один клиент, имеют и все остальные. И все виды обслуживания, которые выполняет один мастер, выполняет и любой другой, хотя, возможно, для других клиентов.

Определение 12.7. Семья $\langle\varphi, \tau\rangle$ называется *примитивной,* если τ — полное отношение.

Может быть полезным исследование ситуаций, когда описание семьи сводится к заданию одной или нескольких примитивных. Такое сведение возможно для случая транзитивности отношения τ (теорема 12.3).

Теорема 12.1. *Если в семье* $\langle\varphi, \tau\rangle$ *отношение* τ *транзитивно и существует элемент* $\xi \in L$ *такой, что для любых* $x \in M$ *и* $y \in M$ *из* $\langle x, \xi\rangle \in \mathfrak{M}$ *и* $\langle y, \xi\rangle \in \mathfrak{M}$ *вытекает* $\langle x, \xi\rangle \tau \langle y, \xi\rangle$, *то семья* $\langle\varphi, \tau\rangle$ *примитивна.*

Доказательство. Покажем, что для любых $\langle x, \eta\rangle \in \mathfrak{M}$ и $\langle y, \zeta\rangle \in \mathfrak{M}$ выполнено $\langle x, \eta\rangle \tau \langle y, \zeta\rangle$. По определению 12.6 найдутся такие $z \in M$ и $u \in M$, что $\langle x, \eta\rangle \tau \langle z, \xi\rangle$ и $\langle y, \zeta\rangle \tau \langle u, \xi\rangle$. Так как по условию $\langle z, \xi\rangle \tau \langle u, \xi\rangle$ и отношение τ симметрично и транзивно, получаем $\langle x, \eta\rangle \tau \langle y, \zeta\rangle$.

Эта теорема допускает следующую наглядную интерпретацию: если есть один элемент $\xi \in L$, выражающий для всех элементов из *М* общий смысл, и *х* транзитнвно, то все пары выражают один и тот же смысл.

Тот же вывод верен, если семью можно пополнить таким оператором ξ. Например, если к средствам естественного языка добавить формальные выражения операторов типа Мельчука—Жолковского и включение такого оператора в семью позволяет выразить тот же смысл для любого слова, то (разумеется, если отношение τ было транзнтивпо) исходная семья автоматически оказывается примитивной.

Т е о р е м а 12.2. *Если в семье* $\langle\varphi, \tau\rangle$ *множество L одноэлементное, то семья* $\langle\varphi, \tau\rangle$ *примитивна.*

Действительно, пусть $L = \{\xi\}$ и $x \in M,\ y \in M$. По определению 12.5 $\langle x, \xi\rangle \in \mathfrak{M}$ и $\langle y, \xi\rangle \in \mathfrak{M}$. Из $\langle x, \xi\rangle \in \mathfrak{M}$ по определению 12.6 вытекает существование такого $\eta \in L$, что $\langle x, \xi\rangle\tau\langle y, \eta\rangle$. Поскольку по условию теоремы $\xi = \eta$, $\langle x, \xi\rangle\tau\langle y, \xi\rangle$, ч. т. д.

Здесь мы доказали примитивность семьи $\langle\varphi, \tau\rangle$, не мспользуя транзитивность отношения τ.

Для транзитивного отношения τ любая семья может быть представлена как несложная композиция примитивных. Разберем этот случай немного подробней. В этом случае множество пар (ребер графа) разбивается на непересекающиеся классы эквивалентности.

Определение 12.8. Семья $\Sigma_1 = \langle\varphi_1, \tau_1\rangle = \langle\langle\mathfrak{M}_1, M_1, L_1\rangle, \tau_1\rangle$ называется *простым сужением* семьи $\Sigma = \langle\varphi, \tau\rangle = \langle\langle\mathfrak{M}, M, L\rangle, \tau\rangle$, если $M_1 \subseteq M,\ L_1 \subseteq L,\ \mathfrak{M}_1 \subseteq \mathfrak{M}$ и $\tau_1 \subseteq \tau$.

Пусть в семье $\Sigma = \langle\varphi, \tau\rangle$ отношение τ трапзитнвно и *K* — некоторый класс эквивалентности для этого отношения. Тогда можно рассмотреть простое сужение $\Sigma_K = \langle\langle K, M, L\rangle, \tau_K\rangle$ семьи $\Sigma = \langle\langle\mathfrak{M}, M, L\rangle, \tau\rangle$, получаемое, если оставить только пары, входящие в класс *K*. Легко проверить, что Σ_K действительно является семьей. В самом деле, существует хотя бы одна пара $\langle x, \xi\rangle$, принадлежащая классу *K*. Но тогда по определению семьи для всякого $y \in M$ существует $\eta \in L$ такое, что $\langle x, \xi\rangle\tau\langle y, \eta\rangle$. Следовательно, $\langle y, \eta\rangle \in K$.

Аналогично, для любого $\eta \in L$ существует $y \in M$ такое, что $\langle x, \xi\rangle\tau\langle y, \eta\rangle$. Таким образом, сужение Σ_K является семьей, причем, очевидно, примитивной семьей. Беря все классы эквивалентности, мы приходим к результату, формулируемому как

Теорема 12.3. *Пусть* S *— семья с транзитивным отношением* τ. *Тогда существует множество примитивных семей* $\Sigma_{K_1}, \Sigma_{K_2}, \ldots$ *таких, что*

1) *каждая семья* Σ_{K_i} *есть простое сужение семьи* ∑;

2) *для каждой пары* $\langle x, \xi\rangle \in \mathfrak{M}$ *существует ровно одно* K_i, *для которого* $\langle x, \xi\rangle \in K_i$;

3) *если* $\langle x, \xi \rangle \tau \langle y, \eta \rangle$, *то пары* $\langle x, \xi \rangle$ *и* $\langle y, \eta \rangle$ *принадлежат общему* K_i.

Теорема 12.3 дает в сущности перечисление всех возможпых семей с транзитивным отношением τ. Геометрически такие семьи строятся так: берутся множества *M* и *L* и строятся *m* графов. В каждом из графов любая вершина из *M* соединена с некоторой вершиной из *L* и любая вершина из *L* соединена с некоторой вершиной из *M.* В каждом графе ребра окрашены в свой цвет. Наконец, каждая пара $\langle x, \xi \rangle$ может соединяться только в одном из графов, т. е. каждая пара может соединяться ребром только одного цвета. Теперь берется объединение всех ребер и полагается $\langle x, \xi \rangle \tau \langle y, \eta \rangle$, если соответствующие ребра одинаково окрашены. Эта конструкция и дает общий вид транзитивной семьи. Ее одноцветные части являются составляющими примитивными семьями.

В нпетранзитнвном случае роль примитивных семей играют неразложимые семьи. Именно, семья $\langle \varphi, \tau \rangle$ называется *неразложимой,* если транзитивное замыкание $\bar{\tau}$ отношения τ является полным отношением. Для произвольной семьи имеет место точный аналог теоремы 12.3 с заменой термина «примитивная» на «неразложимая». Таким образом, все сводится к алгебраической проблеме описания всех неразложимых семей.

12.4. Формальная задача теории дешифровки

В области дешифровки неизвестных письменностей и языков (и в других родственных лингвистических проблемах) явным образом возникают задачи об установлении изоморфных соответствий между множествами с отношениями.

Ранее было определено, что такое изоморфизм двух множеств, на которых задано по отношению. Пусть теперь имеется два множества, на каждом из которых определено по *n* отношений: $\langle M^1, A_1^1, A_2^1, \ldots, A_n^1 \rangle$ и $\langle M^2, A_1^2, A_2^2, \ldots, A_n^2 \rangle$. Мы говорим, что эти два множества с отношениями *изоморфны,* если существуют такие взаимно-однозначное соответствие ψ между множествами M^1, M^2 и взаимно-однозначное соответствие θ между множествами $\{A_1^1, A_2^1, \ldots, A_n^1\}$ и $\{A_1^2, A_2^2, \ldots, A_n^2\}$, что между соответствующими друг другу элементами множеств выполнены соответствующие друг другу отношения. Именно, если x^1 и y^1

принадлежат M^1 и выполнено соотношение $x^1 A_i^1 y^1$, то для их образов $x^2 = \psi(x^1)$ и $y^2 = \psi(y^1)$ должно выполняться соотношение $x^2 A_j^2 y^2$, где $A_j^2 = \theta(A_i^1)$. Обратно, из $x^2 A_j^2 y^2$ должно следовать $x^1 A_i y^1$.

В задачах дешифровки часто возникает задача поиска соответствия (перевода) между двумя множествами (слов или других элементов языка) и между отношениями на этих множествах так, чтобы это соответствие устанавливало изоморфизм множеств с отношениями. В качестве примера мы приведем искусственно придуманную задачу, которая давалась на Второй традиционной олимпиаде по языковедению и математике на филологическом факультете Московского государственного университета.

Дан перечень из следующих десяти арабских слов, записанных в латинской транскрипции (знак ᶜ означает специфический согласный арабского языка):

miγzal, maᶜbud, mahzan, maᶜmil, mirgab, maᶜbar, maγzul, maᶜbad, miᶜbar, maᶜmal.

Это множество мы обозначим через $M_{ар}$. Множество $M_{рус}$ русских слов состоит из переводов этих слов на русский язык: кумир, рабочий, речная переправа, склад, пряжа, паром, завод, веретено, святилище (место поклонения), телескоп. Требуется для каждого из арабских слов определить его русский перевод. Иными словами, требуется (не обращаясь к словарям и лицам, знающим оба языка) найти правильное соответствие:

$$\psi: M_{ар} \to M_{рус}.$$

Задача на первый взгляд не может иметь однозначного решения. Любое из взаимно-однозначных отображений $M_{ар} \to M_{рус}$ в равной степени может быть формальным ответом. Общее число возможных отображений равно числу перестановок из 10 элементов, т. е. 10! =3 628 800. Оказывается, простой факт, что мы имеем множества из осмысленных слов, сокращает степень неопределенности задачи в более чем три с половиной миллиона раз и позволяет получить однозначное решение задачи с высокой степенью надежности. Дело в том, что в нашем множестве русских слов отчетливо выделяются некоторые смысловые отношения. Это отношения R_1 —«относиться к общему семантическому полю» и R_2 — «выражать общий семантический класс». Оба эти отношения суть эквивалентности и определяют разбиения множества $M_{рус}$ на классы. Классы по R_1 имеют вид:

{веретено, пряжа}, {телескоп}, {паром, переправа}, {кумир, святилище}, {склад}, {завод, рабочий}.

Классы по R_2 имеют вид: {веретено, телескоп, паром} — инструмент, которым производится действие, {пряжа, кумир} — объект, над которым производится действие, {переправа, святилище, склад, завод} — место действия, {рабочий} — субъект действия. Но между соответствующими арабскими словами выполнены те же самые смысловые отношения. ***Правдоподобно предположить, что эти отношения как-то выражены во внешней форме слов.*** Посмотрим теперь, какие формальные отношения существуют между арабскими словами из $M_{\text{ар}}$. На множестве $M_{\text{ар}}$ легко выделяются два отношения: Q_1 — «иметь одинаковую структуру согласных» и Q_2 — «иметь одинаковую структуру гласных». Оба эти отношения суть отношения эквивалентности.

(Разумеется, эти отношения легче выделить, если заранее знать, что в семитских языках (к этому классу языков относятся арабский, иврит, эфиопский, аккадский и многие другие живые и мертвые языки народов передней Азии и северо-восточной Африки) последовательности согласных и гласных имеют смыслоразличнтельпый характер.)

По отношению Q_2 мы имеем следующие классы:

{miγzal, mayzul}, {mirgab}, {miʿbar, maʿbar}, {maʿbud, maʿbad}, {mahzan}, {maʿmal, maʿmil}.

По отношению Q_1 получаем классы {miγzal, mirgab, maʿmal}, {maʿmil}.

Сравнивая числа элементов в классах разбиений множеств $M_{\text{ар}}$ и $M_{\text{рус}}$, мы видим, что отождествить следует отношения R_1 и Q_1, а также R_2 и Q_2. Теперь надо установить такое соответствие между арабскими и русскими словами, чтобы словам, входящим в общий класс по Q_1, соответствовали слова, входящие в общий класс по R_1. Аналогично, если арабские слова имеют общую структуру гласных (находятся в отношении Q_2), то их русские переводы должны выражать общий семантический класс (находиться в отношении R_2). Соберем арабские и русские слова в таблицы, где столбцы и строки обязаны соответствовать друг другу в смысле совпадения числа элементов в соответствующих классах:

Гласные / Согласные	*ia*	*au*	*aa*	*ai*
mγzl	*miγzal*	*maγzul*		
mrgb	*mirgab*		(*margab*)	(*margib*)
$m^{c}br$	$mi^{c}bar$		$ma^{c}bar$	
$m^{c}bd$		$ma^{c}bud$	$ma^{c}bad$	($ma^{c}bid$)
mhzn			*mahzan*	
$m^{c}ml$			$ma^{c}mal$	$ma^{c}mil$

Класс / Поле	инструмент	обьект	место	субъект
Прядение	Веретено	Пряжа		
Астрономия	Телескоп		(Обсерватория)	(Астроном)
Перевоз	Паром		Переправа	
Культ		Кумир	Святилище	(Жрец)
Хранение			Склад	
Производство			Завод	

Из этих таблиц видно, что изоморфное соответствие (перевод с сохранением указанных отношений) возможно только при выбранном соответствии строк (классов по R_1 и Q_1) и столбцов (классов по R_2 и Q_2).

Более того, в скобках мы указали русские и арабские слова, на которые наша таблица дает основание экстраполировать полученное соответствие. Эта процедура напоминает заполнение Менделеевым пустых мест в открытой им таблице химических элементов. Заметьте, что таблицу Менделеева тоже можно интерпретировать как установление соответствия между классами элементов с данными химическими свойствами и классами элементов с данными типами атомных весов и номеров.

В реальных задачах дешифровки трудность заключается в том, что чистого изоморфизма никогда не бывает, а нужно искать простые множества слов (слогов, букв) и их соответствия, для которых изоморфизм выполняется.

12.5. О дистрибуциях

В структурной лингвистике широко используется понятие так называемой *дистрибуции,* или *дистрибутивного отношении.* Это понятие применимо к любым элементам, образующим тексты: словам, слотам. морфемам, буквам, звукам и т. д. Мы дадим здесь основные определения, связанные с этим понятием.

Пусть задан некоторый язык *Я,* т. е. некоторое множество текстов, принадлежащих фиксированной знаковой системе. Таким образом, мы здесь понимаем **язык как запас текстов определенного вида**.

Определим теперь *операцию замены* $(x;\ a \to b)$ (Здесь *a* и *b* — элементы алфавита $\mathfrak{A}$, а *x* — элемент несущего множества *M* для текста *T*).

Пусть имеется некоторый текст $T = \langle S,\ \varphi\rangle$, где $S = \langle M,\ A_1,\ A_2,\ \ldots,\ A_n\rangle$. *Результатом замены* $(x;\ a \to b)$ мы будем называть текст $T' = \langle S,\ \varphi'\rangle$, где $\varphi'(y) = \varphi(y)$ для всех элементов *y* несущего множества *M,* отличных от *x,* и

$$\varphi'(x) = \begin{cases} \varphi(x), & \text{если} \quad \varphi(x) \neq a, \\ b, & \text{если} \quad \varphi(x) = a. \end{cases}$$

В случае $\varphi(x) \neq a$ результирующий текст *T'* совпадает с исходным. В этом случае замену мы будем называть *фиктивной.* Иначе говоря, замена $(x;\ a \to b)$ состоит в том, что текст в фиксированном месте *x* меняется: если в этом месте текста *T* стоял знак *a*, то в *T'* на этом же месте будет стоять *b*.

Например, пусть синтаксическая схема есть множество {1, 2, 3, 4} с совершенным порядком, а текст *T* есть цепочка *abca*. Тогда замены

$$(1;\ a \to b),\ (2;\ a \to b),\ (3;\ a \to b),\ (4;\ a \to b)$$

дают, соответственно, следующие цепочки:

$$bbca,\ abca,\ abca,\ abcb.$$

Текст *T',* полученный в результате данной замены $(x;\ a \to b)$ из текста *T,* вообще говоря, не обязан принадлежать тому же самому языку. Поэтому, если дан алфавит $\mathfrak{A}$ и зафиксирован язык, то возможность выполнять замены некоторых знаков алфавита, не выходя за пределы этого языка, определяет *дистрибутивные отношения* на алфавите $\mathfrak{A}$, т. е. отношения, связанные со свойствами распределения знаков алфавита в тексте. Введем соответствующие определения, считая каждый раз, что язык *Я* уже зафиксирован.

Определение 12.9. Элемент *a* *мажорирует* элемент *b*, если для всякого текста $T \in Я$ результат замены вида $(x;\ a \to b)$ при любом *x* есть текст *T'*, принадлежащий языку Я.

Это отношение мы будем обозначать через $\Rightarrow$. Легко проверить, что оно рефлексивно и транзитивно, т. е. является квазипорядком. Отношение

$$\Leftrightarrow = \Rightarrow \cap (\Rightarrow)^{-1}$$

является отношением эквивалентности (теорема 7.6) и означает *взаимозаменимость.* Именно, $a \Leftrightarrow b$ означает, что текст *T* и результат любой замены $(x;\ a \to b)$ одновременно принадлежат или не принадлежат языку *Я*. Действительно, соотношение $a \Leftrightarrow b$ означает, что одновременно выполнено $a \Rightarrow b$ и $b \Rightarrow a$. Таким образом, если $T \in Я$, то и результат замены $(x;\ a \to b)$ $T' \in Я$. Если же $T' \in Я$, то результат обратной замены $(x;\ b \to a)$, совпадающий с исходным текстом *T,* также принадлежит *Я*.

Например, в языке, состоящем из цепочек *abb. bbb, aba* и *bba,* над алфавитом из двух элементов *a* и *b* выполнено соотношение $a \Rightarrow b$, но не выполнено $b \Rightarrow a$; таким образом, отношение $\Rightarrow$ является здесь настоящим порядком.

Определение 12.10. Элементы *a* и *b* находятся в *отношении общей дистрибуции,* если существуют такой текст $T = \langle S,\ \varphi \rangle \in Я$ и такая замена $(x;\ a \to b)$, что результат замены *T'* принадлежит языку *Я*, причем φ*(x) = a.*

Последнее означает, что замена $(x;\ a \to b)$ не фиктивна, но в позиции *x* действительно присутствует элемент алфавита *a,* заменяемый на *b*. Отношение общей дистрибуции симметрично, так как, если существует нефиктивная замена $(x;\ a \to b)$ в тексте *T,* то в результирующем тексте *T'* существует нефиктивная замена $(x;\ b \to a)$. Отношение общей дистрибуции, вообще говоря, не рефлексивно. Для рефлексивности этого отношения необходимо и достаточно, чтобы в алфавите $\mathfrak{A}$ не было «безработных» элементов, т. е. чтобы для всякого элемента $a \in \mathfrak{A}$ существовал текст- $T = \langle S,\ \varphi \rangle \in Я$ и позиция *x,* для которой $\varphi(x) = a$. Тогда в этом тексте возможна замена $(x;\ a \to a)$. Таким образом, отношение общей дистрибуции является (при разумных ограничениях на язык) толерантностью.

Мы могли бы ввести другой вариант этого отношения, потребовав в определении 12.10, чтобы *T* не совпадал с *T'*. Условие φ*(x)* = *a* было бы этим обеспечено, но никакой элемент *a* не был бы тогда в отношении общей дистрибуции сам с собой. Такое отношение было бы симметричным и антирефлексивным.

Наконец, важный тип дистрибутивных отношений даст

Определение 12.11. Элементы *a* и *b* находятся в *отношении дополнительной дистрибуции,* если для всякого текста $T = \langle S,\ \varphi \rangle \in Я$ результат *T'* любой замены $(x;\ a \to b)$ при φ*(x)* = *a* не принадлежит языку *Я*.

Соотношение дополнительной дистрибуции между элементами алфавита *a* и *b* мы будем обозначать через *a* Com *b*. Отношение дополнительной дистрибуции очевидным образом антирефлексивно (если в алфавите $\mathfrak{A}$ нет «безработных» элементов.)

Докажем, что отношение дополнительной дистрибуции симметрично. Пусть выполнено *a* Com *b,* но неверно, что *b* Com *a.* Тогда существует текст $T \in Я$ и нефиктивная замена $(x;\ b \to a)$, дающая текст $T' \in Я.$

Тогда в тексте *T'* можно проделать нефиктивную замену $(x;\ a \to b)$, т. е. соотношение *a* Com *b* не выполнено. Полученное противоречие доказывает симметричность дополнительной дистрибуции. По теореме 6.1 на алфавите $\mathfrak{A}$ можно ввести такую систему признаков, что соотношение *a* Com *b* будет выполнено в том и только том случае, когда *a* и *b* имеют ровно один общий признак.

В качестве примера рассмотрим язык *Я*, состоящий из цепочек $abbc,\ bbbc,\ baba,\ abbb,\ abbd$ и $bbbd.$ В этом языке *a* и *b* находятся в отношений общей дистрибуции, *b* и *c* также находятся в отношении общей дистрибуции, но *a* и *c* находятся в дополнительной дистрибуции. Элементы *c* и *d* взаимозаменимы: $c \Leftrightarrow d.$ Элемент *d* находится с элементами *a* и *b* в тех же дистрибутивных отношениях, что и элемент *c*

Нетрудно доказать, что справедлива следующая

Лемма 12.1. *Если элементы a и b взаимозаменимы, а элемент c находится в одном из двух дистрибутивных отношений с a, то он находится в том же отношении с b.*

Так как отношение взаимозаменимости есть эквивалентность, то можно ввести разбиение алфавита $\mathfrak{A}$ на классы эквивалентности по этому отношению. Эти классы называются *дистрибутивными классами.*

В качестве примера возьмем множество синтаксически правильных русских фраз. Это множество можно рассматривать как язык $Я_г$ над алфавитом $\mathfrak{A}$, состоящим из всех русских словоформ. Этот язык, вообще говоря, нельзя отождествлять с русским языком в классическом понимании. Так, в язык $Я_г$ нам пришлось бы включить тексты типа «Огненный стул задумчиво преобразовывал сапоги».

С другой стороны, в $Я_г$ не войдет, скорее всего, фраза «Я для нее все равно», хотя этот пример взят пз Лермонтова («Княгиня Литовская»).

В языке $Я_г$ дистрибутивные классы состоят из словоформ, имеющих тождественную грамматическую структуру — совпадающий набор грамматических признаков. (Мы не уточняем здесь полного списка признаков: в зависимости от того, что мы принимаем в качестве признаков, исчерпывающих грамматическую характеристику словоформы, мы можем получить разные языки $Я_г$.*)* Дистрибутивные классы будут состоять из множеств словоформ вида {стул, стол, столб, катер, .. .} или {зеленого, большого, прекрасного, .. .}. Эти множества состоят из грамматически одинаковых форм разных слов, поскольку, заменяя слово в некоторой форме на другое слово в той же форме, мы будем получать грамматически правильную фразу. Так, приведенный выше пример имеет вполне осмысленный прототип:

«Огненный столб медленно окутывал дома».

Если мы возьмем две словоформы существительных, стоящие в разных падежах, то они будут находиться, вообще говоря, в дополнительном распределении, поскольку, изменив в правильной русской фразе падеж существительного, мы получаем обычно неправильную фразу.

Формально, существуют примеры текстов, где различные падежи существительных могут заменять друг друга. Скажем, одинаково возможны фразы «Врач осматривает глаза» и «Врач осматривает глазами». Но по существу здесь взаимозаменимость происходит от незаполненности всех мест. В «полных» фразах, где явно указаны «мнимые» члены предложения, взаимозаменимости уже не будет: «Врач осматривает чем-то глаза» и «Врач осматривает глазами что-то». Аналогично, прилагательные, не совпадающие по роду, числу и падежу, также находятся в дополнительном распределении.

Представим теперь, что в некотором языке *Я* в каждом дистрибутивном классе выбран эталонный элемент. Тогда в любом тексте языка любой элемент алфавита может быть заменен эталонным элементом из того же класса. В силу взаимозаменимости любых элементов одного класса мы получим опять текст из языка *Я*. Запас таких текстов мы будем называть *эталонным.* Обратно, любой текст языка *Я* можно получить из эталонного с помощью ряда замен вида $(x;\ a \to b)$, где a и b лежат в одном классе.

Итак, вместо того, чтобы указывать все множество текстов языка, достаточно задать эталонный запас и правила замены, т. е. дистрибутивные классы. Тем самым более экономно кодируется информация об языке. В сущности изучение грамматики чужого языка обычно идет путем задания эталонных текстов и дистрибутивных классов (перечня типов склонений и спряжений). Разумеется, так можно изучить лишь грамматическую структуру языка, а не фразеологические обороты, смысловую сочетаемость и оттенки значений.

М. В. Арапов заметил, что школьное обучение грамматике родного языка путем так называемой *постановки вопросов* есть в сущности обучение подстановке вопросительного местоимения вместо словоформы. Например, переход от фразы «Маша кушает кашу» к вопросительным фразам «Кто кушает кашу?» и «Маша кушает что?» есть переход к эталонным местоимениям, для которых ученик заранее запоминает название падежа. Заметим, что можно изучать дистрибутивные отношения в фиксированной позиции текста, зафиксировав во всех предыдущих определениях значение x в заменах $(x;\ a \to b)$.

Следующие рассуждения годятся лишь для случая, когда язык *Я* есть множество конечных цепочек (синтаксическая схема есть конечное множество с одним отношением — совершенным порядком).

Будем далее употреблять знак включения $A \subseteq B$ для того, чтобы указать тот факт, что цепочка *А* является *связной частью* цепочки *В,* т. е. что знаки алфавита, стоящие в серии последовательных позиций цепочки *В,* образуют цепочку *А* ($A \subseteq B$ означает существование таких цепочек С, *D,* что цепочка *В* получается последовательным приписыванием к цепочке *С* сначала *А,* потом *D(B = CAD).*).

Например, если $B = abcabb$ и $A = cabb$, то $A \subseteq B$.

Пусть *А, В* и *А'* — три цепочки и $A \subseteq B$. *Результатом замены* $A \to A'$ называется цепочка *В',* полученная из *В* зачеркиванием цепочки *А* и вписыванием на ее место цепочки *А'* (Поскольку цепочка *А* может входить в цепочку *В* несколько раз, результат замены не однозначно определяется цепочками *А, В, А'.*). При такой замене длина цепочки может измениться.

Определение 12.12. Цепочки *А* и *А'* называются *взаимозаменимыми,* если для любой замены $A \to A'$ в любой цепочке *В* такой, что $A \subseteq B$, результат замены принадлежит *Я* в том и только том случае, когда $B \in Я$, и обратно: для любой замены $A' \to A$ в любой цепочке *B* такой, что $A' \subseteq B$, результат замены принадлежит *Я* в том и только том случае, когда $B \in Я$.

Лемма 12.2. *Отношение взаимозаменимости цепочек является эквивалентностью.*

Доказательство предоставляем читателю.

Обозначим отношение взаимозаменимости цепочек тем же символом $\Leftrightarrow$, что и аналогичное отношение для элементов алфавита.

Множество всех «безработных» цепочек, т. е. цепочек, не входящих ни в одну цепочку языка *Я*, образует класс эквивалентности по отношению взаимозаменимости, который мы обозначим $K_{бр}$.

Пример 1. Язык $Я_1$ состоит из всех цепочек вида $a^m b^n$, т. е. $\underbrace{aa\ \dots\ a}_{m \text{ раз}}\ \underbrace{bb\ \dots\ b}_{n \text{ раз}}$ $(m \geqslant 0,\ n \geqslant 0, m + n > 0)$. В данном случае имеются четыре следующих класса взаимозаменимых цепочек:

$$K_{бр},\ K_a = \{a, aa, \dots\},\ K_b = \{b, bb, \dots\},\ K_{ab} = \{a^m b^n\},$$

где $m > 0$ и $n > 0$.

Пример 2. Язык *Я* состоит из всех цепочек вида $a^m b^m$. В этом случае количество классов оказывается бесконечным: $K_{бр},\ K_n = \{a^p b^q\}$, где $q - p = n$ $(n = 0,\ \pm 1,\ \pm 2,\ \dots)$, а также одноэлементные классы $K_a^i = \{a^j\}$ и $K_b^i = \{b^i\}$ $(i,\ j = 1,\ 2,\ 3\ \dots)$.

Аккуратное вычисление классов в этих примерах предоставляем читателю.

Обозначим через $\mathfrak{G}$ множество всех конечных цепочек над алфавитом $\mathfrak{A}$. Нетрудно видеть, что если $\mathfrak{G}$ рассматривать как язык, то в нем есть ровно один класс взаимозамснимых цепочек. Множество $\mathfrak{G}$ является полугруппой относительно операции приписывания цепочек (так называемой *конкатенации).*

Результат приписывания цепочки *B* справа к цепочке *A* мы будем обозначать через *A B*. Например, если *A* = *aba* и *B* = *aab*, то *AB* = *abaaab* и *BA* = *aababa*.

Лемма 12.3. *Пусть* $A \Leftrightarrow A'$ *и* $B \Leftrightarrow B'$. *Тогда*

$$AB \Leftrightarrow A'B'.$$

Доказательство предоставляем читателю.

Итак, результат операции приписывания разных представителей классов лежит всегда в одном и том же классе. Это означает, что можно определить операцию приписывания самих классов взаимозаменимости. Именно, пусть даны два класса K_1 и K_2. Выберем в них по представителю $A \in K_1$' и $B \in K_2$. Тогда через $K_1 K_2$ мы обозначим класс, в котором находится цепочка *AB*. В силу леммы 12.3

определение класса K_1K_2не зависит от выбора цепочек в классах K_1 и K_2. (Заметим, что класс K_1K_2может быть шире, чем множество всех цепочек AB при $A \in K_1,\ B \in K_2$.)

В первом примере мы имеем следующие правила «перемножения» классов:

$$K_{бр}K_a = K_{бр}K_b = K_{бр}K_{ab} = K_{бр}K_{бр} =$$
$$= K_aK_{бр} = K_bK_{бр} = K_{ab}K_{бр} = K_{бр};$$
$$K_aK_b = K_{ab};\quad K_bK_a = K_{бр};\quad K_bK_b = K_b;\quad K_aK_a = K_a$$
$$K_{ab}K_b = K_aK_{ab} = K_{ab};$$
$$K_bK_{ab} = K_{ab}K_a = K_{ab}K_{ab} = K_{бр}.$$

Во втором примере «таблица умножения» имеет вид

	$K_{бр}$	K_n	K_a^m	K_b^m
$K_{бр}$	$K_{бр}$	$K_{бр}$	$K_{бр}$	$K_{бр}$
K_r	$K_{бр}$	$K_{бр}$	$K_{бр}$	K_{r+m}
K_a^i	$K_{бр}$	K_{n-i}	K_a^{i+m}	K_{m-i}
K_b^i	$K_{бр}$	$K_{бр}$	$K_{бр}$	K_b^{i+m}

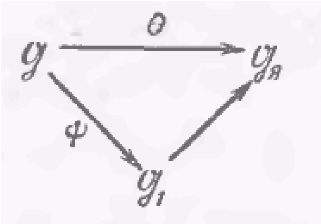
Обозначим через $\mathfrak{S}_Я$полугруппу классов, определяемую языком *Я*. Очевидно, отображение

$$\theta\colon \mathfrak{S} \to \mathfrak{S}_Я,$$

которое каждой цепочке ставит в соответствие ее класс, является гомоморфизмом свободной полугруппы $\mathfrak{S}$ на полутруппу $\mathfrak{S}_Я$ классов взгимозаменимости относительно языка *Я*.

Известно, что множество классов взаимозаменимости конечно в том и только том случае, когда язык *Я* принадлежит к типу так называемых *автоматных языков.* Интересно выяснить, какие условия на полутруппу $\mathfrak{S}_Я$ следуют из условия, что язык *Я* описывается некоторым типом порождающих грамматик. Е. Путинский в дипломной работе описал класс таких полутрупп, для которых существует язык *Я*, при котором исходная полутруппа изоморфна $\mathfrak{S}_Я$.

Рассмотрим, наконец, некоторый гомоморфизм $\psi\colon \mathfrak{G} \to \mathfrak{G}_1$ свободной полугруппы $\mathfrak{G}$ в какую-то полугруппу $\mathfrak{G}_1$. Будем называть этот гомоморфизм *нормальным* относительно языка *Я*, если из того, что $A \in Я$ и $\psi(B) = \psi(A)$, следует, что $B \in Я$.
Иначе говоря, если *A* и *B* имеют общий образ, то они одновременно принадлежат или не принадлежат языку *Я*. Оказывается, любой нормальный гомоморфизм продолжаем до гомоморфизма θ в полугруппу классов:



Нетрудно видеть, что, наоборот, всякий гомоморфизм, для которого нарисованная диаграмма коммутативна, нормален относительно языка *Я*. В этом построении интересно, как произвольному языку *Я* сопоставляются алгебраические объекты: полугруппа классов и нормальные гомоморфизмы свободной полугруппы. Интересно было бы исследовать, как связаны алгебраические свойства этих объектов и свойства языка. Например, что означает изоморфизм полугрупп $\mathfrak{G}_Я$?

12.6. Индивидуальные тестовые задачи

Сводка основных типов отношений и их свойств

В нижеследующей таблице мы для сравнения приводим список основных типов отношений и определяющих их свойств. Знаком + мы обозначаем, что данное свойство входит в определение данного типа отношения. Знак (+) показывает, что данное свойство отношения вытекает из остальных.

Тип отношения	Рефлексивность	Симметричность	Транзитивность	Антирефлексивность	Асимметричность	Антисимметричность
Эквивалентность	+	+	+			
Толерантность	+	+				
Строгий порядок			+	+	(+)	(+)
Квазипорядок	+		+			
Нестрогий порядок	+		+			+

Законы композиции

1. На трехэлементном множестве $G=\{a, b, c\}$ задан закон композиции одной из следующих таблиц:

а) $\begin{array}{c|ccc} \top & a & b & c \\ \hline a & b & c & a \\ b & c & c & b \\ c & b & b & c \end{array}$; б) $\begin{array}{c|ccc} \top & a & b & c \\ \hline a & a & c & b \\ b & c & b & a \\ c & b & a & c \end{array}$; в) $\begin{array}{c|ccc} \top & a & b & c \\ \hline a & b & c & a \\ b & c & a & b \\ c & a & b & c \end{array}$.

Определите свойства каждого из этих законов и тип соответствующей алгебраической системы. Укажите нейтральный и симметричные элементы, если они существуют..

2. Дана группа $G = \{a, b, c, d, e, f\}$ относительно закона $\top$, определенного таблицей

T	a	b	c	d	e	f
a	e	f	a	b	c	d
b	d	c	b	a	f	e
c	a	b	c	d	e	f
d	f	e	d	c	b	a
e	c	d	e	f	a	b
f	b	a	f	e	d	c

а) Является ли данная группа коммутативной?

б) Какой элемент из G играет роль нейтрального элемента?

в) Для каждого элемента из G определите симметричный элемент.

г) Покажите, что множества $G' = \{a, c, e\}$ и $G'' = \{c, d\}$ относительно данного закона композиции являются подгруппами данной группы.

3. Укажите тип алгебраической системы, которая образует каждое из приведенных ниже множеств с определенными на нем внутренними операциями:

а) натуральные числа (сложение);

б) натуральные числа (умножение);

в) целые числа (сложение);

г) целые числа (умножение);

д) четные числа (сложение и умножение);

е) действительные числа (умножение);

ж) целые числа (сложение и умножение);

з) действительные или комплексные числа (сложение и умножение).

4. К каким алгебраическим системам относится множество квадратных матриц n-го порядка (n — фиксированное число) относительно операций:

а) сложения;

б) умножения;

в) сложения и умножения?

5. На множестве $G=(a, b, c)$ заданы два внутренних закона композиции — аддитивный и мультипликативный:

+	a	b	c
a	b	c	a
b	c	a	b
c	a	b	c

;

	a	b	c
a	a	b	c
b	b	a	c
c	c	c	c

.

а) Покажите, что G с заданными на нем законами образует коммутативное тело.

б) Определите нейтральные элементы относительно заданных законов композиции.

в) Для элементов из G определите симметричные элементы относительно заданных законов композиции.

г) Замените элементы a, b, c соответственно на числа 1, 2 и 3 и истолкуйте заданные операции в числовом множестве.

6. Покажите, что целые числа, кратные некоторому числу p, составляют идеал в кольце целых чисел.

Составные отношения

1. Пусть R и S определены на P, где P — множество всех людей, следующим образом:

$$R=\{(x, y): x, y \in P \text{ и } x \text{ является отцом } y\},$$

$$S = \{(x, y): x, y \in P \text{ и } x \text{ — дочь } y\}.$$

Описать явно следующие отношения:

а) R^2; б) S^2; в) $R\circ S$; г) $S \circ R$;

д) $S \circ R^{-1}$; е) $R^{-1}\circ S$; же) $R^{-1}\circ S^{-1}$;

з) $S^{-1}\circ R$; и) $S^{-1}\circ S^{-1}$; к) $S^{--1}\circ R^{-1}$.

2. Используя отношения R и S упражнения 1, описать замыкание следующих отношений:

а) R^+; б) S^+; в) R^*; г) S^*;

д) $(S \circ S^{-1})^*$; е) $(R^{-1}\circ R)^*$; $(S^2 = R^2)^+$.

Матрицы и бинарные отношения

1. Пусть A — конечное множество и $|A|=n$, а M — матрица над (Z_2, $\wedge$, $\vee$), соответствующая некоторому бинарному отношению на A. Доказать, что

$$M^+ = \sum_{p=1}^{n} M^p.$$

(Следствием этого является тот факт, что вместо полукольца (Z_2, $\wedge$, $\vee$) мы можем рассматривать булеву алгебру (B, $\wedge$, $\vee$, $\neg$), где

$B = \{0, 1\}$. Поэтому мы часто будем обозначать множество матриц $n \times n$ через $M(n, B)$ и называть их *булевими матрицами.)*

2. Доказать, что если M — конечная квадратная матрица над $(Z_2, \wedge, \vee)$, то

$$M^* = (I + M)^+.$$

Указание: см. задачу 1 из п. 3.13.

3. Показать, что если матрица M над $(Z_2, \wedge, \vee)$ такая, что $I \le M$, то $M^n \le M^{n+1}$ для любого $n \in N$. В качестве следствия доказать, что если M имеет размер $n \times n$ и $p \ge m$, то

$$M^* = (I + M)^p, \qquad M^* = M^{2^q}$$

для некоторого q такого, что $2^q \ge m$.

4. Показать, что если существует обратная к M матрица M^{-1} (т.е. $M^{-1}M = MM^{-1} = I$), то она единствена.

Доказать также, что если N обратима и согласована с M, то

$$(NM)^{-1} = M^{-1}N^{-1}.$$

5. Доказать, что если A, B и C — согласованные матрицы такие, что

$$A^*B = 0 = A^*C,$$

то отсюда не следует равенство

$$B = C.$$

13.Геометрический подход к проблеме группового выбора

Современный этап в развитии теории систем и системного анализа отмечен обращением к чрезвычайно сложным и плохо определенным системам, связанным с деятельностью людей. Для решения задач, возникающих в процессе управления такими системами, имеющейся объективной информации оказывается явно недостаточно и тогда дополнительно привлекается субъективная информация — индивидуальные суждения высококвалифицированных специалистов (экспертов). В этих условиях особенно повышается роль субъективной информации, а также выдвигаются новые требования к методам ее обработки и к обоснованности решенийг принимаемых на ее основе.

В реальных задачах принятия решений с использованием экспертных суждений трудно рассчитывать на то, что при поиске решения можно будет ограничиться формальными математическими методами, однако их использование зачастую становится полезным и необходимым на разных этапах фомирования приемлемого коллективного решения.

Очевидно, что при решении многих задач системного анализа тот или иной способ принятия решения должен опираться на групповые суждения экспертов. Попытка усреднить отдельные мнения часто приводит к тому, что с этим средним мнением каждый эксперт в отдельности согласен не больше, чем, может быть, с точкой зрения другого эксперта. Поэтому особое внимание привлекают такие способы построения групповых решений, которые опирались бы не столько на формальные правила, сколько на «разумные», легко воспринимаемые людьми принципы, позволяющие придавать содержательное толкование получаемым решениям.

К таким «разумным», «естественным» принципам относится принцип единогласия Парето для согласования индивидуальных предпочтений. Пожалуй, ни один другой принцип согласования не обсуждался в научной литературе так глубоко и детально, как этот, и в то же самое время он, пожалуй, единственный, для которого не существовало техники получения групповых решений. Это обстоятельство, ввиду практической важности проблемы построения групповых решений предопределяет интерес к поиску математического аппарата для получения групповых решений, удовлетворяющих принципу единогласия Парето. Использование принципа Парето позволяет сделать акцент в сторону максимального учета индивидуальных предпочтений экспертов и разработать такую общую процедуру, чтобы усредненное мнение более или менее удовлетворяло все множество экспертов.

Настоящий раздел, основное содержание которого принадлежат В.Б.Кузьмину, лежит в русле того направления в моделировании группового выбора, которое **трактует индивидуальные мнения (предпочтения) как <u>точки в пространстве соответствующих бинарных отношений</u>.** Распространенное понимание принципа Парето как принципа отбраковки тех и только тех альтернатив, которые «единогласно подчинены» другим альтернативам, сводится к построению «пересечения» индивидуальных бинарных отношений. В работе используется более **широкое истолкование принципа Парето как лишь <u>необходимого (не обязательно достаточного</u>) условия для группового решения, но зато отнесенного не только к «отбраковке», но и к «принятию» альтернатив.** Это истолкование сводится к выделению в качестве группового предпочтения некоторого «промежуточного» бинарного отношения «между» заданными индивидуальными отношениями. **Изучение формализации понятий «промежуточности» и производных от него понятий «выпуклости»**

и т. п. в пространстве отношений позволяет развить «геометрический» взгляд на изучаемые объекты.

При разработке этого «геометрического» подхода В.Б.Кузьмин исходил из более широкой задачи: **провести общее исследование геометрии пространств отношений индивидуального предпочтения, наиболее распространенных в практике экспертного метода**. Надо отметить, что для решения задачи группового выбора в пространствах отношений традиционно применяется метрическая структура, которая используется в общеизвестных «механических» правилах нахождения группового решения — в медианах и средних. При этом в стороне остаются такие геометрические структуры, как «между», выпуклости, частичного порядка и др. Исследование именно таких структур является целью настоящего раздела работы.

Исследования геометрических структур проводятся для **пространств двух типов бинарных отношений: <u>обычных (четких) и нечетких</u>**.

В первой части растоящей работы будут рассмотрены **пространства бинарных отношений: <u>обычных (четких)</u>**.

Общая схема исследований следующая: рассматриваются два варианта понятия «между» для отношений и соответственно два варианта понятия «выпуклости» множества отношений, исследуется взаимосвязь между ними и предлагается алгоритм построения выпуклой оболочки множества индивидуальных отношений и его «ядра», содержащего искомые групповые решения, т. е. решения, удовлетворяющие принципу единогласия Парето.

Исследование геометрических структур проводится на двух уровнях общности: сначала проводится общее рассмотрение про странств бинарных отношений (четких или нечетких) без конкретизации типа отношений, а затем результаты переносятся на конкретные пространства. Так, для пространств четких бинарных отношений вводится понятие полных пространств, рассмотрение которых допускает изучение конкретных пространств с общих позиций и упрощает доказательства; доказывается теорема о существовании линейного сегмента в произвольном пространстве бинарных отношений; для полных пространств доказывается эквивалентность двух определений выпуклости, одно из которых есть полный аналог соответствующего понятия в евклидовой геометрии, а другое —- экспликация принципа Парето и т. д. Вместе с тем результаты общего изучения нельзя перенести механически в некоторые конкретные пространства и здесь требуется проведение дальнейших исследований. Так, например, остался невыясненным вопрос о том, является ли пространство линейных квазипорядков полным пространством; понятие «ядра» исходной совокупности индивидуальных предпочтений может рассматриваться в различных модифицированных вариантах, отвечающих специфике решаемой задачи, и требует поиска структурных характеристик.

В настоящей работе результаты общего изучения геометрии пространств предпочтений исследуются в приложении к пространствам частичных порядков и квазитранзитивных отношений, и решение задачи группового выбора доводится в них до машинных алгоритмов.

Методологическая особенность развиваемого в работе подхода состоит в том, что поиску единственного группового решения сначала предшествует построение множества «допустимых» групповых решений, удовлетворяющих принципу Парето. Выбор единственного группового решения производится уже из построенного множества допустимых групповых решений.

Интерес к пространствам нечетких отношений связан с тем, что язык теории нечетких отношений позволяет получать оценки той степени (меры, вероятности), с которой исследуемые объекты находятся в данном отношении, и поэтому во многих практических задачах оказывается наиболее адекватным условиям экспертного оценивания и целям экспертизы.

Проблема агрегирования индивидуальных суждений о порядке предпочтений на конечном множестве исследуемых объектов в содержательном плане состоит в нахождении такого суждения, которое было бы «средним» относительно исходных или, другими словами, находилось бы «в середине между ними». Для поиска и оценки суждений, претендентов на роль такого среднего, часто используется метрический подход. При этом подходе исходные данные задачи агрегирования, например, «погружаются» в подходящую абстрактную модель, скажем, взятую из физики модель пространственного расположения единичных масс, и задачу нахождения среднего трактуют как задачу нахождения центра тяжести. Подобный прием был использован, например, Кемени и Снеллом, которые использовали понятие расстояния между суждениями, а в качестве среднего выбиралось суждение, сумма расстояний до которого от исходных суждений минимальна. Именно с этим метрическим подходом будет производиться сравнение подхода, развиваемого далее.

Отметим, что для решения проблемы агрегирования индивидуальных предпочтений используются и некоторые другие подходы, в книге не

рассматриваемые. Среди них можно упомянуть аксиоматический подход и статистические методы.
В этом разделе мы попытаемся выявить причины обращения к метрическим моделям и пространствам предпочтений при решении проблемы группового выбора. При этом будет показано, что к понятию модели приводят два пути, использующие различные способы представления индивидуальных предпочтений. Кроме того, такое рассмотрение позволит понять некоторые аспекты метрического подхода к проблеме группового выбора, присущие также и развиваемому подходу.

13.1. Представление предпочтений

Этапу обработки индивидуальных предпочтений неизбежно предшествует этап измерений самих предпочтений. В современной науке под **измерением понимают строгую и для данного вида измерений почти всегда воспроизводимую с «близким» результатом процедуру, при которой измеряемое свойство объекта (признак, фактор) сравнивается с некоторым эталонным его проявлением и в результате получается число, характеризующее степень выраженности исследуемого свойства.** Такая трактовка измерения заимствована из классической физики и относится к случаю, когда исследуемые свойства обладают количественной природой или мы умеем их наделять таковой, как, например, в случае измерения количества тепла.
В случае качественных свойств корректно говорить об измерении, если в приведенной выше трактовке понятия «измерение» переставить некоторые акценты. Поскольку структура качественных свойств, как правило, «неосязаема», то «легализация» свойств начинается — явно или неявно — с построения некоторого соответствия между совокупностью наших содержательных представлений о природе свойств и описанием этой совокупности в точных терминах.
В теории измерений качественных свойств такое «описание в точных терминах» структуры исследуемых свойств производится с использованием понятия *системы с отношениями.* При этом, **говоря, что свойство имеет определенную структуру, подразумевают любую структуру, детерминированную эмпирическими связями между эмпирическими объектами**. Этот общий подход распространяется также на случай «эмпирических отношений», которые являются субъективными утверждениями об отношениях между эмпирическими объектами. Когда имеют дело с качественными

свойствами, объектом измерения как раз и являются эмпирические отношения, а результат выражается через посредство субъективных утверждений об измеряемых отношениях. **Под самим измерением понимают в этом случае установление соответствия между эмпирической и некоторой числовой системами с отношениями, а под шкалой — совокупность из обеих систем с отношениями и соответствия между ними.**

Рассмотренную ситуацию формально можно описать следующим образом. Пусть *A* — конечное множество объектов, называемое *носителем* системы. Объекты из *A* являются носителями исследуемого свойства (признака, фактора). Множество *A* вместе с фиксированным на нем множеством отношений R_i $(i \in I)$ называется *системой с отношениями* и обозначается $\mathrm{A}=<(A; R_i\ (i \in I)>$, $I = \{1, ..., n\}$, $n= |A|$. В зависимости от того, какова природа объектов и отношений (эмпирическая или числовая), система с отношениями называется *эмпирической* или *числовой системой с отношениями* соответственно. Простейшей эмпирической системой с отношениями, служащей основанием для шкал наименований, является система $\mathrm{A} = <A; \approx>$, где $\approx$ есть отношение эквивалентности.

В большинстве встречающихся на практике случаев измерение качественных свойств производится в шкалах порядка, приводящих к представлению результатов измерений в виде ранжирований. В этом случае эмпирическая система с отношениями содержит по крайней мере отношения эквивалентности и порядка такие, что соответствующая система с отношениями представляет собой упорядоченное множество. Примером такой системы может служить множество индивидуумов, сравниваемых, скажем, по уровню умственных способностей.

На языке введенных понятий теории измерений отображение m: $\mathrm{A} \to \mathrm{R}$ называется *шпалой порядка,* если m — монотонно возрастающее преобразование, где А — эмпирическая система $<A; \approx, \mathrm{p}>$ с отношениями эквивалентности $\approx$ и строгого предпочтения p, а R — числовая система $<R; =, <>$ с отношениями равенства = и порядка <. Такое отображение называется *гомоморфизмом* системы.

Широкое использование шкал порядка в практике получения данных об индивидуальных предпочтениях объясняет интерес многих исследователей к методам обработки измерений в этих шкалах, в том числе и при поиске групповых предпочтений. Поэтому естественно обращение к теории, в рамках которой рассматриваются условия, налагаемые на методы обработки измерений в различных шкалах, т. е. обращение к теории измерений.

Отметим далее, что, вообще говоря, существует целый класс шкал, отображающих данную эмпирическую систему в числовую. Эти шкалы между собой связаны допустимыми преобразованиями, которые переводят данную шкалу в эквивалентные ей шкалы.
Для шкалы порядка класс допустимых преобразований составляют все монотонно возрастающие преобразования, ими же ограничиваются и возможности обработки результатов измерений в этой шкале. По этой причине многие традиционные способы обобщенного описания (осреднения) результатов таких измерений непригодны, поскольку для статистик, включающих средние стандартные отклонения, знание лишь порядка следования недостаточно. К допустимым статистикам в данном случае относятся только квантили и процент присвоения данному объекту определенного ранга, чего явно недостаточно для решения практических задач.
Таким образом, обращение к теории измерений позволяет нам только выявить те ограничения, которые необходимо соблюдать при «арифметической» обработке индивидуальных суждений о порядке предпочтений на исследуемых объектах, и практически не дает нам никаких средств, позволяющих находить групповое предпочтение. Поэтому невольно возникает мысль искать такое представление для измерений предпочтений, которое, с одной стороны, гарантировало бы удовлетворение выявленных ограничений на методы обработки измерений, а с другой стороны, давало бы возможность применения других «неарифметических» методов обработки измерений.
В соответствии с тем, что исходными данными в проблеме агрегирования служат совокупности индивидуальных отношений предпочтения, мы будем рассматривать совокупности U из m индивидуальных суждений о порядке предпочтений на множестве $A = \{a_1, ..., a_n\}$. Поэтому числовая система R, которую мы будем рассматривать, состоит из m-мерного линейного пространства $\mathbf{R}^m$ — носителя числовой системы и совокупности из m отношений предпочтения на $\mathbf{R}^m$, причем i-е отношение задается сравнением двух векторов из $\mathbf{R}^m$ по величине i-й координаты.
Любой гомоморфизм эмпирической системы в выбранную числовую систему задается функцией f, которая каждому объекту из эмпирической системы ставит в соответствие точки в носителе числовой системы, т. е. точки в пространстве $\mathbf{R}^m$. Задающие эти гомоморфизмы числовые функции f называются *представляющими функциями* (*другими* словами, действительная функция f на объектах $\{a_1, a_2, \ldots\}$ представляет данное упорядочение множества

A, если числа $\{f(a_1), f(a_2), \ldots\}$ упорядочены по величине так же, как объекты множества *A* в данном упорядочении.). Для системы U представляющие функции являются действительными *m*-мерными вектор-функциями. Построенная шкала определяется упорядоченной тройкой $\langle \mathcal{U}, \mathcal{R}, f \rangle$.

Для данных эмпирической и числовой систем могут существовать различные представляющие функции *f*. Обозначим через φ преобразование, связывающее два некоторых гомоморфизма *f* и *g* данной эмпирической системы в числовую систему. Это преобразование задает изоморфизм числовой системы на себя. Для системы R такие изоморфизмы могут, например, задаваться монотонными преобразованиями координат векторов из R^m, причем каждая координата подвергается своему монотонному преобразованию.

Множество всех изоморфизмов φ образует группу Φ.

Мы скажем, что две шкалы $\langle \mathcal{U}, \mathcal{R}, f \rangle$ и $\langle \mathcal{U}, \mathcal{N}, g \rangle$ принадлежат шкалам одного и того же типа Φ, если существует такое $\varphi \in \Phi$, что диаграмма на рис. 1.1 коммутативна.

$$\begin{tikzcd} & R^m \arrow[dd, "\varphi"] \\ A \arrow[ur, "f"] \arrow[dr, "g"'] & \\ & R^m \end{tikzcd}$$

Рис. 1.1

Рассмотрим теперь векторное пространство **E** всех функций на *A* со значениями в R^m. Размерность **E** равна $m \times n$. Для каждой фиксированной системы U множество ее представляющих функций является подмножеством в пространстве **E**. Пространство **E** будем называть *пространством шкал.* Группа Φ действует на пространстве шкал **E** как группа взаимнооднозначных преобразований. Именно, каждой $f \in$ **E** ставится в соответствие $f' = \varphi^{\circ} f$ для $\varphi \in \Phi$. Все пространство распадется на орбиты (*орбитой* элемента под действием группы преобразований называется множество всех его образов при действии элементами этой группы) относительно группы Φ, т. е. представляется в виде объединения непересекающихся минимальных подмножеств. Каждая такая орбита состоит из множества всех представляющих функций для некоторой системы U, и, наоборот, для каждой системы U множество всех ее представляющих функций составляет орбиту группы Ф. Тем самым между множеством орбит пространства шкал **E** и множеством всех эмпирических систем *{*U*}* установлено взаимнооднозначное соответствие. Каждую орбиту в **E** можно задать с помощью уравнений и неравенств.

В теории измерений вопросы обработки данных исследуются в связи с так называемой *проблемой адекватности.* Проблема адекватности возникает в общем случае при рассмотрении вопроса о том, какие отношения в числовой системе соответствуют «истинным» отношениям в эмпирической системе. Простые примеры показывают, что не любые отношения между результатами измерений соответствуют «истинным» отношениям между объектами. Более того, указания на некоторые отношения между измерениями могут оказаться бессмысленными без указания на то, в какой шкале были произведены измерения. Например, справедливость числового утверждения, что квадрат массы одного тела меньше массы другого тела, зависит от шкалы измерения масс. Другими словами, возникает проблема адекватности операций, которые производятся над измерениями, тем отношениям, которые существуют в эмпирической системе.

Общий подход к определению адекватности намечен в литературе по теории измерений. Принимая во внимание, что основной объект наших рассмотрений — это совокупности утверждений о порядке предпочтений, определение адекватности мы дадим здесь в удобной для нас форме.

Прежде всего условимся, что под *числовым утверждением* для системы U будем понимать отношение или набор отношений вида $F(x)P0$, где $x \in$ **E**, F — некоторая числовая функция на E, P — одно из отношений <, ≤, = (например, $F(x) < 0$).

Определение 1.1. Числовое утверждение $F(x)P0$ называется *адекватным,* если для любого $\varphi \in Ф$ отношения $F(x)P0$ и $F(\varphi(x))P0$ эквивалентны.

Рассмотрим множество тех $x \in$ **E**, для которых выполняется $F(x)P0$. Наше определение требует, чтобы это множество было инвариантным относительно преобразований $\varphi \in Ф$. Геометрически это означает, что это множество — объединение некоторого числа орбит пространства **E**. Последнее вытекает из следующего утверждения, которое в дальнейшем будет существенно использоваться.

Утверждение 1.1. *Для того чтобы числовое утверждение было адекватно, необходимо и достаточно, чтобы множество тех x, для которых F(x) P*0 *справедливо, было объединением некоторого числа орбит.*

Доказательство этого утверждения в силу его тривиальности мы опускаем.

Для иллюстрации понятия адекватности разберем три примера.

Пример 1.1. Дано $A = \{a, b, c\}$, проверяется адекватность числового утверждения $x_a + x_b > x_c$. Для этого примера **E** — 3-мерное пространство, $F(x) = x_a + x_b — x_c$, P — отношение >. Рассматривается шкала отношений. Так как для этого типа шкал допустимые преобразования — это преобразования подобия $\varphi(x) = kx$ $(k > 0)$, то для этой шкалы орбитами в пространстве **E** будут лучи, выходящие из начала координат. Неравенство $x_a + x_b — x_c > 0$ определяет полупространство в пространстве **E**, которое, естественно, представляется объединением некоторого множества орбит. В силу утверждения 1.1 рассматриваемое числовое утверждение адекватно в шкале отношений.

Пример 1.2. Исследуем то же самое утверждение в шкале интервалов, допустимые преобразования в которой составляют линейные положительные преобразования $\varphi(x) = kx+l$, $k> 0$.

Рассмотрим орбиту точки $\bar{X}_0 = (x_a^0, x_b^0, x_c^0)$. Тогда для любой точки $\bar{X}$, принадлежащей орбите $\bar{X}_0$, имеем $\bar{X} = k\bar{X}_0 + l$, или, в координатной записи, $x_a = kx^0_a + b$ и т. д. Вектор **1** = (*l, l, l*). Если $\bar{X}_o$ пропорционален **1**, то $\bar{X}$ пропорционален **1** тоже, поэтому орбитой такой точки будет $x_a = x_b = x_c$. Если же это не выполняется, то множество векторов $\bar{X}$ (орбита $\bar{X}_0$) заполняет полуплоскость, проходящую через прямую $x_a = x_b = x_c$ и вектор $\bar{X}_o$. Пространство **E** в этом случае расслаивается на множество плоскостей (орбит), опирающихся на орбиту $x_a = x_b = x_c$. Для того чтобы $x_a + x_b — x_c > 0$ было адекватно, согласно утверждению 1.1 необходимо, чтобы определяемое этим неравенством полупространство представлялось в виде объединения некоторого числа описанных полуплоскостей. Это не так в силу того, что направляющий вектор этого полупространства (1, 1, —1) не ортогонален направляющему вектору (1, 1, 1) орбиты $x_a = x_b = x_c$. Полупространство, определяемое неравенством $x_a + x_b — x_c > 0$, пересекает каждую орбиту $\varphi(x)= kx+l$ и ни одной не содержит целиком. Следовательно, утверждение $x_a + x_b — x_c > 0$ неадекватно в шкале интервалов.

Пример 1.3 мы «оживим» фабулой из экспертной практики. Пусть два эксперта производят упорядочение трех объектов a_1, a_2, a_3. В качестве основания числовой системы используется множество натуральных чисел от 1 до 10. В таблице 1.1 приведены оценки $f(a_i)$; i = 1, 2, 3, отражающие степени предпочтения экспертов.

Определим понятие обобщенного среднего по Колмогорову:

$$f_p(X, F) = F^{-1}\left(\frac{1}{p}\sum_{i=1}^{p} F(x^i)\right),$$

где $F(z)$ —строго монотонная функция, $F^{-1}(z)$ —обратная к ней. Зададимся целью выбрать одно из возможных средних для описания «среднего» группового упорядочения. Подсчитаем значения для четырех наиболее распространенных видов средних: квадратичного, арифметического, геометрического и гармонического (см. таблицу 1.2, где для наглядности эти значения приводятся с точностью до постоянных множителей).

Таблица 1.1

Эксперты	Представляющие функции	Значение представляющих функций на объектах			Упорядочение
		a_1	a_2	a_3	
1	f_1	10	9	5	$a_3<a_2<a_1$
2	f_2	1	3	5	$a_1<a_2<a_3$

Таблица 1.2

Вид среднего	Значения среднего на объектах			Упорядочения, соответствующие виду среднего
	a_1	a_2	a_3	
$\sqrt{f_1^2+f_2^2}$	$\sqrt{101}$	$\sqrt{90}$	$\sqrt{50}$	$a_3<a_2<a_1$
f_1+f_2	11	12	10	$a_3<a_1<a_2$
$\sqrt{f_1 \cdot f_2}$	$\sqrt{10}$	$\sqrt{27}$	$\sqrt{25}$	$a_1<a_3<a_2$
$\left(\frac{1}{f_1}+\frac{1}{f_2}\right)^{-1}$	0,91	2,25	2,50	$a_1<a_2<a_3$

Из таблицы 1.2 непосредственно видно, что выбранные средние дают нам четыре различных групповых упорядочения, из которых никакое не имеет очевидного преимущественного права представлять совокупность исходных упорядочений.

В соответствии с определением 1.1 для того чтобы определить пригодность того или иного среднего представлять всю совокупность исходных упорядочений, нужно проверить адекватность набора утверждений, состоящего из совокупности исходных упорядочений и утверждения, полученного на основе выбранного среднего. Упорядочения должны быть записаны в виде числовых утверждений. Так, например, для среднего арифметического нужно рассматривать следующий набор числовых утверждений:

$$\left.\begin{aligned}&f_1(a_3) < f_1(a_2) < f_1(a_2),\\&f_2(a_1) < f_2(a_2) < f_2(a_3),\\&\frac{f_1(a_3)+f_2(a_3)}{2} < \frac{f_1(a_1)+f_2(a_1)}{2} < \frac{f_1(a_2)+f_2(a_2)}{2}.\end{aligned}\right\} \quad (*)$$

Легко привести пример, когда значение истинности утверждения (*) изменяется в результате применения монотонного преобразования. Рассмотрим кусочно-линейную монотонную функцию, которая оставляет на месте точки *1, 3, 5* и *10,* а точку *9* переводит в точку *6* (рис. 1.2).

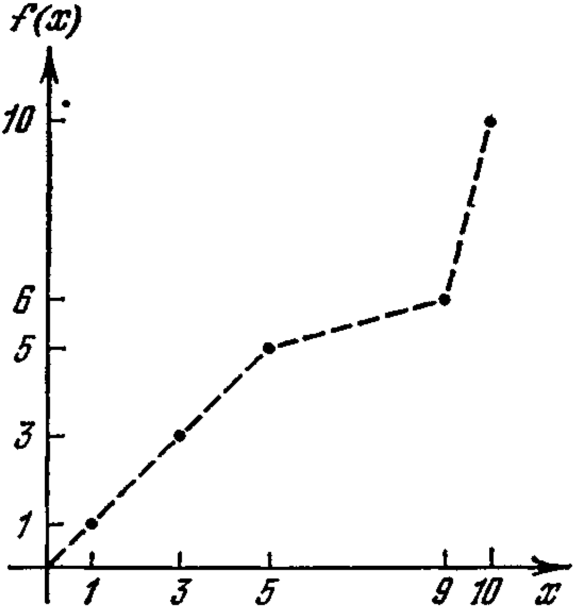


Рис 1.2

Тогда два первых неравенства в системе (*) не изменятся, а третье примет вид

$$\frac{f_1(a_2)+f_2(a_2)}{2}<\frac{f_1(a_2)+f_2(a_3)}{2}<\frac{f_1(a_1)+f_2(a_1)}{2},$$

что соответствует изменению среднего упорядочения с $a_3 < a_1 < a_2$ на $a_2 < a_3 < a_1$ и, следовательно, значение истинности утверждения (*) изменится. Отсюда следует известный факт: среднее арифметическое неадекватно в шкале порядка. Легко привести примеры, показывающие, что другие виды обобщенных средних тоже неадекватны.

Возпикает вопрос, почему использование обобщенных средних для характеристики совокупностей ранжировок дает разные результаты. Глубокая причина этого заключается в том, что при такой обработке ранги или значения представляющих их функций выступают в роли количественной меры оцениваемых объектов, тогда как, по существу, они являются только метками, позволяющими лишь расставить объекты в соответствии с предпочтениями оценивающих их экспертов. Последовательное развитие этого замечания приводит к заключению, что под *измерением* в шкале порядка следует понимать в общем случае не значение представляющей функции на объекте, а *всю ранжировку в целом.* Другими словами, каждое отдельное измерение в шкале порядка множеству оцениваемых объектов ставит в соответствие некоторое бинарное отношение, например, квазипорядок. На введенном выше языке орбит это означает, что порядковая информация о каждой орбите в пространстве шкал **E** представляется соответствующим орбите квазипорядком.

Установление изоморфизма между множеством орбит в пространстве шкал и множеством всех возможных (на фиксированном носителе *A)* отношений данного типа позволяет, следовательно, каждую орбиту заменить ее математической моделью — точкой, помеченной соответствующим орбите отношением. Заметим, что орбиты в пространстве шкал располагаются определенным образом, т. е. в каждом отдельном случае (см. примеры 1.1 и 1.2) в их расположении есть определенная структура. Для наведения на множестве помеченных точек структуры, удобной для наших дальнейших рассмотрений, естественно обратиться к привычному, сложившемуся геометрическому представлению о расстоянии. Таким образом, определение на множестве U всех бинарных отношений данного типа метрической структуры *d* дает возможность исследовать совокупность индивидуальных предпочтений в рамках метрической модели.

Такая модель, следовательно, имеет два геометрических представления. Одно, как бы внешнее относительно модели, является следствием показанного изоморфизма между моделью и пространством шкал **E**, другое, внутреннее, индуцируется метрической структурой в самой модели.

Представление измерений в шкале порядка точками в модели заведомо снимает проблему адекватности, поскольку при работе внутри модели исходные точки уже не затрагиваются, не сдвигаются: все дальнейшие операции производятся лишь с внутренней относительно модели характеристикой взаимного расположения точек — попарными расстояниями между ними. Проблема группового выбора решается при этом указанием такой точки T модели, в которой достигается минимум суммы расстояний или суммы квадратов расстояний от T до данных точек. (Существуют и другие подходы к решению проблемы адекватности. Так, изложение теории измерений на языке допустимых преобразований, нацеленное на изучение адекватности, и решение задачи нахождения адекватных средних для наиболее распространенных шкал см. в работах А. И. Орлова; статистический подход к решению задачи нахождения адекватного среднего в шкалах порядка описывается в работах других авторов).

При анализе данной совокупности ранжировок сами ранжировки не подвергаются обработке — функция расстояния обеспечивает основные данные, на которых уже работают самые разные аналитические аппараты.

Достоинства метрических моделей выявляются в приложениях к многообразным теоретическим и практическим задачам в различных областях от «чистой» математики до полевых социологических исследований. Однако, как отмечается в литературе, существуют некоторые трудности использования таких моделей при анализе совокупностей экспертных суждений и поиске группового решения.

Как указывается в литературе, недостатки данных о расстоянии с первого взгляда не кажутся очевидными. Для двух или другого небольшого числа ранжировок расстояния между ними или до какой-либо внутренней точки модели еще дают представление о структуре связей между ними. Однако, при большом числе оцениваемых объектов (>4) и большом числе исходных ранжировок никакого реального представления об их расположении и структуре связей составить невозможно. Ценность анализа таких данных, например, методами факторного анализа или распознавания образов в значительной степени зависит уже от искусства, проявляемого

исследователем при интерпретации результатов применения этих методов.
При построении метрической модели приходится вводить и определять довольно много понятий: «между», линейный сегмент, путь, масштаб расстояния. При этом модель наделяется теми характеристиками, которые были использованы в определяемых понятиях и вводимых аксиомах. Второй недостаток метрической модели как раз и состоит в том, что при аксиоматическом определении функции расстояния вводятся положения, не всегда имеющие эмпирическое обоснование или подтверждение. К таким положениям можно отнести, например, условия равноправия объектов и равноценности места в ранжировании или связь масштаба расстояния с числом оцениваемых объектов для метрики в пространстве разбиений.

13.2. Геометрический подход

Богатство геометрических структур в моделях для различных отношений предпочтения не исчерпывается одной метрической структурой. В качестве примера можно указать на работы, посвященные изучению топологических структур на множествах предпочтений. Большой интерес представляет структура выпуклости, которую можно определить во множестве предпочтений. Наше основное внимание, начиная с главы 15, будет посвящено общему изучению этой структуры.
Интерес к выпуклым структурам продиктован, во-первых, простыми аналитическими соображениями. Поскольку неадекватные операции по обработке измерений в шкале порядка все-таки применяются, то исследовался вопрос о том, как результаты применения этих операций соотносятся (геометрически) с исходными данными. Отметим, что неадекватная арифметическая обработка качественных данных (определяемых в общем случае значениями представляющих функций) с помощью произвольных математических средних, дает результат, принадлежащий выпуклой оболочке исходных данных. Во-вторых, **выпуклые структуры, как показывают проводимые далее исследования, помогают раскрыть качественно новые и важные эмпирические связи в практических приложениях.**
Для терминологического различения в нашем подходе вместо понятия метрической модели используется понятие ***пространства*** **как множества всех возможных на фиксированной совокупности объектов отношений данного типа, <u>наделенного структурой выпуклости</u>.**

Единственным общим понятием метрического и развиваемого в настоящей работе подхода является понятие «между», используемое при определении структуры выпуклости. Главное «техническое» отличие рассматриваемого здесь подхода к анализу индивидуальных отношений предпочтения от аналитических методов, используемых в метрическом подходе, заключается в **последовательном использовании исключительно геометрических <u>понятий и структур</u>**. Естественно, что при построении алгоритмов, реализующих развиваемый подход, и интерпретации результатов его практического использования мы оперируем только геометрическими (внутренними относительно рассматриваемых пространств) понятиями, что дает основание весь подход в целом назвать *геометрическим* подходом к анализу индивидуальных отношений предпочтения.

В заключение этого параграфа остановимся на одном обстоятельстве, связанном с общей методологией решения проблемы группового выбора. **В канонической теории принятия решений под *групповым выбором* понимается задание такой функции, которая исходному множеству индивидуальных предпочтений ставит в <u>соответствие одно единственно групповое предпочтение</u>.** Вопросы существования и свойств таких функций обсуждались в литературе очень широко, и они не входят в круг рассматриваемых здесь вопросов. Отметим лишь, что в большинстве исследованных случаев попытки построения функции группового выбора, как правило, приводили либо к ее неоднозначности, либо к доказательству того, что при постулированных условиях такой функции не существует (теорема Эрроу о «невозможности»). Так, в метрической модели в качестве группового выбора используют медиану или среднее, которые определяются неоднозначно. Кроме того, если пытаться рассматривать, например, множество всех медиан для исходной совокупности точек, то описание этого множества наталкивается на существенные аналитические трудности.

В этой связи общую проблему группового выбора предлагается рассматривать <u>на двух уровнях общности</u>.

На первом уровне целесообразно полностью абстрагироваться от конкретного физического содержания решаемой практической задачи (в частности — от конкретных шкал, в которых произведены измерения) и рассматривать индивидуальные суждения как наборы точек в некоторой модели или пространстве. Как указывалось в 13.1, при этом подходе полностью снимаются вопросы, связанные с решением проблемы адекватности в данной шкале. При рассмотрении задачи группового выбора в такой модели или пространстве нет оснований требовать от функции группового выбора однозначности. Более естественным представляется подход, при котором исходному набору индивидуальных суждений ставится в соответствие некоторое множество альтернатив, в определенном смысле *допустимых* для поиска среди них группового решения. К такому же результату, как это будет показано ниже, приводит также содержательное и формальное исследование понятий выпуклой оболочки множества исходных предпочтений и некоторого выпуклого подмножества этой оболочки — так называемого «ядра».

На втором уровне общности рассматривается задача построения функции группового выбора, которая учитывает все специфические особенности задачи. Сюда входит, например, учет компетентности экспертов, учет всех параметров, по которым производится оценивание, процедурные вопросы, предпочтения лица, выбирающего окончательное решение и т. п. Несмотря на то, что в каждой конкретной ситуации, по-видимому, удается найти групповое решение, задача построения на этом уровне общей функции группового выбора представляется бесперспективной. Заметим, что числовые данные, которые получаются в результате измерений в определенных шкалах, являются данными для решения задачи группового выбора на втором уровне.

Предлагаемый подход позволяет решать задачу агрегирования на первом из указанных выше уровней. Применение разработанных методов позволяет опираясь на исходную совокупность индивидуальных предпочтений, выделить некоторое множество «допустимых» групповых решений. Задача нахождения единственного группового решения может рассматриваться как задача выбора из этого множества «допустимых» решений на втором уровне, т. е. с учетом всех специфических условий.

14. Бинарные отношения

В предыдущей главе при рассмотрении некоторых понятий теории измерений мы ввели понятие «тип шкалы», в которой измеряются предпочтения. Со шкалой каждого типа связан определенный способ оценивания. Так, при измерении в шкале порядка объекты расставляются последовательно в соответствии с убыванием степени их предпочтительности или значений представляющих функций на объектах. С измерениями в интервальной шкале мы сталкиваемся в повседневной жизни, когда, например, измеряем длины, температуру,

время. Способ измерений в шкале отношений хорошо иллюстрирует процедура взвешивания на обычных весах.
С каждым способом измерения в шкале данного типа связана определенная форма представления результатов измерений. Так, результаты измерений в шкале порядка на данном множестве объектов могут быть представлены, например, в виде значений представляющих функций на этих объектах или в виде последовательности номеров мест, упорядоченных так же, как и значения представляющих функций, или же могут быть заданы в виде неравенств. Выбор той или иной формы для представления результатов измерений предпочтений в шкале данного типа в конечном итоге определяется тем математическим аппаратом, который предполагается использовать для анализа и обработки измерений, или принципом их агрегации.
В последующих главах этой работы мы будем систематически использовать аппарат обычной теории множеств и теоретико-множественное представление измерений предпочтений. Язык бинарных отношений и булевых матриц позволяет рассматривать более широкий спектр качественных данных об объектах, чем терминология теории измерений. При этом определенному типу качественной информации соответствует определенный класс отношений. Не останавливаясь на обосновании такого выбора, в этой главе мы введем понятие бинарного отношения, дадим краткий перечень действий (операций) над отношениями и основных свойств отношений.

14.1. Понятие бинарного отношения

Пусть **A** — некоторое конечное множество **A** = *(x, y,* z, ...). Рассмотрим множество всех пар элементов *(x,* у), где *x* и $y \in$ **A**. При условии, что *x* не совпадает с *y,* мы будем различать пары *(x, y)* и *(y, x),* т. е. считать элементы в парах упорядоченными. Множество всех упорядоченных пар (*x*, *y),* где (*x*, *y)* $\in$ **A**, обозначается **A** × **A**. Любое подмножество *P* множества **A** × **A** называется *бинарным отношением* на множестве **A.** Множество **A** иногда называют областью задания отношения *P.* Тот факт, что пара *(x, y)* элементов *x*, *y* $\in$ **A** состоит в отношении Р, будет записываться (*x*, *y)* $\in P$ или *xPy*. Бинарное отношение может задаваться или непосредственным указанием пар элементов, состоящих в данном отношении, или правилом, которое позволяет для каждой пары установить, находится или нет данная пара в данном отношении.
Пусть на *n*-элементном множестве **A** определено бинарное отношение *P*. Перенумеруем элементы множества **A** числами от 1 до *n* и каждому

i-му элементу поставим в соответствие i-й столбец и i-ю строку квадратной таблицы размером $n\times n$. Тот факт, что для пары элементов с номерами i и j выполняется отношение P: $(x_i, x_j) \in P$ будем отмечать единицей на пересечении i-й строки и j-го столбца и нулем — если $(x_i, x_j) \notin P$, т. е. если отношение P для элементов x_i и x_j не выполняется. Обозначая элементы построенной таким образом матрицы через a_{ij} правило задания отношения можно сформулировать так:

$$a_{ij} = \begin{cases} 1, & \text{если выполняется } x_i P x_j, \\ 0, & \text{если не выполняется } x_i P x_j. \end{cases}$$

Такой способ представления бинарного отношения называется *матричным.* Полученная матрица называется *матрицей бинарного отношения P.* Мы будем обозначать ее (a_{ij}^P) или просто (P).

14.2. Действия над бинарными отношениями

В этом параграфе мы определим действия над бинарными отношениями, которые опять приводят к бинарным отношениям.

1. *Пересечение* $P \cap Q$ отношений называется отношение, которое содержит только общие для P и Q пары

$$(x, y) \in \mathbf{A}: P \cap Q = \{(x, y): (x, y) \in P \text{ и } (x,y) \in Q\}.$$

Когда P и Q не имеют общих пар, т. е. не пересекаются, то говорят, что их пересечение пусто и записывают $P \cap Q = \varnothing$.

Пример 2.1. Пусть матрицы отношений $P \cap Q$ имеют вид (для большей наглядности нули в матрицах отношений всюду опущены.)

$$(P) = \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline & 1 & \\ \hline & & 1 \\ \hline \end{array} \qquad (Q) = \begin{array}{|c|c|c|} \hline 1 & & \\ \hline 1 & 1 & 1 \\ \hline & 1 & \\ \hline \end{array}$$

Тогда очевидно, что матрица отношения $P \cap Q$ имеет вид

$$(P \cap Q) = \begin{array}{|c|c|c|} \hline 1 & & \\ \hline & 1 & \\ \hline & & \\ \hline \end{array}$$

Таким образом, матрица отношения $P \cap Q$ есть булево пересечение матриц отношений P и Q.

2. *Объединением* отношений $P \cup Q$ называется отношение, которое включает все пары, содержащиеся или в подмножестве P или в подмножестве Q: $P \cup Q = \{(x, y)\colon\ (x, y) \in P$ или $(x, y) \in Q\}$. Когда объединение $P \cup Q$ содержит все возможные пары из $\mathbf{A} \times \mathbf{A}$, а пересечение P и Q пусто, то говорят, что отношения P и Q образуют *разбиение* $\mathbf{A} \times \mathbf{A}$, а их объединение есть полное отношение.

Пример 2.2. Для отношений P и Q из примера 2.1 очевидно имеем

$$(P \cup Q) = \begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline & 1 & 1 \\ \hline \end{array}$$

Таким образом, матрица отношения $P \cup Q$ есть булева сумма матриц отношений P и Q.

3. *Разностью* $P \backslash Q$ отношений P и Q называется отношение, состоящее из тех пар $(x, y) \in P$, которые не содержатся в Q:

$$P \backslash Q = \{(x, y)\colon (x, y) \in P \text{ и } (x, y) \notin Q\}.$$

Частный случай разности двух отношений представляет собой операция взятия дополнения к отношению P (см. ниже п. 5).

Пример 2.3. Для отношений P и Q из примера 2.1 имеем

$$(P \backslash Q) = \begin{array}{|c|c|c|}\hline & 1 & 1 \\ \hline & & \\ \hline & & 1 \\ \hline \end{array} \qquad (Q \backslash P) = \begin{array}{|c|c|c|}\hline & & \\ \hline 1 & & 1 \\ \hline & 1 & \\ \hline \end{array}$$

4. *Симметрической разностью* $P \Delta Q$ называется отношение, состоящее из тех пар (x, y), содержащихся в объединении $P \cup Q$, которые не содержатся в пересечении $P \cap Q$. Другими словами,

$$P \Delta Q = (P \cup Q) \backslash (P \cap Q) = (P \backslash Q) \cup (Q \backslash P).$$

Пример 2.4. Для отношений P и Q из примера 2.1 имеем

$$(P \Delta Q) = \begin{array}{|c|c|c|}\hline & 1 & 1 \\ \hline 1 & & 1 \\ \hline & 1 & 1 \\ \hline\end{array}$$

5. *Дополнением* $\overline{P}$ называется отношение, состоящее из тех пар $(x, y) \in \mathbf{A} \times \mathbf{A}$, которые не входят в P: $\overline{P} = \{(x, y)\colon (x, y) \in \mathbf{A} \times \mathbf{A}$ и $(x, y) \notin P\}$. Отношения P и $\overline{P}$ образуют разбиение $\mathbf{A} \times \mathbf{A}$, т. е.

$$P \cup \overline{P} = \mathbf{A} \times \mathbf{A} \text{ и } P \cap \overline{P} = \varnothing.$$

Пример 2.5. Для отношения P из примера 2.1 имеем

$$(\overline{P}) = \begin{array}{|c|c|c|}\hline & & \\ \hline 1 & & 1 \\ \hline 1 & 1 & \\ \hline\end{array}$$

6. *Обратным отношением* P^{-1} к отношению P называется отношение, которое содержит пару (x, y) тогда и только тогда, когда $(y, x) \in P$, т. е. $P^{-1} = \{(x, y)\colon (y, x) \in P\}$.

Пример 2.6. Для отношения P из примера 2.1

$$(P^{-1}) = \begin{array}{|c|c|c|}\hline 1 & & \\ \hline 1 & 1 & \\ \hline 1 & & 1 \\ \hline\end{array}$$

матрица обратного отношения P^{-1} является транспонированной к исходной матрице отношения P.

7. *Композицией* (произведением) $P \circ Q$ отношений P и Q называется отношение, которое содержит пару (x, y) тогда и только тогда, когда существует $z \in \mathbf{A}$ такое, что $(x, z) \in P$ и $(z, y) \in Q$, т. е. $P \circ Q = \{(x, y)\colon$ найдется z такое, что $(x, z) \in P$ и $(z, y) \in Q\}$,

К частным случаям композиции относится *квадрат отношения P*: $P^2 = \{(x, y)\colon$ найдется z такое, что $(x, z) \in P$ и $(z, y) \in P\}$. По индукции определяется *n-я степень отношения P:*

$$P^n = P^{n-1} \circ P.$$

Тот факт, что $(x, y) \in P^n$ означает, что существует цепочка элементов $x_0 = x, x_1, ..., x_n = y$ такая, что $(x_i, x_{i+1}) \in$ Р для $i = 1, 2, . . ., n$-1.

Пример 2.7. Для отношений *P* и *Q* из примера 2.1 композиция этих отношений имеет матрицу

$$(P \circ Q) = \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline & 1 & \\ \hline \end{array}$$

Матрица композиции отношений *P* и *Q* есть булево произведение матриц этих отношений.

8. *Сужением* отношения *P* на подмножество **B** $\subset$ **A** называется отношение $P_{\mathbf{B}}$ на множестве **B**, которое состоит из всех тех пар $(x, y) \in P$ таких, что $x \in \mathbf{B}$ и $y \in \mathbf{B}$. Другими словами, $\mathbf{P_B} = P \cap (\mathbf{B} \times \mathbf{B})$.

Пример 2.8. Сужение отношений *P* из примера 1 на подмножество **B**, состоящего из первого и третьего элементов, имеет матрицу

$$(P_B) = \begin{array}{|c|c|} \hline 1 & 1 \\ \hline & 1 \\ \hline \end{array}$$

Поскольку бинарные отношения мы рассматриваем как *подмножества* прямого произведения, то для них определено

Отношение включения $P \subseteq Q$, которое выполнено тогда и только тогда, когда каждая пара (x, y), принадлежащая *P*, принадлежит также и отношению *Q*.

14.3. Свойства бинарных отношений

В этом параграфе мы приведем краткие определения важнейших свойств отношений, которые понадобятся нам в дальнейшем.

1. *Рефлексивность* отношения *R* означает, что $(x, x) \in R$, т. е. рефлексивное отношение выполняется между элементом и им самим (xRx). В матрице рефлексивного отношения на главной диагонали всегда стоят единицы.

2. *Антирефлексивность* отношения *R* означает, что $(x, x) \notin R$, т. е. отношение *R* выполняется только для несовпадающих элементов. В матрице антирефлексивного отношения на главной диагонали стоят нули.

3. *Симметричность* отношения *R* означает, что если $(x, y) \in R$, то $(y, x) \in R$, т. е. если для пары (x, y) выполнено отношение *R*, то для

пары (y, x) также выполнено отношение R. В матрице симметрического отношения элементы a_{ij} и a_{ji}, расположенные симметрично относительно главной диагонали, равны между собой.

4. *Антисимметричность* отношения R означает, что если $(x, y) \in R$ и $(y, x) \in R$ R, то $x = y$. Для элементов матрицы антисимметричного отношения выполняется следующее условие: $a_{ij} \bullet a_{ij} = 0$ при $i \neq j$.

5. *Транзитивность* отношения R означает, что из $(x, z) \in R$ и $(z, y) \in R$ следует $(x, y) \in R$. Если в матрице транзитивного отношения элементы $a_{ij} = 1$ и $a_{ij} = 1$, то обязательно $a_{ij} = 1$. Квадрат R^2 транзитивного отношения R, вообще говоря, содержится в самом отношении R: $R^2 \subseteq R$. В случае, если отношение R рефлексивно, то $R^2 = R$.

С понятием транзитивного отношения связано понятие операции *транзитивного замыкания*. Именно, для каждого отношения R определим отношение $\hat{R}$, как наименьшее транзитивное отношение, содержащее данное. Можно показать, что такое отношение определяется единственным образом и $\hat{R} = R \cup R^2 \cup \ldots \cup R^n$, где n — мощность множества **A** — области задания отношения R.

Пример 2.9. Проиллюстрируем это свойство. Пусть отношение R определяется матрицей

$$(R) = \begin{array}{|c|c|c|}\hline & 1 & \\ \hline 1 & & 1 \\ \hline & & 1 \\ \hline \end{array}$$

Тогда имеем

$$(R^2) = \begin{array}{|c|c|c|}\hline 1 & & 1 \\ \hline & 1 & 1 \\ \hline & & 1 \\ \hline \end{array}, \quad (R^3) = \begin{array}{|c|c|c|}\hline & 1 & 1 \\ \hline 1 & & 1 \\ \hline & & 1 \\ \hline \end{array}$$

и транзитивное замыкание отношения R есть матрица

$$(\hat{R}) = \begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline & & 1 \\ \hline \end{array}.$$

6. *Линейность (связность)* отношения R означает, что для любых $x, y \in \mathbf{A}$, $x \neq y$ или $(x, y) \in R$, или $(y, x) \in R$. В матрице линейного отношения или a_{ij}=1, или a_{ji}=1 для любых $i \neq j$.

Поскольку бинарные отношения, с которыми нам придется встречаться в этой книге, как правило, возникают в результате формализации условий выбора, то бинарные отношения будем называть *предпочтениями.* Это могут быть, например, субъективные предпочтения экспертов, лица, принимающего решение, или предпочтения, основанные на функции полезности. Простым примером такого отношения предпочтения является упорядочение некоторого фиксированного множества объектов по степени выраженности признака, определенного на этих объектах, например, по их стоимости.

15. Пространства бинарных отношений

Во многих теоретических исследованиях и практических приложениях приходится рассматривать не произвольные бинарные отношения предпочтения на данном множестве объектов, а отношения предпочтения, на которые наложены некоторые дополнительные условия. Другими словами, обычно рассматривают специальные типы отношений предпочтений, обладающие некоторыми из перечисленных в главе 14 свойств. Особенности решаемой практической задачи предопределяют тип рассматриваемых отношений. Например, в задачах группового выбора обычно используют отношение линейного квазипорядка, соответствующее числовому отношению ≥ (не меньше). Реже используются отношения частичного порядка, совершенного строгого порядка, толерантности и т. п.

Множество всех бинарных отношений предпочтения данного типа с геометрической точки зрения, изложенной в главе 13, образует пространство — пространство бинарных отношений предпочтения данного типа.

В этой главе мы рассмотрим абстрактную модель пространства отношения и двенадцать конкретных реализаций этой модели.

Основное внимание будет сосредоточено на анализе взаимосвязей, существующих между пространствами отношений различных классов.

15.1. Три класса отношений

Пусть **A** — конечное множество объектов. Бинарное отношение R называется *отношением слабого предпочтения* на множестве **A**, если для любых $x, y \in$ **A** выполняется либо xRy, либо yRx.

Из нашего определения немедленно следует, что отношение слабого предпочтения рефлексивно, т. е. xRx для любых $x \in$ **A**, Для остальных пар (x, y) могут иметь место два случая: либо мы имеем xRy, и не выполнено yRx, либо одновременно выполняется xRy и yRx. В первом случае мы говорим, что x строго предпочитается y и пишем xPy. Тем самым *отношение строгого предпочтения* P определяется следующим образом: объекты $x, y \in$ **A** находятся в отношении P (xPy) тогда и только тогда, когда выполнено xRy и не выполнено yRx. Отношение P, очевидно, антиреф-лексивно и антисимметрично.

Во втором случае мы говорим, что выбор между x и y для нас безразличен и будем обозначать это xIy. *Отношение безразличия* I тем самым определено следующим образом: $(x, y) \in I$ тогда и только тогда, когда xRy и yRx. Очевидно, что отношение I рефлексивно и симметрично.

В качестве примера рассмотрим отношение R с матрицей

$$(R) = \begin{array}{|c|c|c|}\hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline & & 1 \\ \hline \end{array}$$

Этому отношению R соответствует строгое отношение предпочтения P с матрицей

$$(P) = \begin{array}{|c|c|c|}\hline & & 1 \\ \hline & & 1 \\ \hline & & \\ \hline \end{array}$$

и отношение безразличия *I* с матрицей

$$(I) = \begin{array}{|c|c|c|} \hline 1 & 1 & \\ \hline 1 & 1 & \\ \hline & & 1 \\ \hline \end{array}$$

Используя определение бинарного отношения как подмножества прямого произведения, взаимосвязь введенных отношений можно сформулировать в виде следующего утверждения:

Утверждение 3.1.*Множества P, I, P^{-1} образуют разбиение прямого произведения* **A**× **A**. *При этом* $R = P \cup I$, $P = \overline{R^{-1}}$ *и* $1 = R \cap R^{-1} = P \cup P^{-1}$.

Справедливость этого утверждения легко проверить для предыдущего примера.

15.2. Пространства предпочтений и безразличия

В практических задачах, которые нам придется рассматривать, мы будем иметь дело не с отдельно взятыми отношениями, а с совокупностями таких отношений. Например, можно было бы рассматривать множество *всех* отношений предпочтения или *всех* отношений безразличия. Однако особенности каждой отдельной задачи обычно сужают это множество в силу того, что на отношения накладываются дополнительно условия, например условие транзитивности и т. п. Основным объектом наших исследований будут именно *подмножества* множества всех бинарных отношений. Дадим следующее общее определение.

Определение 3.2.1. *Пространством бинарных отношений* с носителем **А** называется произвольное подмножество множества всех бинарных отношений на **А**.

Несмотря на большую общность этого определения, на основе его можно получить содержательные результаты для произвольных пространств бинарных отношений. В этой книге, однако, нас будут интересовать лишь отношения слабого и строгого предпочтений и безразличия. В соответствии с этим мы будем рассматривать лишь подмножества, состоящие целиком из элементов одного класса и соответственно этому пространства будем называть *пространствами предпочтения* (слабого или строгого) или *пространствами безразличия.*

Обозначим через R произвольное пространство отношений слабого предпочтения. С каждым пространством *R* связаны пространство P отношений строгого предпочтения *P* и пространство I отношений безразличия *I*. Таким образом, пространство P образовано всеми отношениями *P* такими, что $P = \overline{R^{-1}}$, где $R \in$ R, а пространство I состоит из всех *I* таких, что $I = R \cap R^{-1}$.

Указанные взаимосвязи между пространствами R, P и I легко описать следующим образом. Введем отображения *i*, α и β, отображающие множество всех бинарных отношений, определенных на множестве **A**, в себя:

$$i:\quad \mathbf{M} \mapsto \overline{\mathbf{M}^{-1}},$$
$$\alpha:\quad \mathbf{M} \to \mathbf{M} \cap \mathbf{M}^{-1},$$
$$\beta:\quad \mathbf{M} \to \overline{\mathbf{M} \cup \mathbf{M}^{-1}},$$

где $\mathbf{M} \subseteq \mathbf{A} \times \mathbf{A}$. Сужение отображения *i* (которое мы будем обозначать той же буквой) на пространство R отображает это пространство биективно на пространство P. Аналогично α и β (точнее их сужение на соответствующее пространство) отображают сюрьективно R и P на I.

Эти отображения можно представить в виде следующей диаграммы:

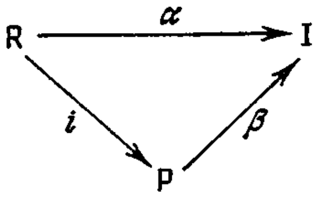


Диаграмма 3.1

Мы закончим этот параграф доказательством того, что эта диаграмма коммутативна.

Утверждение 3.2. *Диаграмма* 3.1 *коммутативна, т. е.* $\alpha = \beta \circ i.$

Доказательство. Пусть $R \in$ R. Имеем $\alpha(R) = R \cap R^{-1}$ и

$$(\beta \circ i)(R) = \beta(i(R)) = \beta(\overline{R^{-1}}) = \overline{\overline{R^{-1}} \cup (\overline{R^{-1}})^{-1}} = R^{-1} \cap R,$$

откуда следует утверждение теоремы,

15.3. Диаграмма пространств

В практических задачах, где используются отношения предпочтения и безразличия, обычно интересно рассматривать не произвольные подмножества таких отношений, а совокупности отношений, обладающие определенными свойствами.

Рассмотрим сначала конкретные пространства слабых предпочтений.

$\mathcal{P}$ — *пространство всех отношений слабого предпочтения.*

$\mathcal{QT}$ — *пространство всех квазитранзитивных отношений,* т. е. пространство всех слабых предпочтений, удовлетворяющих условию квазитранзитивности: для любого $R \in \mathcal{QT}$ отношение $P = \overline{R^{-1}}$ транзитивно.

Пример 3.1. Примером отношения из пространства $\mathcal{QT}$ может служить отношение *R,* заданное матрицей

$$(R) = \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline & 1 & 1 \\ \hline \end{array}$$

Легко проверить, что это отношение нетранзитивно (так как $R \subset R^2$), а отношение $P = \overline{R^{-1}}$ имеет матрицу

$$(P) = \begin{array}{|c|c|c|} \hline & & 1 \\ \hline & & \\ \hline & & \\ \hline \end{array}$$

и транзитивно.

$\mathcal{QO}$ — *пространство линейных квазипорядков.* Получается из $\mathcal{QT}$ требованием транзитивности отношения *R*.

Пример 3.2. Примером отношения из $\mathcal{QO}$ может служить транзитивное отношение, заданное матрицей

$$\begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 1 & 1 & 1 \\ \hline & & 1 \\ \hline \end{array}$$

$\mathcal{TO}$ — *пространство совершенных порядков.* Получается из $\mathcal{QO}$ требованием антисимметричности отношений R.

Пример 3.3. Примером отношения из $\mathcal{TO}$ может служить антисимметричное отношение, заданное матрицей

1	1	1
	1	1
		1

Связь между введенными пространствами слабого предпочтения можно изобразить в виде диаграммы 3.2, где стрелки указывают отображение вложения φ пространств: каждое предыдущее пространство есть подмножество последующего.

$$\mathcal{TO} \xrightarrow{\varphi} \mathcal{QO} \xrightarrow{\varphi} \mathcal{QT} \xrightarrow{\varphi} \mathcal{P}$$

Диаграмма 3.2

Пример 3.4. Теперь приведем пример отношения из пространства $\mathcal{P}$, которое не принадлежит ни одному из вложенных в него пространств (см. диаграмму 3.2);

1	1	1
	1	1
1		1

Это отношение не принадлежит пространству $\mathcal{QT}$, так как отношение $\overline{R^{-1}}$, имеющее матрицу

	1	
		1

нетранзитивно.

Теперь рассмотрим пространства строгого предпочтения, соответствующие, согласно диаграмме 3.1, указанным пространствам слабого предпочтения.

$\mathcal{SP}$ — *пространство всех отношений строгого предпочтения.*

$\mathcal{PO}$ — *пространство всех отношений строгого частичного порядка,* т. е. транзитивных отношений строгого предпочтения.

Пример 3.5. Пусть $R \in \mathcal{QT}$ есть отношение из примера 3.1. Тогда отношение $P = \overline{R^{-1}}$ принадлежит пространству $\mathcal{PO}$ и имеет матрицу

		1

Очевидно, что P есть частичный порядок.

$\mathcal{QS}$ — *пространство всех квазисерий,* т. е. строгих частичных порядков P таких, что $I = \overline{P \cup P^{-1}}$ — эквивалентность.

Пример 3.6. Пусть $R \in \mathcal{QO}$ есть отношение из примера 3.2. Тогда отношение $P = \overline{R^{-1}}$ принадлежит пространству $\mathcal{QS}$ и имеет матрицу

		1
		1

$\mathcal{TSO}$ — *пространство всех совершенных строгих порядков.*

Пример 3.7. Пусть $R \in \mathcal{TO}$ есть, отношение из примера 3.3. Тогда отношение $P = \overline{R^{-1}}$ принадлежит пространству $\mathcal{TSO}$ и имеет матрицу

	1	1
		1

Связь введенных пространств изображена на диаграмме 3.3, где стрелки указывают вложение пространств.

$$\mathcal{TYO} \xrightarrow{\varphi} \mathcal{QY} \xrightarrow{\varphi} \mathcal{PO} \xrightarrow{\varphi} \mathcal{YP}$$

Диаграмма 3.3

Пример 3.8. Для *R* из примера 3.4 отношение $P = \overline{R^{-1}}$ с матрицей

	1	
		1

служит примером отношения из пространства $\mathcal{YP}$, которое не принадлежит ни одному из вложенных в него пространств из диаграммы 3.3.

Наконец, в соответствии с диаграммой 3.1, определим пространства отношений безразличия.

$\mathcal{T}$ — *пространство всех отношений безразличия,* т. е. симметричных и рефлексивных отношений. Такие отношения называются *отношениями толерантности*. Поэтому мы в дальнейшем будем называть это пространство *пространством толерантностей.*

$\mathcal{TOT}$ — *пространство всех транзитивно ориентируемых отношений толерантности,* т. е. таких отношений толерантности *T*, что дополнение к *T* представляется в виде объединения взаимно обратных транзитивных отношений.

Пример 3.9. Для отношения *R* из примера 3.1 отношение $I = R \cap R^{-1}$ имеет матрицу

1	1	
1	1	1
	1	1

Дополнение к этому отношению имеет матрицу

		1
1		

и представляется в виде объединения взаимнообратных отношений P и P^{-1} с матрицами

1		

		1

соответственно.

$\mathcal{I}$ — *пространство всех отношений эквивалентности.*

Пример 3.10. Для отношения R из примера 3.2 отношение I имеет матрицу

1	1	
1	1	
		1

Очевидно, что I есть эквивалентность.

$\mathcal{E}$ — *пространство всех отношений равенства.* Очевидно, что оно состоит из одной точки, матрица отношений которой есть диагональная матрица.

Так же как и в предыдущих случаях укажем связь введенных пространств на одной диаграмме

$$\mathcal{E} \xrightarrow{\varphi} \mathcal{I} \xrightarrow{\varphi} \mathcal{T}O\mathcal{T} \xrightarrow{\varphi} \mathcal{T}$$

Диаграмма 3.4

Пример 3.11. Подобно тому, как это было сделано в примерах 3.4 и 3.8, приведем пример отношения из пространства $\mathcal{T}$, не принадлежащего ни одному из вложенных в него пространств на диаграмме 3.4. Бели мы возьмем для этой цели отношение R из примера 3.4, то получим отношение I с матрицей

1		1
	1	
1		1

Легко видеть, что отношение с такой матрицей является эквивалентностью и, следовательно, принадлежит пространству $\mathcal{J}$. Таким образом, полученное отношение, против ожиданий, не является примером отношения, принадлежащего пространству $\mathcal{T}$ и не принадлежащего вложенным в него пространствам.
В данном случае этот факт объясняется тем, что отображение α (диаграмма 3.1) является сюръективным, а не биективным отображением. Другими словами, прообраз точки из пространства $\mathcal{T}$ в пространстве $\mathcal{P}$ состоит, вообще говоря, из нескольких точек. В нашем примере в этот прообраз, наряду с отношением R из примера 3.4 входит, скажем, отношение R^7 с матрицей

1	1	1
	1	
1	1	1

А это отношение принадлежит пространству $\mathcal{QO}$, что и объясняет принадлежность отношения I пространству $\mathcal{J}$.
Искомый пример можно привести лишь в случае, когда мощность n носителя **A** не менее 5. В случае $n = 5$ таким примером может служить отношение I с матрицей

1		1	1	
	1		1	1
1		1		1
1	1		1	
	1	1		1

Подобные отношения, входящие среди прочих в пространство $\mathcal{T}$, называются *транзитивно неориентируемыми.* В терминах теории графов транзитивно ориентируемые и неориентируемые отношения изучались в работе Гилмора и Гофмана.

Диаграммы 3.1—3.4 можно объединить в одну трехмерную диаграмму 3.5. Вертикальные отображения на этой диаграмме являются вложениями пространств, а типы горизонтальных отображений определены в 15.2.

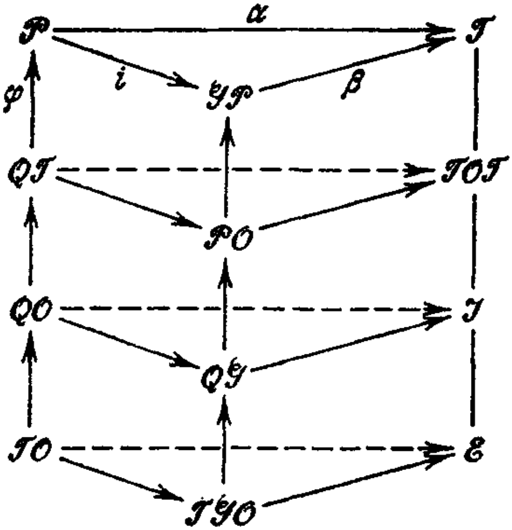


Диаграмма 3. 5

Покажем теперь, что диаграмма 3.5 коммутативна. Это означает, что любые два пути, идущие в направлении стрелок из одного пространства в другое, определят одно и то же отображение этих пространств.

Лемма 3.1. *Любой квадрат отображений на диаграмме* 3.5 *коммутативен.*

Доказательство. Утверждение леммы непосредственно следует из того, что вертикальные стрелки суть вложения, определяемые одинаковыми формулами.

Лемма 3.2. *Любой треугольник отображения на диаграмме* 3.5 *коммутативен.*

Доказательство непосредственно следует из утверждения 3.2.

Теорема 3.1, *Диаграмма* 3.5 *коммутативна.*

Доказательство. Нам надо показать, что любые два пути, идущие из одного пространства в другое, задают одно и то же отображение. Для примера рассмотрим следующие два пути:

$$\mathcal{TO} \to \mathcal{QO} \to \mathcal{QT} \to \mathcal{TOT} \text{ и } \mathcal{TO} \to \mathcal{TPO} \to \mathcal{E} \to \mathcal{I} \to \mathcal{TOT}.$$

Последний путь в силу леммы 3.2 задает то же отображение, что и путь

$$\mathcal{TO} \to \mathcal{E} \to \mathcal{I} \to \mathcal{TOT}.$$

Последний путь в силу леммы 3.1 задает то же отображение, что и путь

$$\mathcal{TO} \to \mathcal{QO} \to \mathcal{T} \to \mathcal{TOT}.$$

В силу той же леммы 3.1 последний путь задает то же отображение, что и первый путь $\mathcal{TO} \to \mathcal{QO} \to \mathcal{QT} \to \mathcal{TOT}$.

Сводная схема из двенадцати пространств (диаграмма 3.5) содержит пространства многих отношений, используемых на практике и исследуемых в теории полезности и группового выбора. Четыре пространства из этой системы изучались в рамках метрического подхода. Так впервые на аксиоматической основе в пространстве $\mathcal{QO}$ была введена и использована концепция расстояния между ранжированиями для построения ранжирования, «согласованного» с данными. Аналогичный подход был развит для расстояния в пространстве эквивалентностей $\mathcal{I}$, а так же — для пространства квазисерий $\mathcal{QP}$ (изоморфного пространству $\mathcal{QO}$). В пространстве строгих частичных порядков $\mathcal{P}$ расстояние аксиоматически было введено в ряд работ. Проблема группового выбора в этих пространствах всеми авторами ставилась так, как она была сформулирована в главе 13 для первого уровня общности.

Таким образом, в метрическом подходе изученными оказались все пространства второго (снизу) «этажа» диаграммы 3.5 и одно пространство — с третьего. Интересным фактором представляется появление в этой схеме двух не изучавшихся ранее пространств $\mathcal{OT}$ и $\mathcal{TOT}$ третьего «этажа». Точки пространства квазитранзитивных отношений $\mathcal{QT}$ в общем случае представляют собой отношения менее «жесткие» в части требований транзитивности, чем отношения из $\mathcal{QO}$, и имеют содержательный, эмпирический эквивалент в теории экономического поведения. Более подробно пространство $\mathcal{Q}\mathcal{T}$ будет изучаться в главе 16. Пространство $\mathcal{T}\mathcal{O}\mathcal{T}$ содержит такие отношения толерантности из $\mathcal{T}$, которые совпадают с отношением $\overline{P \cup P^{-1}}$, где P — квази-транзитивное отношение. Другими словами, отношения из $\mathcal{T}\mathcal{O}\mathcal{T}$ суть отношения неразличимости для отношений из $\mathcal{Q}\mathcal{T}$. В данной работе это пространство не изучается.

Самостоятельной и интересной представляется задача расширения этой схемы или за счет введения в нее других, часто используемых в приложениях и теоретических исследованиях отношений, или за счет пространств с несколькими отношениями, например, с отношением древесного порядка и лексикографического порядка и т. п.

Таким образом, изображенная на диаграмме 3.5 система пространств бинарных отношений в наглядной форме представляет мир таких пространств. Она позволяет указать место как уже изученных в метрическом подходе пространств, так и тех, которые будут рассматриваться в данной работе. Выявленные связи между пространствами будут использованы для переноса постановок задач и методов их решения из одних пространств в другие.

16. Геометрические структуры пространств бинарных отношений

При обсуждении основных положений Геометрического подхода в главе 13 мы уже отмечали, что наиболее распространенным является подход, при котором для решения задачи группового выбора в пространстве отношений данного типа вводится метрическая структура. Она вводится на основе общих сложившихся геометрических представлений и концепций. В основе вводимых расстояний лежит система аксиом, одна из которых вводит важное геометрическое понятие «между». Это понятие является единственным общим понятием для метрического и разрабатываемого здесь геометрического подхода.

Понятие «между» лежит в основе теоретических построений в геометрическом подходе и в нашем контексте несет естественную смысловую нагрузку, например, в утверждениях типа «одно отношение предпочтения лежит между двумя другими» или в заданиях вида «найти отношение предпочтения, лежащее между данными». Эти высказывания не содержат в себе ничего необычного.

Ситуация, стоящая за ними, чрезвычайно жизненна и с давних времен привлекает к себе внимание. Для примера укажем на, по-видимому, первую и простейшую как по форме, так и по содержанию постановку задачи принятия решения, сформулированную (как принято считать), в первой половине XIV века французским философом Ж. Буриданом в известной притче об осле: осел, очутившись между двумя совершенно одинаковыми охапками сена не мог ни одну из них предпочесть другой, т. е. решить, какую из них выбрать, и околел от голода. Несмотря на, казалось бы, шутливый характер этого примера, можно сказать, что изложенная в нем в аллегорической форме ситуация выбора с незначительными вариациями составляет основу доброй половины задач, решаемых с помощью экспертов. Другая половина представляет собой «зеркальное отражение» этой ситуации в том смысле, что организаторы экспертиз хотят попасть в положение героя притчи, т. е. найти решение, которое лежало бы одновременно между всеми суждениями экспертов. (Пожалуй, и в наше время привлечение большинства современных научных методов для решения задачи «буриданова осла» вряд ли помогло бы последнему избежать летального исхода, исключая разве что только голосование при нечетном числе экспертов и запрете воздерживаться от голосования.)

В обычной евклидовой геометрии мы имеем представление о расположении одной точки пространства между двумя другими: точка *А* лежит между точками *В* и *С,* если она лежит на отрезке прямой, соединяющей *В* и С, при этом сумма расстояний от точйи *А* до точек *В* и *С,* равна расстоянию между *В* и С. В терминах понятия расстояния между отношениями тот факт, что отношение *А* находится между отношениями *В* и *С* характеризуется аналогично. Различие в природе евклидова пространства и пространства отношений в последнем случае проявляется в определении понятия «линейный сегмент», заменяющего для пространств отношений понятие «отрезок прямой».

Наш особый интерес к понятию «между» связан с тем, что при согласовании индивидуальных предпочтений групповое решение естественно искать среди множества всех тех предпочтений, которые расположены «в середине между» исходными предпочтениями. Такие, лежащие между исходными, множества предпочтений названы в данной работе *выпуклыми множествами.*

В этой главе теория выпуклых множеств строится для произвольных пространств бинарных отношений. В дальнейшем мы не будем различать точки пространства и соответствующие им бинарные отношения. Всюду в этой главе рассматривается фиксированное пространство R. Если это особо не оговорено, то мы считаем, что все рассматриваемые точки принадлежат этому пространству.

16.1. Отношение «между»

Начнем изучение структур пространств предпочтений с введения двух определений понятия «между».

Определение 4.1. Отношение R лежит между отношениями R_1 и R_2, если

$$R_1 \cap R_2 \subseteq R \subseteq R_1 \cup R_2.$$

То обстоятельство, что отношение R лежит между отношениями R_1 и R_2 будет записываться так: $R \in [R_1, R_2]$.

Первое определение понятия «между» для случая трех отношений R, R_1 и R_2 содержательно означает, что отношение R, если оно лежит между отношениями R_1 и R_2, должно содержать то общее, что есть в отношениях R_1 и R_2 (т. е. содержать в себе пересечение этих отношений), и само содержаться в отношении, аккумулирующем R_1 и R_2. Эта интерпретация становится совершенно прозаической, если в ней всюду слово «отношение» заменить на одно из синономичных в данном контексте слов: «суждение», «мнение», «высказывание».

Введенное определение «между» допускает естественное обобщение на случай произвольного числа отношений.

Определение 4.2. Отношение R лежит между отношениями $R_1, R_2, .., R_k$, если

$$\bigcap_i R_i \subseteq R \subseteq \bigcup R_i.$$

Запись $R \in [R_1, ..., R_k]$ будет обозначать, что точка R лежит между точками $R_1, R_2, ..., R_k$.

Понятие «между» определяет некоторую геометрическую структуру пространства R. Наряду с этой структурой в пространстве R имеется структура частично упорядоченного множества. Именно, предпочтение R_1 предшествует (нестрого) предпочтению R_2, если $R_1 \subseteq R_2$. Это отношение частичного порядка на R индуцирует отношение частичного порядка на любом подмножестве в R. В дальнейшем, говоря о максимальных и минимальных элементах различных подмножеств в пространстве R, мы будем иметь в виду минимальные и максимальные элементы этих подмножеств относительно этого порядка. Структура «между» и структура порядка на R согласованы в том смысле, что из $R_1 \subseteq R \subseteq R_2$ следует, что $R \in [R_1, R_2]$.

Докажем вспомогательное утверждение, устанавливающее связь двух определений «между». Пусть $R_1, R_2, ..., R_k$ — точки пространства R.

Лемма 4.1. *Пусть* $Q_j \in [R_1 ..., R_k]$ *для* $j = 1,2, …, m$ *и* $R \in [Q_1 ,..., Q_m]$. *Тогда* $R \in [R_1, ..., R_k]$.

Доказательство. Согласно определению 4.2 имеем

$$\bigcap_{i=1}^{k} R_i \subseteq Q_1 \subseteq \bigcup_{i=1}^{k} R_i,$$

$$\cdots\cdots\cdots$$

$$\bigcap_{i=1}^{k} R_i \subseteq Q_m \subseteq \bigcup_{i=1}^{k} R_i.$$

Согласно определению 4.1

$$Q_1 \cap \ldots \cap Q_m \subseteq R \subseteq Q_1 \cup \ldots \cup Q_m.$$

Из выписанных включений следует

$$\bigcap_{i=1}^{k} R_i \subseteq Q_1 \cap \ldots \cap Q_m \subseteq R \subseteq Q_1 \cup \ldots \cup Q_m \subseteq \bigcup_{i=1}^{k} R_i,$$

откуда $R \in [R_1, ..., R_k]$, что и требовалось доказать.

Следствие. Пусть R' и R'' лежат между R_1 и R_2. Тогда любое R, лежащее между R' и R'', лежит также и между R_1 и R_2.

Пусть R' и R'' — две различные точки в пространстве R.

Определение 4.3. *Линейным сегментом* между R^f и R'' назовем последовательность различных точек R_u ..., R_k такую, что

1. $R_1 = R', \quad R_k = R''$;
2. $R_i \in [R_m, R_l]$ для $m \leqslant i \leqslant l$;
3. из $R \in [R_i, R_{i+1}]$ следует, что либо $R = R_i$, либо $R = R_{i+1}$.

Две различные точки R' и R'' назовем *соседними,* если они сами бразуют линейный сегмент. Очевидно, что между соседними точками не лежит ни одна точка, отличная от них. Линейный сегмент является последовательностью соседних точек, лежащих «между» двумя данными и «соединяющей» их.

Теорема 4.1. *В произвольном пространстве бинарных отношений* R *для любой пары различных точек существует линейный сегмент между ними.*

Доказательству теоремы предпошлем следующую лемму.

Лемма 4.2. *Пусть R' и R'' —различные точки в пространстве* R. *Существует точка, лежащая между R' и R'' и соседняя к R'.*

Доказательство. Пусть $R \in [R', R'']$ и $R \neq R''$. Согласно следствию из леммы 4.1 множество всех точек, лежащих между *R'* и *R*, содержится в множестве точек, лежащих между *R'* и *R''* .

Покажем, что $R'' \notin [R', R]$. Предположим противное. Тогда справедливы включения

$$R' \cap R'' \subseteq R \subseteq R' \cup R''$$

и

$$R' \cap R \subseteq R \subseteq R' \cup R.$$

Из этих включений имеем

$$R = (R \cap R') \cup (R \cap R'') \subseteq R'' \cup (R \cap R'') = R''$$

и

$$R'' = (R'' \cap R') \cup (R'' \cap R) \subseteq R \cup (R'' \cap R) = R.$$

Отсюда *R* = *R''*, что противоречит условию *R≠ R''*. Итак, множество точек, лежащих между *R'* и *R,* строго содержится в множестве точек, лежащих между *R'* и *R ''*. Так как множество точек, лежащих между *R'* и *R'',* конечно, то отсюда непосредственно следует утверждение леммы.

Доказательство теоремы 4.1. Пусть *R'* и *R''* — точки пространства R. Положим R_1= *R'* и определим R_2 как соседнюю к R_1 точку, лежащую между *R'* и *R'* . Как показано в лемме 4.2, такая точка существует. Если R_2= *R'',* то теорема доказана. В противном случае мы применяем предыдущее рассуждение к точкам R_2 и *R''*. По индукции, пусть R_{i+1} есть соседняя точка к R_i, лежащая между R_i и *R''*. Легко видеть, что на некотором шаге мы получим R_k = *R''*. Согласно следствию из леммы 4.1, все построенные точки лежат между *R'* и *R''*.

Покажем, что построенная последовательность точек задает линейный сегмент между *R'* и *R''*. Очевидно, достаточно проверить выполнение условия 2 из определения 4.3. Пусть R_m, R_i и R_l такие, что $m \le i \le l$. Из построения последовательности точек $\{R_i\}$ следует, что $R_i \in [R_m, R'']$, $R_l \in [R_i, R'']$, т. е. справедливы включения

$$R_m \cap R'' \subseteq R_i \subseteq R_m \cup R'' \quad \text{и} \quad R_i \cap R'' \subseteq R_l \subseteq R_i \cup R''.$$

Используя эти включения, получаем

$$R_m \cap R_l \subseteq R_m \cap (R_i \cup R'') =$$
$$= (R_m \cap R_i) \cup (R_m \cap R'') \subseteq (R_m \cap R_i) \cup R_i = R_i$$

и

$$R_m \cup R_l \supseteq R_m \cup (R_i \cap R'') =$$

$$= (R_m \cup R_i) \cap (R_m \cup R'') \supseteq (R_m \cup R_i) \cap R_i = R_i,$$

откуда следует, что $R_i \in [R_m, R_l]$, что и требовалось доказать.

Введенное в этом параграфе понятие линейного сегмента является аналогом привычного геометрического понятия отрезка прямой, соединяющей две заданные точки. Существенное отличив состоит в том, что, вообще говоря, существуют различные линейные сегменты, соединяющие две заданные точки. Это обстоятельство накладывает специфический оттенок на интерпретацию геометрических построений в пространствах бинарных отношений.

16.2. Выпуклые множества и выпуклые оболочки

Наличие понятия «между» в пространстве предпочтений позволяет ввести естественное понятие выпуклого множества. Соответственно тому, что у нас имеется два определения «между», мы дадим два определения выпуклости.

Определение 4.4. Множество **X** точек пространства **R** называется *выпуклым,* если вместе с любыми двумя точками оно содержит любую точку, лежащую между ними.

Используя наши обозначения, можно сказать, что множество X выпукло, если из $R', R'' \in X$ и $R \in [R', R'']$ следует, что $R \in X$.

Определение 4.5. Множество $\mathbf{X}=\{R_1, \ldots, R_k\}$ точек пространства **R** называется *выпуклым,* если из $R \in [R_1, \ldots, R_k]$ следует, что $R \in X$.

Другими славами, множество отношений выпукло, если любое отношение, лежащее одновременно между всеми отношениями из этого множества, принадлежит этому же множеству.

Исследуем связь между двумя определениями выпуклости. Следующее утверждение непосредственно вытекает из леммы 4.1.

Лемма 4.3. *Из выпуклости в смысле определения* 4.5 *следует выпуклость в смысле определения* 4.4 *для любого пространства бинарных отношений* R.

До сих пор на пространство **R** не накладывалось никаких ограничений. В дальнейшем мы будем в основном рассматривать пространства бинарных отношений, для которых выполнено следующее условие.

Условие полноты. *Для любых двух соседних точек R' и R'' пространства* **R** *симметрическая разность $R'\Delta R''$ есть одноэлементное множество.*

Другими словами, соседние отношения различаются на одной паре элементов множества *A*. Пространства, удовлетворяющие условию полноты, будем называть *полными пространствами.*

Если вспомнить, что в пространстве отношений предпочтения может быть определена метрическая структура, то полные пространства характеризуются тем, что в этих пространствах соседние точки отстоят друг от друга на минимальное единичное расстояние. Другими словами, полные пространства плотно, без «дыр» заполнены точками этого пространства.

Для полных пространств существует более глубокая связь между двумя определениями выпуклости, чем установленная в общем случае в лемме 4.3.

Лемма 4.4. *Для полного пространства из выпуклости в смысле определения* 4.4 *следует выпуклость в смысле определения* 4.5.

Доказательство. Пусть $\mathbf{X} = \{R_1, ..., R_k\}$ — множество, выпуклое в смысле определения 4.4, т. е. вместе с любыми двумя точками из X содержит все точки, лежащие между ними. Пусть $R\in [R_1, ..., R_k]$, Рассмотрим линейный сегмент между R_1 и R (выбор точки R_1 произволен). Предположим, что $R\notin \mathbf{X}$. Поскольку $R_1\in X$, то в линейном сегменте между R_1 и R найдутся две последовательные точки R' и R'' такие, что $R'\in \mathbf{X}$, а $R''\notin X$. Так как точки R' и R'' — соседние, то возможны два случая: $R''\backslash R' = \{(a, b)\}$ или $R'\backslash R'' = \{(a, b)\}$ для некоторой пары (a, b). Рассмотрим эти случаи отдельно.

1. $R'' = R' \cup \{(a, b)\}$, $(a, b)\notin R'$. Так как $R \in [R', R]$ по определению линейного сегмента, то $R'' \subseteq R' \cup R$, откуда $(a, b) \in R$. Так как $R \in [R_1, ..., R_k]$, то $R \subseteq \bigcup_{i=1}^{k} R_i$. Отсюда $(a, b) \in R_i$ для некоторого i Покажем, что $R''\in[R', R_i]$. Действительно, $R' \cap R_i \subseteq R''$, так как $R'\subset R''$. Далее, $R'' \subseteq R' \cup R_i$, так как $R'' = R' \cup \{(a, b)\}$ и $(a, b) \in R_i$. Итак, $R'' \in [R', R_i]$. Но $R' \in \mathbf{X}$, $R_i \in \mathbf{X}$, а $R'' \notin \mathbf{X}$, что противоречит выпуклости.

2. $R' = R'' \cup \{(a, b)\}$, $(a, b)\notin R''$. Так как $R''\in[R', R_i]$ по определению линейного сегмента, то $R' \cap R \subseteq R''$, откуда $(a, b) \notin R$. Так как $R \in [R_1, ..., R_k]$, то $\bigcap_{i=1}^{k} R_i \subseteq R$. Отсюда найдется номер i такой, что $(a, b) \notin R_i$. Покажем, что $R'' \in [R', R_i]$. Дей-

ствительно, $R' \cap R \subseteq R''$, так как $R'=R'' \cup \{(a, b)\}$ и $(a, b) \notin R$. Далее, $R'' \subseteq R' \cup R_i$, так как $R'' \subset R$. Итак, $R'' \in [R', R_i]$.

Но $R' \in \mathbf{X}$, $R_i \in \mathbf{X}$, а $R'' \notin \mathbf{X}$, что противоречит выпуклости **X**. Полученные противоречия показывают, что $R \in \mathbf{X}$, что и требовалось доказать.

Из доказанных лемм непосредственно следует, что справедлива

Теорема 4.2. *Для полного пространства* R *оба определения выпуклости эквивалентны.*

Доказанные теоремы позволяют исследовать в полных пространствах аналог геометрического понятия выпуклой оболочки множества.

Пусть теперь **X** — произвольное множество в пространстве бинарных отношений R.

Определение 4.6. *Выпуклой оболочкой множества* **X** *в пространстве* R называется наименьшее выпуклое множество, содержащее **X** (понимая выпуклость в смысле определений 4.4 или 4.5).

Выпуклые оболочки множества **X** будем обозначать через **C(X)** и **П(X)** соответственно определениям выпуклости 4.4 и 4.5. Так как само пространство R — выпуклое множество и содержит **X**, а пересечение выпуклых множеств, как нетрудно видеть,— выпуклое множество, то выпуклая оболочка существует для любого множества **X**. Легко проверить, что она определяется единственным образом и совпадает с пересечением всех выпуклых множеств, содержащих **X**.

Наличие двух определений выпуклости и соответственно — двух вариантов выпуклой оболочки позволяет дать два способа построения выпуклой оболочки. Используем сначала определение **C(X)**. Для данного множества **X** обозначим через **X'** множество, полученное добавлением к **X** всех точек, лежащих между парами точек из **X**. Введем обозначения: $\mathbf{X}_0 = \mathbf{X}$, $\mathbf{X}_1 = \mathbf{X}'$, $\mathbf{X}_2 = (\mathbf{X}')'$, ..., $\mathbf{X}^m = (\mathbf{X}^{m-1})'$. Очевидно, что $\mathbf{X}_i \subseteq \mathbf{X}_{i+1}$. Так как число точек в пространстве R конечно, то последовательность вложенных множеств $\mathbf{X}_i$ стабилизируется, т. е. найдется номер N такой, что $\mathbf{X}_{N-1} \neq \mathbf{X}_N$, а $\mathbf{X}_N = \mathbf{X}_{N+1} = \mathbf{X}_{N+2} = \ldots$

Лемма 4.5. *В предыдущих обозначениях* $\mathbf{X}_N = \mathbf{C}(\mathbf{X})$ *для любого пространства* R.

Доказательство. Пусть $R', R'' \in \mathbf{X}_N$ и $R \in [R', R'']$. Имеем $R \in \mathbf{X}_{N+1}$. Но $\mathbf{X}_{N+1} = \mathbf{X}_N$ по определению N. Итак, $\mathbf{X}_N$ — выпуклое множество. Пусть **Y** — выпуклое в смысле определения 4.4 множество и $\mathbf{X} \subset \mathbf{Y}$, Очевидно, что $\mathbf{X}_i \subseteq \mathbf{Y}$ для всех i. В частности,

$\mathbf{X}_N \subseteq \mathbf{Y}$, откуда следует, что $\mathbf{X}_N$ — минимальное выпуклое множество, содержащее **X**.

Выпуклую оболочку можно построить также, исходя из определения **П**(**X**). Для заданного множества **X** == {R_1, ..., R_k} определим множество $\tilde{\mathbf{X}}$ всех R таких, что $R \in [R_1,...,R_k]$.

Лемма 4.6. *Для полного пространства* $\tilde{\mathbf{X}} = \mathbf{\Pi}(\mathbf{X})$.

Доказательство. Согласно лемме 4.1 $\tilde{\mathbf{X}}$—выпуклое множество в смысле определения 4.5. С другой стороны, если $\mathbf{Y} \supseteq \mathbf{X}$ и **Y** — выпуклое множество, то $\tilde{\mathbf{X}} \subseteq \mathbf{Y}$ в силу определения 4.5. Отсюда следует утверждение леммы.

Лемма 4.7. *Пусть* **X** *— подмножество в произвольном пространстве* R. *Тогда* $\mathbf{C}(\mathbf{X}) \subseteq \mathbf{\Pi}(\mathbf{X})$

Доказательство. Следует непосредственно из следствия к лемме 4.1, леммы 4.5 и леммы 4.6.

Теорема 4.3. *Для произвольного подмножества* **X** *в полном пространстве бинарных отношений*

$$\mathbf{C}(\mathbf{X}) = \mathbf{\Pi}(\mathbf{X}).$$

Доказательство. Как следует из леммы 4.6, достаточно показать, что $\mathbf{C}(\mathbf{X}) \supseteq \mathbf{\Pi}(\mathbf{X})$. Так как $\mathbf{X} \subset \mathbf{C}(\mathbf{X})$, то в силу определения выпуклой оболочки $\mathbf{\Pi}(\mathbf{X})$ имеем $\mathbf{\Pi}(\mathbf{X}) \subseteq \mathbf{\Pi}(\mathbf{C}(\mathbf{X}))$. Но по теореме 4.2 для полного пространства бинарных отношений $\mathbf{\Pi}(\mathbf{C}(\mathbf{X})) = \mathbf{C}(\mathbf{X})$, откуда и следует доказываемое включение. Теорема доказана.

Результаты этой главы, сформулированные здесь во всей общности для произвольных пространств бинарных отношений, применимы, конечно, и к пространствам предпочтений и безразличия.

16.3. Выпуклые оболочки и проблема группового выбора

В предыдущем параграфе мы ввели в рассмотрение класс пространств бинарных отношений — полные пространства. Для полных пространств оказалось, что два различных варианта определения выпуклости дали один и тот же результат, точнее, соответствующие этим определениям выпуклые множества отношений совпадают. Этот факт в свете проблем группового выбора представляет для нас определенный интерес. Если первое определение выпуклой оболочки, базирующееся на первом определении понятия «между» и первом определении выпуклости, имеет ярко выраженную геометрическую основу и является полным аналогом соответствующего понятия в евклидовой геометрии, то второе определение связано с важным понятием в теории группового выбора. Именно, оно представляет собой не что иное, как формализацию хорошо известного условия Парето на принцип согласования отношений индивидуального предпочтения: *«если все индивидуумы предпочитают объект а объекту b, то и в групповом отношении объект а должен быть предпочтительнее b. Точно так же, если все члены группы безразличны в выборе между а и b, таково же должно быть групповое решение»*.

Построенное групповое множество как раз и состоит из отношений, удовлетворяющих этому условию, и поэтому выбор единственного группового решения естественно производить из отношений, составляющих это множество. Однако не все точки выпуклой оболочки «равноправны» в том смысле, что некоторые из них расположены ближе к одним исходным точкам (мнениям), чем к другим. Кроме того, число отношений, составляющих выпуклую оболочку исходного множества отношений, в практических приложениях может оказаться столь велико, что задача выбора окончательного решения на первом уровне будет трудноразрешимой. Поэтому здесь открываются широкие возможности для создания различных способов сокращения числа или направленного отбора отношений, из которых затем будет выбрано единственное групповое решение.

В следующих главах задача направленного формирования множества допустимых групповых решений будет решена для трех конкретных пространств. Сейчас мы проиллюстрируем общую идею решения такой задачи. В конкретных задачах в каждом выпуклом множестве можно выделить подмножество точек, с геометрической точки зрения расположенных однородно относительно исходных точек, порождающих выпуклую оболочку. Такое выделенное подмножество в дальнейшем будет называться ядром выпуклого множества. Ядро имеет простую геометрическую структуру: оно представляет собой выпуклое множество всех точек, лежащих между двумя точками, которые в дальнейшем будут обозначаться $P_{\min}$ и $P_{\max}$.

Ни одна из точек, входящих в ядро, не «тяготеет» к какому-либо одному из исходных суждений. В этом смысле все точки ядра «равноправны». Очевидно, что они также удовлетворяют условию Парето. Исходя из этих свойств точки ядра признаются в геометрическом подходе допустимыми для поиска среди них групповых решений.

17. Теория выпуклых множеств в пространствах частичных порядков и квазитранзитивных отношений

Результаты, полученные ранее для произвольных пространств бинарных отношений, в этой главе будут использованы в приложении к двум конкретным пространствам отношений предпочтения: частичного порядка $\mathcal{PO}$ и квазитранзитивных отношений предпочтения $\mathcal{QT}$. Для этих пространств удается более глубоко исследовать структуру выпуклых множеств и построить алгоритмы нахождения множества допустимых групповых решений.

17.1. Выпуклые множества в пространстве $\mathcal{PO}$

Прежде всего мы установим полноту пространства $\mathcal{PO}$. Напомним, что точками этого пространства служат отношения частичного порядка на некотором конечном множестве **A**, т. е. антирефлексивные и транзитивные бинарные отношения на **A**. Отношение частичного порядка всегда обладает свойством антисимметричности. Очевидно, что пересечение $P \cap Q$ двух отношений частичного порядка P и Q снова является отношением частичного порядка. Напротив, их объединение $P \cup Q$, вообще говоря, уже не является отношением частичного порядка.

Для доказательства полноты необходимо проверить, что любые соседние точки пространства $\mathcal{PO}$ различаются лишь на одноэлементном множестве. Пусть P и Q — две соседние точки в $\mathcal{PO}$. Так как $P \cap Q \in \mathcal{PO}$, и, очевидно, $P \cap Q \in [P,\ Q]$, то либо $P \cap Q = P$, либо $P \cap Q = Q$. Таким образом, из двух соседних отношений в $\mathcal{PO}$ одно обязательно содержится в другом. Пусть, например, $P \subset Q$. Справедлива следующая

Лемма 5.1. *Если* $P \subset Q$ *— две соседние точки в пространстве* $\mathcal{PO}$, *то* $Q = P \cup \{(a,\ b)\}$ *для некоторой пары* $(a,\ b) \in A \times A$.

Доказательство. Согласно известной теореме Шпильрайна элементы множества A могут быть занумерованы так, что из $(a_i, a_j) \in Q$ всегда следует, что $i < j$. Рассмотрим пары (a_i, a_j), принадлежащие Q и не принадлежащие P: $(a_i,\ a_j) \in Q \setminus P$. Пусть i_0 — наибольший из индексов i у таких пар. Положим $a = a_{i0}$. Среди индексов j таких,

что $(a, a_j) \in Q\backslash P$, выберем наименьший j_0. Положим $b=a_{j0}$. Очевидно, $(a, b) \in Q\backslash P$.
Определим отношение R на A следующим образом: R —■ = $Q\backslash\{(a, b)\}$. Покажем, что R — частичный порядок. Очевидно, достаточно установить транзитивность 7?. Пусть (#, y) ^ R, (y, z) ∈ Д. Докажем, что и (ж, z) ∈ Д. Если

$$(x, z) \neq (a, b),$$

то (a:, z) ^ Л, так как Q транзитивно. Предположим, что

$$(x, z) = (a, b).$$

Тогда $(a, y) \in Q$ и $(y, b) \in Q.$
Пусть $y — a_к$. Тогда имеем
$i_o<k$ и $к<]_0$. Так как $k<j_0$, то
$(a, y) = (a, a_k) \notin Q\backslash P$ по
определению $7*_0$. Следовательно, (a, y) ^ P. Так как i_0 < fc, то
$(y, b) = (a_k, a_{j0}) \notin Q\backslash P$ по определению
i_0. Следовательно, (г/, fc)s е P. Отсюда в силу транзитивности P имеем (a, b) ∈ P_t что противоречит тому, что (a, b) ∈е ^\P. Полученное противоречие завершает доказательство.
Для случая небольших кардинальных чисел множества A (лг = 2, 3) пространство $\mathscr{PO}$ можно изобразить графически в виде обычной диаграммы Хассе. Опуская тривиальный случай n = 2, мы приведем эту диаграмму для случая, когда A состоит из трех элементов: A = = (a, Ь, c}.

На рис. 5.1 стрелки указывают вложение

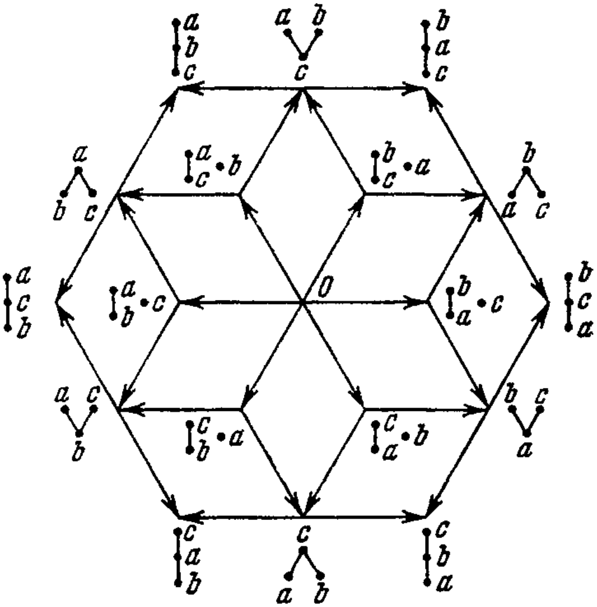


Рис. 5.1.

частичных порядков, рассматриваемых как подмножества в АХ А.

Буквой *O* обозначен тривиальный частичный порядок, совпадающий с пустым подмножеством. Рис. 5.1 является хорошей иллюстрацией к лемме 5.1: соседние точки на нем — это в точности те* которые соединены стрелками.

Из леммы 5.1 непосредственно следует, что справедливаТеорема 5.1.

Пространство $\overrightarrow{\mathcal{PO}}$ *является полным пространством.*

Таким образом, для пространства $\overrightarrow{\mathcal{PO}}$ справедливы все результаты, полученные для полных пространств бинарных отношений в предыдущей главе. В частности, два определения выпуклости эквивалентны в этом пространстве, и для любого множества $\mathbf{X} \subset \overrightarrow{\mathcal{PO}}$ выпуклая оболочка **C(X)** совпадает с **П(X)**.

Изучим теперь подробнее структуру выпуклых множеств в пространстве $\overrightarrow{\mathcal{PO}}$. Пусть **X** — выпуклое множество в пространстве $\bar{\mathcal{P}\mathcal{O}}$. Тогда на X также определено отношение частичного порядка из $\mathcal{P}\mathcal{O}$.

Опишем подробнее структуру частичного порядка на **X**.

Лемма 5.2. *В выпуклом множестве* **X** *имеется единственный минимальный элемент.*

Доказательство. Пусть P и *Q* •— два различных минимальных элемента из **X**. Частичный порядок *R* = *P* Г) *Q* лежит между *P* и *Q* и, следовательно, Й^X. С другой стороны, *R* содержится в *P* и *Q,* но не совпадает с ними, что противоречит минимальности *P* и *Q*.

Множество **X** является выпуклой оболочкой множества, состоящего из минимального элемента и всех максимальных элементов, так как любой элемент множества **X** лежит между минимальным и некоторым максимальным элементом. Максимальных элементов в множестве **X** может быть несколько. Таким образом, выпуклая оболочка множества **X** полностью определяется минимальными и максимальными элементами. Мы будем говорить, что выпуклое множество **X** *порождается* подмножеством **X'**, если **C(X')=X**. В рассматриваемом случае выпуклое множество **X** порождается подмножеством, состоящим из минимального и всех максимальных элементов.

Рассмотрим структуру выпуклого множества," обладающего единственным максимальным элементом. В этом случае, как следует из изложенного, все точки множества X лежат между минимальным и единственным максимальным элементами. Обозначим через *Pmm* и Pтах минимальный и максимальный элементы множества X соответственно. Очевидно, что P_{min} ^ P_{max}. Обозначим через *d* мощность разности отношений $P_{тах}$ и Pтт, т. е.

$d = |P_{max} \setminus P_{min}|$. Рассмотрим структуру множества всех отношений (не обязательно частичных порядков), лежащих между P_{min} и P_{max}. Каждое такое отношение получается добавлением к отношению P_{min} некоторого подмножества из $P_{max} \setminus P_{min}$. Тем самым множество всех отношений, лежащих между P_{min} и P_{max}, имеет ту же структуру, что и множество всех подмножеств $P_{max} \setminus P_{min}$. Поскольку множество всех подмножеств $P_{max} \setminus P_{min}$ естественным образом отождествляется с вершинами i-мерного куба, то и множество всех отношений, лежащих между P_{min} и P_{max}, имеетту же структуру. При этом отношения P_{min} и P_{max} соответствуют противоположным вершинам такого куба.

Однако не все вершины построенного таким образом *d*-мерного куба соответствуют частичным порядкам. Все эти отношения являются антисимметричными, так как они содержатся в частичном порядке Ртах. Так как эти отношения можно выписывать последовательно, то, проверяя каждое из них на транзитивность, мы можем выделить среди 2^d вершин те из них, которые соответствуют отношениям частичного порядка.

В случае, если в выпуклом множестве **X** имеется несколько максимальных элементов, то все множество **X** является подмножеством объединения кубов, соответствующих всем парам (P_{min}, P_{max}). При этом два куба, соответствующие P^1_{max} и P^2_{max}, инцидентны по грани, соответствующей паре $(P_{min}, P^1_{max} \cap P_{max})$.

В общем случае рассмотрим грань, по которой инцидентны все кубы, соответствующие имеющимся в X максимальным элементам. Из предыдущих построений следует, что эта грань явля-^ ется кубом, соответствующим (P_{min}, P_h), где P_h — частичный порядок, ЯВЛЯЮЩИЙСЯ пересечением всех максимальных элементов из X.

17.2. Базис и ядро в пространстве $\mathcal{PO}$

Как уже указывалось выше, любое выпуклое множество $\mathbf{X} \subset \subset \mathcal{PO}$ порождается множеством, состоящим из минимального элемента множества X и всех его максимальных элементов. Однако априори не очевидно, что для построения выпуклого множества X необходимо использовать все максимальные элементы этого множества. В общем случае множество X является выпуклой оболочкой множества, состоящего из минимального и некоторого собственного подмножества максимальных элементов. Введем следующее

Определение 5.1. *Базисом* выпуклого множества X назовем произвольное подмножество множества всех максимальных элементов из X, которое вместе с минимальным элементом порождает X.

Из рассмотрения в предыдущем параграфе структуры выпуклого множества в пространстве $\mathcal{PO}$ следует, что существует максимальная грань, по которой инцидентны все кубы, соответствующие парам $(P_{\min}, P_{\max})$, где $P_m\&x$ пробегает множество всех базисных элементов. Эта грань является кубом с вершинами

P_{mn} и $\widehat{P}_{\max}$, где $\widehat{P}_{\max}$ есть пересечение всех максимальных эле-
ментов в X. С геометрической точки зрения точки этого куба образуют выпуклое множество, расположенное «однородно» относительно исходного выпуклого множества. Это позволяет ввести следующее

Определение 5.2. *Ядром* выпуклого множества $\mathbf{X} \subset \mathcal{PO}$ будем называть множество всех точек, лежащих между минимальной точкой и пересечением всех базисных точек в X.

Пересечение всех максимальных точек из базиса будем называть *ядерным отношением* и обозначать P_h.

Выделим два возможных крайних случая. Первый, когда в множестве X имеется всего лишь одна максимальная точка. В этом случае все точки множества X лежат между минимальной и данной максимальной точками и ядро совпадает с самим множеством X. Второй случай имеет место тогда, когда ядро состоит из одной точки. Эта точка является минимальной точкой, которая является пересечением всех максимальных точек.

Возвращаясь к содержательной постановке задачи (гл. I), напомним, что выпуклая оболочка представляет собой множество всех возможных групповых решений. В силу «однородности» расположения точек ядра относительно выпуклой оболочки исходного множества предпочтений, естественно считать отношения, Принадлежащие ядру, отношениями, *допустимыми для поиска групповых решений.* С этой точки зрения ядро является множеством допустимых групповых решений.

§ 5.3. Геометрические структуры в пространстве $\mathcal{QT}$

Напомним, что точками пространства $\mathcal{QT}$ служат все квази-транзитивные отношения слабого предпочтения R на фиксированном конечном множестве А. Условие квазитранзитивности отношения R состоит в том, что соответствующее ему отношение строгого предпочтения $P = \overline{R^{-1}}$ должно быть транзитивным отношением, т. е. частичным порядком. Тем самым отображение

i, заданное условием $i\colon \mathbf{M} \mapsto \overline{\mathbf{M}^{-1}}$, является биективным отображе-

нием пространства $\mathcal{QT}$ на пространство $\mathcal{PO}$.

Отметим важные свойства отображения i. Во-первых, оно обращает символ включения сг для отношений. Точнее, из Ri^{C1} <=/?$_2$

следует $i(R_1) \supset i(R_2)$. Во-вторых, при отображении i пере-

сечение отображений переходит в объединение образов и наоборот. Таким образом,

$$i(R_1 \cap R_2) = i(R_1) \cup i(R_2) \quad \text{и} \quad i(R_1 \cup R_2)$$

Так как основные геометрические структуры в пространствах бинарных отношений вводились в терминах символов cz, U и П, то, используя терминологию теории структур, можно сказать, что i осуществляет дуальный изоморфизм пространств

$\mathcal{QT}$ и $\mathcal{PO}$. Используя этот дуальный изоморфизм, можно перенести все результаты, полученные для пространства $\mathcal{PO}$ на пространство $\mathcal{QT}$ (разумеется, в двойственной формулировке). Мы проиллюстрируем это положение, доказав полноту пространства $\mathcal{QT}$.

Лемма 5.3. *Объединение двух квазитранзитивных предпочтений есть снова квазитранзитивное предпочтение.*

Доказательство. Пусть $R = Ri\backslash)R_z$, где Ri и R_2 — квазитранзитивные предпочтения.

Имеем $\overline{R^{-1}} = \overline{R_1^{-1}} \cap \overline{R_2^{-1}} = P_1 \cap P_2$. Так как Р$_4$ П P_2 есть частичный порядок, то R есть квазитранзитивное предпочтение.

Лемма 5.4. *Если Ri и R_z — два соседних квазитранзитивных предпочтения, то их симметрическая разность есть одноэлементное множество.*

Доказательство. Рассмотрим отношения $P_1 = i(R_1)$ и $P_2 = i(R_2)$ в пространстве $\mathcal{PO}$. Если в этом пространстве су-
ществует отношение P, отличное от Pi и P_2 и такое, что Ps s $[p_u$ Р$_2]_?$ то отношение $R = i(P)$, в силу биективнасти i, отлич-
но от Ri и R_2 и $R \in [R_1, \ R_2]$, что противоречит тому, что точки
Ri я Rz — соседние. Отсюда следует, что Pi и P_2 — соседние точки в $\mathcal{PO}$.

По определению симметрической разности имеем

$$R_1 \Delta R_2 = (R_1 \cap \overline{R}_2) \cup (\overline{R}_1 \cap R_2) =$$
$$= (\overline{P_1^{-1}} \cap P_2^{-1}) \cup (P_1^{-1} \cap$$

Но $Pi\&P_2$ есть одноэлементное множество в силу леммы 5.1, откуда и следует доказываемое утверждение.

Из доказанной леммы непосредственно следует, что справедлива

Теорема 5.2» *Пространство* $\mathcal{QT}$ *есть полное пространство.*

Точно таким же образом в пространство $\mathcal{QT}$' можно перенести понятия базиса, ядра, ядерной точки, рассмотренные в пространстве $\mathcal{PO}$. Вое эти понятия переносятся из в $\mathcal{PO}$ в $\mathcal{QT}$
двойственной формулировке, т. е. с заменой минимальных элементов на максимальные и наоборот.

Вся теория выпуклых множеств в пространстве $\mathcal{QT}$ могла бы быть построена на основе непосредственного изучения структур этого пространства так, как это было сделано для пространства $\mathcal{PO}$. Мы предпочли использованный здесь подход, так как он естественным образом вытекает из идей, развитых в главе III, где была построена диаграмма пространств бинарных отношений.

§ 5.4. Построение ядра в пространстве $\mathcal{PO}$

Из результатов главы IV, опираясь на два определения понятия «между», можно получить точные алгоритмы для построения выпуклой оболочки исходного множества предпочтений, чтобы затем для этой оболочки построить ядро. Однако основную задачу — построение ядра — можно решить, не используя описания всего выпуклого множества. Опишем алгоритм построения ядра.

Пусть М =* {Pi, ..., P_h} — множество точек в пространстве $\mathcal{PO}$, которые мы интерпретируем как индивидуальные предпочтения, и C(M) — выпуклая оболочка множества М. Пусть $P =$

$$= \bigcup_i P_i,\ i = \overline{1,\ k}.$$

Обозначим для каждого $P_i \in \mathbf{M}\ \ (i = \overline{1,\ k})$ через *Pi* максимальный элемент в *P,* содержащий $P_i : P_i' \supset P_i$. Для некоторых *i* может быть *P[* = *P%* или $P_i' = P_j'\ (i, j = \overline{1, k};\ i \neq j)$.

Теорема 5.3. *Отношение* $P_k = \bigcap\limits_i P_i'\ \ (i = \overline{1,\ k})$ *есть ядерное от-*

ношение для C(M) *и ядро* C(M) *состоит из всех точек, лежащих между* $P_{\min} = \bigcap\limits_i P_i$ *и* P_k.

Доказательство. Покажем сначала, что C(M) совпадает с выпуклой оболочкой множества $\{P_{\min}, P'_1, \ldots, P'_k\}$. Так как

$$P_{\min} \subseteq P_i \subseteq P'_i,$$

то $P_i \in [P_{\min}, P'_i]$; таким образом, все /\ принадлежат выпуклой оболочке $\mathbf{C}(\{P_{\min}, P'_1, \ldots, P'_k\})$. Согласно лемме 4.1 отсюда следует, что

$$\mathbf{C}(\mathbf{M}) \subseteq \mathbf{C}(\{P_{\min}, P'_1, \ldots, P'_k\}).$$

Так как $P_i \subseteq P'_i \subseteq P_j$, то

$$\bigcap_{j=1}^{k} P_j \subseteq P'_i \subseteq \bigcup_{j=1}^{k} P'_j,$$

т. е/ $P'_i \in \mathbf{C}(\mathbf{M})$,

$i = \overline{1, k}$. Так как и $P_{\min} \in \mathbf{C}(\mathbf{M})$, то снова согласно лемме 4.1

$$\mathbf{C}(\mathbf{M}) \supseteq \mathbf{C}(\{P_{\min}, P'_1, \ldots, P'_k\}),$$

откуда следует доказываемое ут* верждение:

$$\mathbf{C}(\mathbf{M}) = \mathbf{C}(\{P_{\min}, P'_1, \ldots, P'_h\}).$$

Из этого равенства с учетом очевидной максимальности всех

$P_1, \ldots, P'_k$ в $\mathbf{C}(\mathbf{M})$ следует, что множество $\{P'_1, \ldots, P'_k\}$ есть базисное множество, откуда и следует утверждение теоремы. Эта теорема (в «двойственной» формулировке) справедлива также и в пространстве $\mathcal{QT}$.

Интерес к ядру, построенному в этой теореме, вызван тем, что базис, на основе которого строится ядро, порождается исходной совокупностью точек P_u ..., P_h. Поэтому с точки зрения геометрического подхода это ядро отражает индивидуальную структуру *исходного* набора предпочтений, а не структуру произвольного базиса, порождающую то же выпуклое множество.

Исходя из свойств максимальных элементов и доказательства теоремы 5.3, легко описать алгоритм поиска максимального элемента *Pi,* ближайшего к данному Р». Для исходной совокупности индивидуальных предпочтений *(P_u ..., P_h)* определим отношение

$$P = \bigcup_{i=1}^{k} P_i.$$

Затем для каждого *Pi* ищем максимальный элемент, добавляя к *Pi* по одной паре предпочтений *{#, у)* из *APi^PXPi* и проверяя полученное отношение на транзитивность. Процесс продолжается до тех пор, пока какую бы пару мы ни добавляли, транзитивное отношение не получается. Если такой пары предпочтений в $AP_{\{}$ не нашлось, то, следовательно, *P{* есть уже максимальная точка.

§ 5.5. Построение ядра в пространстве $\mathcal{QO}$

В задачу настоящей работы не входит подробное изучение структуры выпуклых множеств в пространстве $\mathcal{QO}$ аналогично тому, как это было сделано для пространства $\mathcal{PO}$ и $\mathcal{QT}$. Однако ввиду того, что отношения линейного квазипорядка в практических задачах встречаются довольно часто, здесь будет рассмотрен эвристический алгоритм построении ядра выпуклого множества 1в пространстве $\mathcal{QO}$, осно-* ванный на полученных выше результатах.

Пусть исходное множество отношений индивидуального предпочтения X задано в пространстве $\mathcal{QO}$. В главе III была рассмотрена коммутативная диаграмма 3.5 пространств бинарных отношений. Фрагмент этой диаграммы, содержащий три рассматриваемых пространства, изображен на рис. 5.2.

$$\begin{tikzcd} \mathcal{QT} \arrow[r, "i"] & \mathcal{PO} \\ \mathcal{QO} \arrow[u, "\varphi"] & \end{tikzcd}$$

Рис. 5.2.

Напомним, что отображение <р есть отображение вложения, т. е. в нашем случае пространство линейных квазипорядков $\mathcal{QO}$ есть подмножество пространства квазитранзитивных отношений и вкладывается в последнее целиком. Точки пространства *QO* в пространстве $\mathcal{QT}$ характеризуются как транзитивные отношения квазитраизитивного предпочтения. Используя отображения *i* и q>, можно определить образ множества X в пространстве $\mathcal{PO}$ как множество $\mathbf{Y} = (i \circ \varphi)(\mathbf{X})$.

Для множества Y, согласно изложенному выше, можно построить ядро K(Y) в пространстве $\mathcal{PO}$. В дальнейшем мы будем различать следующие две ситуации:

1. Прообраз множества K(Y) в пространств е $\mathcal{QO}$ относительно отображения $i \circ \varphi$ непуст. В этом случае за ядро исходного множества *X* принимаем прообраз $(\varphi^{-1} \circ i^{-1})\ (\mathbf{K}(\mathbf{Y}))$ ядра множества Y.

2. Прообраз K(Y) в $\mathcal{QO}$ пуст. Это означает, что множество r~4K(Y)) в пространстве $\mathcal{QT}$ _ не содержит транзитивных отношений. В этом случае естественно пополнить множество K(Y) таким отношением, которое, с одной стороны, имело бы прообраз в

$\mathcal{QO}$, а с другой стороны, «не слишком бы отличалось» по своему геометрическому расположению от ядра.

В качестве такой точки предлагается рассматривать максимальную точку Рк, содержащуюся (в смысле «не превосходящую») в $P_{\min} \in \mathbf{K}(\mathbf{Y})$ и такую, что ее прообраз всуществует. $\mathcal{QO}$ Выбор именно такой точки продиктован следующими геометрическими соображениями. Любая другая точка вне ядра содержится в пересечении некоторых (не всех) максимальных элементов выпуклой оболочки С(Х). Следовательно, эта точка «ориентирована» на эти максимальные элементы и расположена «неравномерно» по отношению к исходному множеству.

Следующее утверждение позволит нам указать алгоритм для построения отношения Рк.

Теорема 5.4. *Образ точки Рк при*отображении* i"[1] *совпадает с транзитивным замыканием отношения* $i^{-1}(P_{\min})$.

Доказательство. Поскольку точка *Р*к есть максимальное отношение в $\mathcal{PO}$, содержащееся в $P_m щ$ и имеющее прообраз в $\mathcal{QO}$, то $i^{-1}(P_{\mathrm{K}})$ есть минимальное отношение в $\mathcal{PO}$, содержащее $i^{-1}(P_{\min})$ и такое, что $i^{-1}(P_{\mathrm{K}})$ есть минимальное транзитивное отношение в $\mathcal{QT}$, содержащее *i^iPmm).* Отсюда следует, что $i^{-1}(P_{\mathrm{K}})$ совпадает с транзитивным замыканием $i^{-1}(P_{\min})$.

Итак, для построения ядра множества X в пространстве $\mathcal{QO}$ мы последовательно рассматриваем точку $P_т$т, ее прообраз $\hat{i}^{-1}(P_{\min})$ в пространстве $\mathcal{QT}$, транзитивное замыкание отношения

$\widehat{i^{-1}(P_{\min})}$ и, наконец, в качестве единственной точки — ядра мно-

жества X — образ отношения $\widehat{i^{-1}(P_{\min})}$ в пространстве $\mathcal{QO}$.

Подводя итог исследованиям, проведенным в этой главе, отметим, что в ней предложено решение основной задачи — построение множества допустимых групповых решений для трех пространств предпочтений. В дальнейшем мы рассмотрим алгоритм построения ядер в этих пространствах и различные примеры, иллюстрирующие основные введенные понятия.

§ 5.6. Блок-схема алгоритма «Ядро»

В предыдущих параграфах этой главы было получено строгое решение задачи построения множества допустимых групповых решений в пространствах $\mathcal{QT}$ и $\mathcal{PO}$ и предложен эвристический метод решения этой задачи для случая, когда исходные данные принадлежат пространству $\mathcal{QO}$. Поскольку решение последней задачи включает решение первой, то при разработке алгоритма естественно рассматривать более общий случай, когда исходные данные принадлежат $\mathcal{QO}$ и в этом же пространстве ищутся допустимые групповые решения.

Итак, пусть исходные данные являются отношениями линейного квазипорядка, т. е. экспертные оценки представляются в виде ранжирований исследуемых объектов по предпочтениям. Зафиксируем нумерацию оцениваемых объектов. Будем через / в 1, 2, ..., *n* обозначать текущий индекс объекта, через i = l,2,...,m — текущий индекс эксперта, через

$$(R_i) = (r_{i1},\ r_{i2},\ \ldots,\ r_{ij},\ \ldots,\ r_{in})$$

—

ранжирование, полученное от 1-го эксперта, где *Гц* — ранг /-го объекта, присвоенный ему г-м экспертом. Таким образом, входными

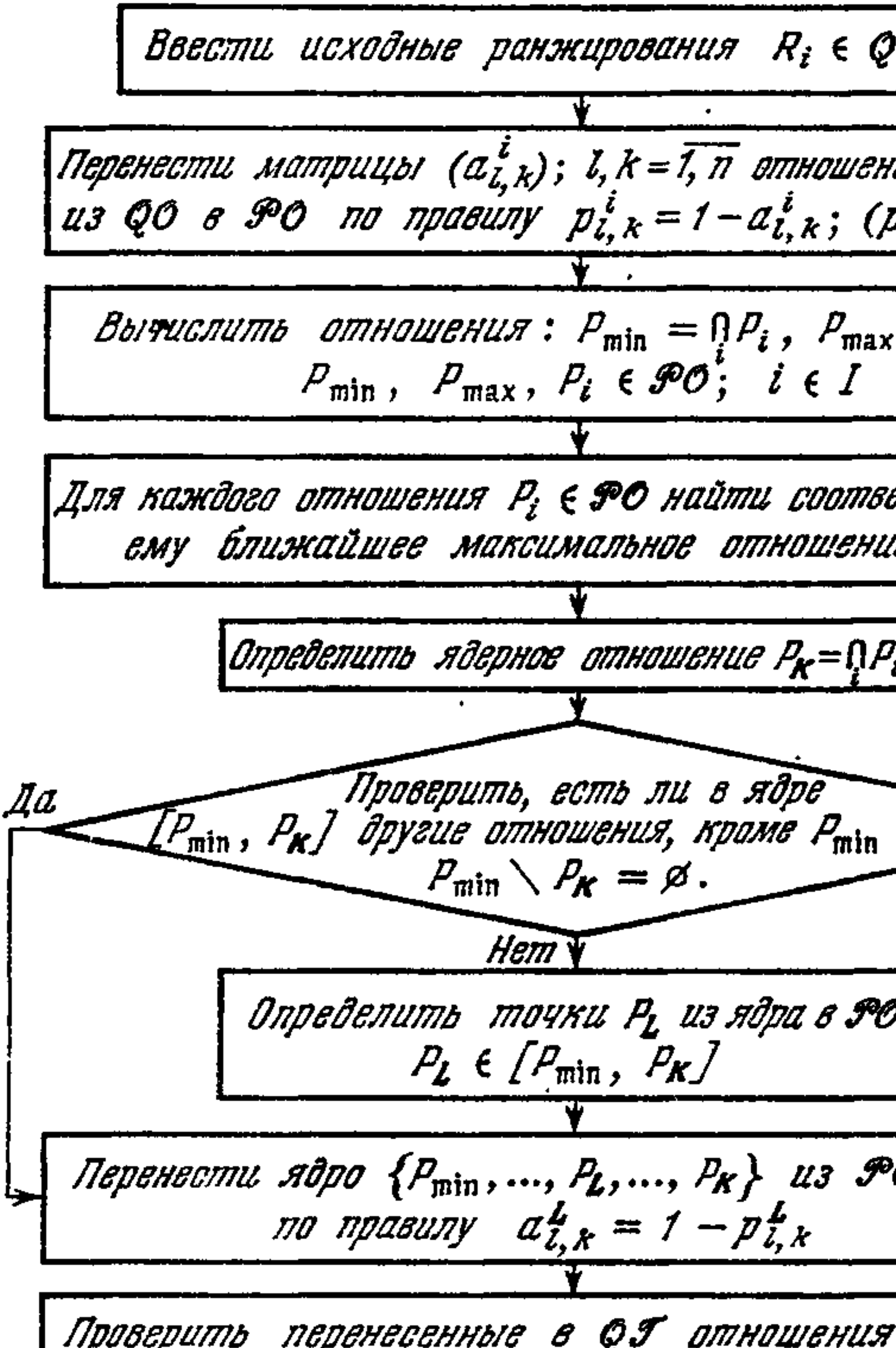

Ввести исходные ранжирования $R_i \in$
Перенести матрицы $(a^i_{l,k})$; $l,k=\overline{1,n}$
из QO в $\mathcal{P}O$ по правилу $p^i_{l,k}=1-a^i_{l,k}$;
Вычислить отношения: $P_{\min}=\bigcap_i P_i$,
$P_{\min}$, $P_{\max}$, $P_i \in \mathcal{P}O$; $i \in I$
Для каждого отношения $P_i \in \mathcal{P}O$ найти
ему ближайшее максимальное
Определить ядерное отношение
Да
Проверить, есть ли в ядре
$[P_{\min}, P_K]$ другие отношения, кроме $P_{\min}$
$P_{\min} \setminus P_K = \varnothing$.
Нет
Определить точки P_L из ядра в
$P_L \in [P_{\min}, P_K]$
Перенести ядро $\{P_{\min}, \ldots, P_L, \ldots, P_K\}$ из
по правилу $a^L_{l,k}=1-p^L_{l,k}$
Проверить перенесенные в $\mathcal{O}\mathcal{T}$

Рис. 5.3.

данными являются: *n,* *m* и $\{(R_1),\ (R_2),\ \ldots,\ (R_m)\}$. Блок-схема на рис. 5.3 описывает работу алгоритма.

Пояснения к блок-схеме будут сопровождаться примерами на трех объектах *а, Ь* и *с.*

1. Для каждой ранжировки $R_{\{}$ выписываем матрицу (a^i_{lk}); Z, *к* = 1, 2, ..., *п* отношения предпочтения иго эксперта по правилу

$$a^i_{lk}=\begin{cases}1, & \text{если} \quad r_{il}-r_{ik}\leqslant 0,\\ 0, & \text{если} \quad r_{il}-r_{ik}>0.\end{cases}$$

Пусть *Ri* = (1, 1, 2), т. е. эксперт считает, что $a\sim b>c.$ Матрица

отношения предпочтения *i-то* эксперта в пространстве $\mathcal{QO}$ имеет вид

$$(a_{lk})^i=\begin{array}{c|c|c|c|}
 & a & b & c\\
\hline
a & 1 & 1 & 1\\
\hline
b & 1 & 1 & 1\\
\hline
c & & & 1\\
\hline
\end{array}$$

2. Переносим исходные данные из $\mathcal{QO}$ в $\mathcal{PO}$ в соответствии с формулой

$$p^i_{lk}=1-a^i_{lk}.$$

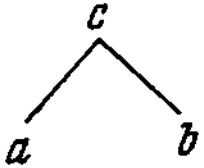
Для выписанной в п. 1 матрицы отношений в $\mathcal{QO}$ пространстве

$\mathcal{PO}$ имеем матрицу *Pc*

$$P_i = \begin{array}{c|c|c|c|} & a & b & c \\ \hline a & & & \\ \hline b & & & \\ \hline c & 1 & 1 & \\ \hline \end{array}$$

а сама точка *Pi* в графическом представлении имеет вид

p_i :

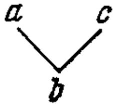


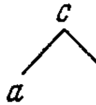
3. Выписываем минимальное отношение $P_{\min} = \bigcap_{i=1}^{m} P_i$ по условию

$$p_{lk}^{\min} = \begin{cases} 1, & \text{если } p_{lk}^{i} = 1 \text{ по всем } i \\ 0, & \text{если } p_{lk}^{i} = 0 \text{ по всем } i \end{cases}$$

Так, для трех отношений в $\mathcal{PO}$

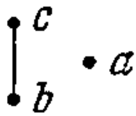
$$P_1 = \begin{array}{c|ccc|} & a & b & c \\ \hline a & & 1 & \\ b & & & \\ c & & 1 & \\ \hline \end{array} \qquad P_2 = \begin{array}{c|ccc|} & a & b & c \\ \hline a & & 1 & \\ b & & & \\ c & 1 & 1 & \\ \hline \end{array} \qquad P_3 = \begin{array}{c|cc} & a & b \\ \hline a & & \\ b & & \\ c & 1 & 1 \end{array}$$





имеем

$$\bigcap_{i=1}^{3} P_i = \begin{array}{c|ccc|} & a & b & c \\ \hline a & & & \\ b & & & \\ c & & 1 & \\ \hline \end{array} ;$$



4.Выписываем максимальное отношение в пространстве $\mathcal{PO}$

$$P_{\max} = \bigcup_{i=1}^{m} P_i$$ по правилу

$$p_{lk}^{\max} = \begin{cases} 1, & \text{если} \quad p_{lk}^{i} = 1 \quad \text{хотя бы} \\ 0, & \text{если} \quad p_{lk}^{i} = 0 \quad \text{для всех} \end{cases}$$

Для отношений из п. 3 имеем   -

$$\bigcup_{i=1}^{3} P_i = \begin{array}{c|c|c|c|} & a & b & c \\ \hline a & & 1 & \\ \hline b & & & \\ \hline c & 1 & 1 & \\ \hline \end{array}$$

5. Для каждой из точек *P{* в пространстве $\mathcal{PO}$ теперь нужно найти соответствующую ей максимальную ближайшую точку

$$P_i' \,(P_i \subset P_i').$$

Обозначим $\bigcup\limits_{j=1}^{m} P_j = (p_{lk}) = \bar{P}.$

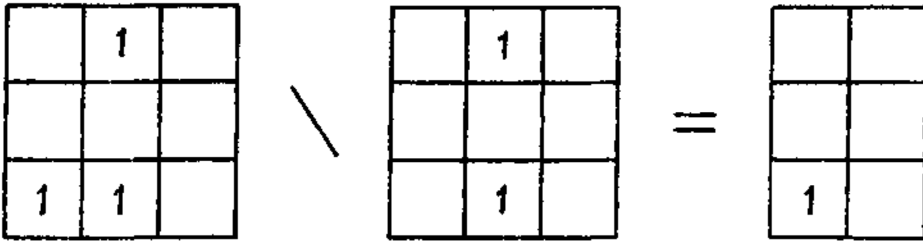
5а) Цикл по *i* =■ 1, 2, ..., *m.* Выход в п. 6.

56) Для каждой точки *Pг* находим матрицу (∇p^i_{lk}) «добавок»,

т. е. матрицу тех ∇p^i_{lk}, добавление которых к P< по описанному ниже правилу позволит найти соответствующий максимальный элемент *P%*

$$(\nabla p^i_{lk}) = P \setminus P_i, \quad \nabla p^i_{lk} = \begin{cases} 1, & \text{если } p \\ 0, & \text{если } p \end{cases}$$

Для примера из п. 3 для P_4, имеем



$$\left(\bigcup_{j=1}^{3} p_j\right) \setminus (p_1) = (\nabla p^1_{ik}$$

5в) Цикл по *I, к* — 1, 2, ..., *п.* Выход в п. 5д. Из матрицы (∇p^i_{ik}) выбираем единичный элемент $\nabla p^i_{ik} = 1_s$

добавляем его к P< и получаем отношение *P%:*

$$P'_i = P_i \cup \nabla p^i_{ik}.$$

Проверяем полученную матрицу *P%* на транзитивность, для чего возводим *P* в квадрат и проверяем включение $P'_i \supseteq (P'_i)^2$.

Если включение выполняется, то переходим к п. 5г, если же — нет, то продолжаем перебор добавок ∇p^i_{ik}. Для нашего примера

$$\begin{array}{|c|c|c|}\hline & 1 & \\ \hline & & \\ \hline & 1 & \\ \hline \end{array} \cup \begin{array}{|c|c|c|}\hline & & \\ \hline & & \\ \hline 1 & & \\ \hline \end{array} = \begin{array}{|c|c|c|}\hline & 1 & \\ \hline & & \\ \hline 1 & 1 & \\ \hline \end{array}$$

$$(P_1) \quad \cup \quad (\nabla p_{31}^1) \quad = \quad (P_1')$$

$$(P_1')^2 = \begin{array}{|c|c|c|}\hline & & \\ \hline & & \\ \hline & 1 & \\ \hline \end{array} \subseteq \begin{array}{|c|c|c|}\hline & 1 & \\ \hline & & \\ \hline 1 & 1 & \\ \hline \end{array} = (P_1')$$

5г) Полагаем $P_i := P_i'$. Повторяем вычисления пункта 56 с
той лишь разницей, что вместо матрицы отношения *Pi* вычитаем из *P* матрицу полученного транзитивного отношения *P%*.

5д) Если ни одно добавление ∇p_{ik}^i к P< не привело к новой транзитивной точке *P*\, т. е. отношение *P*\ уже само есть максимальная точка, то в массив максимальных точек $\{P_i'\}$ записываем *Pi* = *Pi* и продолжаем цикл по t_f п. 5а. В противном случае в массив $\{P_i'\}$ заносим матрицу найденного максимального отношения *Pi* и возвращаемся к п. 5а.

6. К этому моменту сформирован массив максимальных точек

$\{P'_1, P'_2, \ldots, P'_m\}$. Выписываем ядерное отношение $P_{\mathbf{K}} = \bigcap_{i=1}^{m} P'_i$ по условию

$$p_{lk}^{\mathbf{K}} = \begin{cases} 1, & \text{если } p'_{lk} = 1 \text{ для всех } i \\ 0, & \text{если } p'_{lk} = 0 \text{ хотя бы д} \end{cases}$$

Точки $P_m$in и *Pк* определяют ядро в пространстве $\mathscr{PO}$ и представляют собой два допустимых групповых решения в $\mathscr{PO}$. удовлетворяющих принципу Парето. Все остальные решения заполняют линейный сегмент между $P_{\min}$ и $P_{\mathbf{K}}$. Перейдем теперь к построению линейного сегмента. Обозначим множество точек линейного сегмента через $\{P_{\mathbf{L}}\}$. 7. Выписываем матрицу $(\nabla p_{lk}^{\mathbf{L}})$:

$$P_{\mathbf{K}} \setminus P_{\min} = (\nabla p_{lk}^{\mathbf{L}})^{lk},$$

где

$$\nabla p_{lk}^{\mathbf{L}} = \begin{cases} 1, & \text{если } p_{lk}^{\mathbf{K}} - p_{lk}^{\min} = 1, \\ 0, & \text{если } p_{lk}^{\mathbf{K}} - p_{lk}^{\min} = 0. \end{cases}$$

8. Построим массив $\mathbf{L} = \{(l, k)\}$ адресов-индексов тех элементов матрицы $(\nabla p_{lk}^{\mathbf{L}})$, которые равны 1. Например, для матрицы

$$(\nabla p_{lk}^{L}) = \begin{array}{|c|c|c|} \hline & 1 & \\ \hline & & 1 \\ \hline 1 & & \\ \hline \end{array}$$

этот массив будет иметь вид L^Ul, 2), (2, 3), (3, 1)}.

Далее будем перебирать все возможные комбинации «добавок» из (∇p_{lk}^{L}) к отношению P_{min} следующим образом. Пусть $T =$

где через |L| обозначено число элементов в (∇p_{lk}^{L}),

$$= \underbrace{00 \ldots 0}_{|L| \text{ раз}},$$

равных 1. Теперь будем к T прибавлять единицу по модулю 2. Если в получаемых таким образом числах номера позиций, занимаемых единицами, отождествлять с номерами адресов *(I, к)* в массиве L, то таким образом мы переберем все возможные комбинации «добавок». Объединяя P_mm с этими комбинациями и проверяя полученное объединение на транзитивность, мы построим все точки линейного сегмента между P_{mln} и *Рк* и, следовательно, все допустимые групповые решения в пространстве $\mathcal{PO}$.

9. Если по условиям задачи допустимые групповые решения нужно получить в пространстве частичных порядков $\mathcal{PO}$, то массив *{Pi)* нужно вывести на печать.

В качестве дополнительных характеристик полученного множества допустимых групповых решений в $\mathcal{PO}$ можно подсчитать нормированное расстояние между «крайними» мнениями из ядра по формуле

$$d = \frac{\sum_{l,k=1}^{n} \left| p_{lk}^{\min} - p_{lk}^{\mathrm{K}} \right|}{n(n+1)}$$

и коэффициент согласия Кендалла для этих же точек [19]:

$$\tau = 1 - 2d.$$

10. Если групповые решения ищутся в пространстве линейных квазипорядков QO, то перенос точек ядра в пространство QO производится по формуле

$$a_{ij}^{l} = 1 - p_{ij}^{l}; \qquad i, j = \overline{1, n},$$

где (a_{ij}^{l}) есть матрица группового предпочтения в QO, $(p_{ij}^{l}) \in$

$\in \{P_{\mathrm{L}}\}$. Каждая полученная матрица («ij) проверяется на транзитивность.

11. Если множество допустимых групповых решений в пространстве QO пусто, то за групповое решение принимается транзитивное замыкание отношения $\left(1 - \bar{p}_{lk}^{\min}\right)$.

12. Работа алгоритма после выполнения п. 11 заканчивается восстановлением по матрицам групповых решений рангов соответствующих ранжирований. Глава VI

>

ОБЩИЙ АНАЛИЗ

ВЫПУКЛЫХ И МЕТРИЧЕСКИХ СТРУКТУР

В предыдущих главах для решения проблемы группового выбора была развита общая теория выпуклых множеств в пространствах бинарных отношений и рассмотрены ее реализации в трех конкретных пространствах отношений индивидуального предпочтения. Настало время ответить на два вполне уместных здесь вопроса: как реализуются результаты, полученные в рамках этой теории, в остальных пространствах системы пространств, представленной на диаграмме 3.5, и как соотносятся групповые решения, получаемые на основе предложенного подхода, и подхода, при котором для построения групповых решений используется метрическая структура?

Отвечая на эти вопросы, удобно взглянуть на построенную в главе III систему пространств как бы сверху. Мы начнем наш «обзор» с общего рассмотрения метрической структуры в полных пространствах и указания на ранее изученные в этом отношении пространства (§ 6.1). В следующем параграфе будут рассмотрены выпуклые и метрические структуры в нерассматри-вавшихся ранее неполных пространствах. В конце этого параграфа ответ на первый вопрос будет представлен в таблице, подводящей итог изучению свойств полноты и выпуклости в пространствах диаграммы 3.5.

В последнем параграфе этой главы будет проведено сравнение указанных двух подходов к решению проблемы группового выбора на конкретных примерах в двух пространствах диаграммы 3.5.

§ 6.1. Близость и метрика в полных пространствах бинарных отношений

Интуитивно понятие близости существует в любом пространстве с метрикой — мы говорим, что точка *R* «ближе» к точке *P,* чем точка @, если *d(R, P)<d{Q, R),* где *d*— функция расстояния в заданном пространстве. Оказывается, что в любомпространстве бинардых отношений аксиоматическое описание понятия «функция близости», более широкое, чем понятие метрики, в конечном итоге приводит к однозначной метрической структуре. Этим мы, безусловно, обязаны специфике рассматриваемых задач.

В этом параграфе мы рассмотрим функции близости и расстояния для случая полных пространств бинарных отношений, где связь между этими понятиями становится наиболее прозрачной.

Начнем со следующего общего определения.

Определение 6.1. Пусть $\mathbf{R}$— произвольное пространство бинарных отношений. *Функцией близости* на пространстве $\mathbf{R}$ называется каждая функция $\delta(P, Q)$ на $\mathbf{R} \times \mathbf{R}$, удовлетворяющая условиям:

А1. Аддитивность:

$$R \in [P, Q] \Rightarrow \delta(P, R) + \delta(R, Q) = \delta(P$$

А2. Нормировка: для любых двух соседних точек P и Q

$$\delta(P, Q) = 1.$$»

Замечание. В то время как условие А1 носит «универсальный» характер и используется при определении функции близости во *всех* пространствах, условие А2 может видоизменяться в зависимости от конкретного вида пространства. В том виде, как это условие представлено в определении 6.1, оно пригодно для всех полных пространств и, например, для пространств $\mathscr{T}\mathscr{O}$ и $\mathscr{T}\mathscr{P}\mathscr{O}$ (см. диаграмму 3.5).

Нашей ближайшей задачей будет установление существования и единственности функции близости, определенной условиями А1 и А2, для полных пространств бинарных отношений. Сначала установим, что в любом пространстве бинарных отношений справедлива

Лемма 6.1. *Если функция близости существует, то она определяется условиями* А1 *и* А2 *однозначно для любого пространства бинарных отношений* R.

Доказательство.

Пусть $P, Q \in \mathbf{R}$ и $P \neq Q.$ Согласно теореме 4.1 существует линейный сегмент между P и Q:

$$R_0 = P,\ R_1,\ \ldots,\ R_n = Q \qquad (n \geqslant 1).$$

Из условия А2 непосредственно выводим, что

$$\delta(P, Q) = \delta(R_0, R_1) + \delta(R_1, R_2) + \ldots$$

Согласно условию А2 отсюда следует $\delta(P, Q) = n,$. ^ . так как

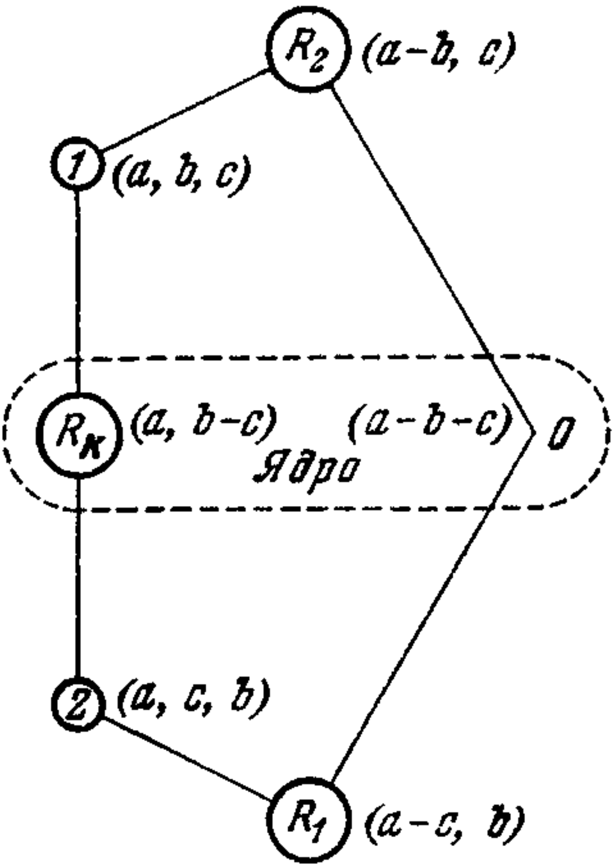
Ri-i и $R_{\{}$ — соседние точки в пространстве $\mathbf{R}$ для всех i. Если же $P = Q$, то из А1 немедленно получаем, что $\delta(P, Q) = 0.$ Следствие 6.1. Если функция б существует, то она удовлетворяет также условиям

А3. Симметрия: $\delta(P, Q) = \delta(Q, P) P, Q.$

А4. Неотрицательность: $\delta(P,\ Q) \geqslant 0$ и $\delta(P,\ Q) = 0 \Leftrightarrow P = Q.$

Наше предположение о существовании функции близости, конечно, было существенным при доказательстве вышеуказанных свойств. Рассмотрим фрагмент пространства QO для трех объектов

QO



Очевидно, что последовательность $R_1,\ 2,\ R_{\text{к}},\ 1,\ R_2$ является линейным сегментом. Но R_{i4} О, R_2 также есть линейный сегмент. Так как 4 Φ 2, то мы делаем вывод, что в пространстве QO не существует функции б, удовлетворяющей условиям А1 и А2. Здесь дело в том, что, как мы уже указывали выше, формулировка условия А2 выбирается в зависимости от конкретного пространства бинарных отношений.

Теперь установим *существование* функции близости, удовлетворяющей условиям А1 и А2, в полных пространствах бинарных отношений. С этой целью напомним сначала читателю определение расстояния Хемминга между булевыми матрицами. Пусть $p = \{p_{ij}\}$ и $q = \{q_{ij}\}$ — две булевы матрицы размера *nХn*. Расстояние Хемминга %(/?, *q)* между ними определяется формулой

$$\chi(p, q) = \sum_{i,j=1}^{n} |p_{ij} - q_{ij}|. \quad (6.1)$$

Расстояние Хемминга удовлетворяет всем обычным свойствам геометрического расстояния.

Так как каждое бинарное отношение определяет булеву матрицу (см. § 2.1), то можно считать, что на любом пространствебинарных отношений существует метрика $d_\chi(P,\ Q)$, определяемая формулой

$$d_\chi(P,\ Q) = \chi(p,\ q), \quad (6.2)$$

где p и q — булевы матрицы отношений P и Q.

Лемма 6.2. *В полном пространстве бинарных отношений функция* $\delta(P,\ Q) = d_\chi(P,\ Q)$ *удовлетворяет условиям* А1 *и* А2.

Доказательство. Пусть $R \in [P,\ Q]$, т. е.

$$P \cap Q \subseteq R \subseteq P \cup Q. \quad (6.3)$$

Переходя к булевым матрицам отношений, перепишем 6.3 в виде

$$p_{ij} \wedge q_{ij} \leqslant r_{ij} \leqslant p_{ij} \vee q_{ij}, \quad (6.4)$$

где $\{p_{ij}\},\ \{q_{ij}\},\ \{r_{ij}\}$ — матрицы отношений *P, Q* и *R* соответствен-
но. Имеем в силу (6.1), (6.2)

$$(P, Q) = d_\chi(P, Q) = \sum_{ij} |p_{ij} - q_{ij}| =$$

$$= \sum_{ij} (|p_{ij} - r_{ij}| + |r_{ij} - q_{ij}|) = d_\chi(P$$

Тем самым установлено, что А1 выполняется.

Пусть теперь *Р* и *Q* — соседние точки. Так как пространство R полно, отсюда следует, что симметрическая разность *PAQ* есть одноэлементное множество. В терминах матриц *{pu}* и {д«} бинарных отношений *Р* и *Q* это означает, что *{рц)* и {g2j} различаются ровно одним элементом, т. е. для всех *i* и /, кроме одной пары, *\pij~ Цц* =0. Из (6.1) и (6.2) получаем

что и требовалось

$$\delta(P, Q) = d_\chi(p, q) = \sum_{ij} |p_{ij} - q_{ij}| =$$

доказать.

Из доказанного выше следует, что справедлива следующая Теорема 6.1. *На любом полном пространстве* **R** *существует единственная функция близости* 6СР, *Q), удовлетворяющая ус-ловиям* А1 *и* А2.

Эта функция совпадает с расстоянием Хеммин-га между матрицами бинарных отношений из пространства $\mathbf{R}$:

$$\delta(P, Q) = d_\chi(P, Q) = \sum_{ij} |p_{ij} - q_{ij}|,$$

где {/?ij, *qij} — матрицы отношений P и Q.*

§ 6.2. Пространства $\mathcal{TO}$ и $\mathcal{TPO}$

В этом параграфе мы довольно подробно рассмотрим структуру пространства $\mathcal{TPO}$ и изоморфного ему пространства— $\mathcal{TO}$ пространств, которые нам в этой работе еще не встречались. Эти пространства (вместе с тривиальпым пространством $\mathcal{E}$) образуют«первый этаж» диаграммы 3.5. Сначала мы рассмотрим выпуклые структуры, а завершим параграф изучением структур близости и метрики в пространстве $\mathcal{TPO}$.

Напомним определение пространства $\mathcal{TPO}$.

Определение 6.2. Пространство $\mathbf{R}$ бинарных отношений называется *пространством совершенных строгих порядков* $(\mathcal{TPO})$, если каждая его точка *L* есть бинарное отношение совершенного строгого порядка, т. е. удовлетворяет условиям

1. Антирефлексивность: $(x, x) \in L \ \forall x$
2. Транзитивность:

$$(x, y) \in L, (y, z) \in L \Rightarrow (x, z) \in L.$$

3. Связность: для любых *x* и *y* либо *(x, y)* s *L* либо $(y, x) \in L$. *Отношения совершенного строгого порядка,* или, как еще говорят, *отношения строгого линейного порядка* на конечном множестве А из *n* элементов, обладают довольно простой структурой. Так как, очевидно, каждое отношение *L* строгого линейного порядка на А есть в то же самое время отношение строгого частичного порядка, то в силу теоремы Шпильрайна на А существует

нумерация $\mathbf{A} = (a_1,\ a_2,\ \ldots,\ a_n)$, согласованная с L. Учитывая свойство связности строгого линейного порядка, получаем следующее утверждение.

Лемма 6.3. *Для любого линейного порядка* L *на конечном множестве* A *существует нумерация* $\mathbf{A} = (a_1,\ a_2,\ \ldots,\ a_n)$ *такая,*

что $(a_i,\ a_j) \in L \Leftrightarrow i < j.$

Очевидно, что предыдущая лемма устанавливает взаимнооднозначное соответствие между точками пространства $\mathcal{T}\mathcal{S}\mathcal{O}$ и различными нумерациями множества А. Фиксируя какую-либо нумерацию $\mathbf{A} = (a_1,\ \ldots,\ a_n)$, мы видим, что любая другая нумерация получается в результате некоторой *перестановки* элементов множества А. Поэтому число различных точек пространства $\mathcal{T}\mathcal{S}\mathcal{O}$ равно *n*\.

Изучим подробнее структуру *соседних* точек в пространстве $\mathcal{T}\mathcal{S}\mathcal{O}$. Напомним, что точка *L* лежит между точками L_4 и $Ь_2$ тогда и только тогда, когда

$$L_1 \cap L_2 \subseteq L \subseteq L_1 \cup L_2. \quad (6.5)$$

Для любых двух элементов *x* и *y* множества А пусть г, i_u *Ч* и *U Iu U* — их номера в нумерациях, согласованных соответственно с *L*, L_u L_2. Из (6.5) немедленно получаем

1) из $i_1 < j_1$ и $i_2 < j_2$ следует $i <$

2) из $i < j$ следует, что $i_1 < j_1$ или i

(6.6)

причем условия (6.6) являются необходимыми и достаточными для того, чтобы $L \in [L_1,\ L_2]$.

Лемма 6.4. *Согласованные нумерации соседних точек в* $\mathcal{T}\mathcal{S}\mathcal{O}$ *отличаются перестановкой двух последовательных элементов.*

Доказательство. Пусть L_t и L_2- две соседние точки и (#i, a_2, ..., a_n) — нумерация, согласованная с L_u а (s_h s_2, ..., s_n) — нумерация тех же элементов в нумерации, согласованной с L_2. Так как $Ь_2ФЬ_u$ то найдется i такой, что s_{i+i}<Si. Рассмотрим точку L, для которой номера элементов есть $(1, \ldots$

$\ldots, i-1, i+1, i, i+2, \ldots, n)$.

Для L легко проверяется выполнение условий (6.6) и, следовательно, $L \in [L_1, L_2]$. Так как $L_{\{}$ и L_2 — соседние точки и $ЬФЬ_u$ то L_2 = L, откуда и следует утверждение леммы.

.Теперь мы обратимся к выпуклым структурам пространства $\mathcal{TPO}$. Для этого пространства, конечно, справедливы все общие определения и результаты §§ 4.1—4.2. Однако пространство $\mathcal{TPO}$ очевидным образом не является полным пространством — матрицы двух соседних отношений, как это следует из леммы 6.4, различаются ровно на двух элементах. Несмотря на это здесь удается установить эквивалентность двух определений выпуклости (см. § 4.2).

Лемма 6.5. *В любом пространстве* $\mathcal{TPO}$ *из выпуклости в смысле определения* 4.4 *следует выпуклость в смысле определения* 4.5.

Доказательство.

Пусть $X = \{L_1, L_2, \ldots, L_k\}$ — множество, выпуклое в смысле определения 4.4, и пусть $L \in [L_1, \ldots, L_k]$. Рассмотрим линейный сегмент между Li и L_2. Предположим, что L & X. Так как L_t е X, то в линейном сегменте между L_4 и L найдутся две соседние точки Z/ и L'' такие, что Z/^X, а L'' $Ф$ §£X. Пусть (#i, ..., a_n) — нумерация элементов, согласованная с Z/. Тогда в силу леммы 6.4 номера тех же элементов в нумерации, согласованной с Ь", будут

$$(1, 2, \ldots, i-1, \underline{i+1}, \underline{i}, i+2, \ldots$$

$\ldots, n)$ для некоторого i. Пусть Ui, s_2, . •., s_n) — номера тех же элементов в нумерации, согласованной с L. Так как $L'' \in [L', L]$, то в силу второго условия из (6.6) имеем $s_{i+1} < s_i$. Пусть теперь $(s_1^r, s_2^r, \ldots, s_n^r)$ — номера тех же элементов в нумерации, согласованной с $L_r \in \mathbf{X}$ $(r = 1, 2, \ldots, k)$. Так как $L \in [L_1, \ldots, L_k]$, то $L \subseteq \bigcup_r L_r$.

Отсюда в силу $s_{i+i} < s_{\{}$ найдется номер г, для которого $s_{i+1}^r < s_i^r$. Покажем, что $L'' \in [L', L_r]$. Условие 1 из (6.6), оче-

видно, выполнено, так как $s_{i+1}^r < s_i^r$, at" отличается от *V* пере-
становкой объектов с номерами *i* и *i* +1. Из тех же соображений следует выполнение условия 2 из (6.6). Но $L' \in \mathbf{X}$ и $L_r \in \mathbf{X}$, а *L"* ^ X, что противоречит выпуклости X. Полученное противоречие показывает, что LeX, т. е. множество X выпукло в смысле определения 4.5. Лемма доказана.
Объединяя результат леммы 6.5 с результатом леммы 4.4, получаем теорему.

Теорема 6.2. *В пространстве* $\mathscr{TPO}$ *(и изоморфном ему пространстве* $\mathscr{TO}$*) оба определения выпуклости эквивалентны.* На этом мы закончим рассмотрение выпуклых структур в пространствах $\mathscr{TPO}$ и перейдем к изучению структуры близости
и метрики в этих пространствах.

Несмотря на то, что пространства $\mathscr{TPO}$ не являются полны-
ми, для них возможно определение функции близости. Так как в § 6.1 было показано, что из существования функции близости, удовлетворяющей условиям А1 и А2, следует ее единственность в любом пространстве бинарных отношений, то достаточно показать, что

в пространстве $\mathcal{TPO}$ существует функция $\delta(L_1,\ L_2)$, удовлетворяющая А1 и А2.

Определим $\delta(L_1,\ L_2)$ формулой

$$\delta(L_1, L_2) = \frac{1}{2}\chi(L_1,\ L_2). \qquad (6.7)$$

Условие А1 проверяется так же, как и в произвольном пространстве (см. § 6.1). Пусть теперь 1ч и *Ь*$_2$ — соседние точки в пространстве $\mathcal{TPO}$. Из леммы 6.4 следует, что матрицы отно-
шений L_4 и *Ь*$_2$ различаются ровно на двух элементах. Поэтому

$d_\chi(L_1,\ L_2) = 2$, откуда следует А2. Таким образом, мы доказали следующую теорему.

Теорема 6.3. *В пространстве* $\mathcal{TPO}$ *существует единственная функция близости, удовлетворяющая условиям* А1 *и* А2. *Эта функция близости задается формулой* (6.7).

Так как d_x является метрикой, то справедлива

Теорема 6.4. *Функция близости, определенная в теореме* 6.3, *задает метрическую структуру пространства* $\mathcal{TPO}$.

Полученные выше в этом параграфе результаты вполне аналогичны результатам для полных пространств. Возвращаясь к диаграмме 3.5, можно сказать, что для пространств предпочтений 1-го, 3-го и 4-го «этажей» этой диаграммы два общих определения выпуклости оказываются эквивалентными, что позволяет развивать для них геометрический подход к проблеме группового выбора, предложенный в данной работе. Далее, для этих же пространств оказался возможным общий подход к попятию функции близости (условия А1 и А2) и построение на его основе метрики в этих пространствах. Эта метрика для единственного из этих пространств, изучавшегося ранее,— пространства $\mathcal{PO}$, совпадает с введенной в работе [15] и позволяет развивать в этих пространствах метрический подход.

Обратимся теперь к пространствам 2-го «этажа» диаграммы 3.5. Это пространство $\mathcal{QO}$ (и изоморфное ему пространство $\mathcal{OP}$)

и пространство безразличия *0*~. Эти пространства рассматривались в работах [1] (пространство $\mathcal{QO}$) и [13, 26] (пространство $\mathcal{S}$, где оно названо пространством разбиений). В этих работах был намечен метрический подход к проблеме группового выбора в данных пространствах. Относительно сложная структура самихотношений в этих пространствах я, особенно, соседних отношений, не позволяет реализовать для них наш подход, основанный на функциях близости. Введение метрики в этих пространствах возможно только лишь на основе довольно громоздкой системы аксиом, хотя в конечном итоге эти метрики совпадают с расстоянием Хемминга, что имеет место и при нашем подходе.

§ 6.3. Полные и неполные пространства

Как видно из изложенного выше, в основе всех построений в рамках геометрического подхода лежит эквивалентность двух понятий выпуклости — геометрического и опирающегося на принцип единогласия Парето. В § 4.2 был доказан простой критерий, позволяющий установить наличие такой эквивалентности. Для этого *достаточно* потребовать, чтобы рассматриваемое пространство бинарных отношений удовлетворяло условию полноты Сем. § 4.2). Однако полнота пространств не является *необходимым* условием эквивалентности двух понятий выпуклости. Так, например, пространство $\mathcal{T}$ отношений толерантности не является полным, однако для него можно доказать эквивалентность двух понятий выпуклости.

На диаграмме 3.5 можно легко указать пространства, являющиеся полными. Это пространства $\mathcal{P}$, $\mathcal{QT}$, $\mathcal{SP}$, $\mathcal{PO}$ и $\mathcal{E}$, последнее — тривиально полное. Все остальные пространства полными не являются. Однако для некоторых из них довольно легко устанавливается эквивалентность двух понятий выпуклости. Такими пространствами являются два изоморфных пространства $\mathcal{TO}$ и $\mathcal{TSO}$, изученные в предыдущем параграфе, а также пространства $\mathcal{T}$ и $\mathcal{TOT}$.

Пространство разбиений $\mathcal{S}$ дает нам пока единственный пример пространства, в котором два определения выпуклости не совпадают и приводят, вообще говоря, к разным выпуклым множествам. Покажем это на следующем примере.

Пусть $\mathbf{A} = \{a,\ b,\ c\}$. Рассмотрим множество $\mathbf{X} = \{I_0,\ I_1,\ I_2,\ I_3\}$, элементы которого задаются матрицами отношений

$$(I_0) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (I_1) = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$(I_2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad (I_3) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Легко проверить, что множество X является выпуклым в смысле определения 4.4. Однако оно не является выпуклым в смысле определения 4.5, так как отношение эквивалентности / с матрицей

$$(I) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

лежит между всеми элементами из X (оно совпадает с $I_0 \cup I_1 \cup$ $\cup\, I_2 \cup I_3$), но не принадлежит X.

Следующая таблица подводит итог рассмотрению свойств полноты и выпуклости для пространств из диаграммы 3.5.

Свойство	П	
	$\mathcal{P}$, $\mathcal{E}$, $\mathcal{QT}$, $\mathcal{SP}$, $\mathcal{PO}$	$\mathcal{QO}$,
Пространство полно	**Да**	**Не**
Два определения выпуклости совпадают	**Да**	**?**

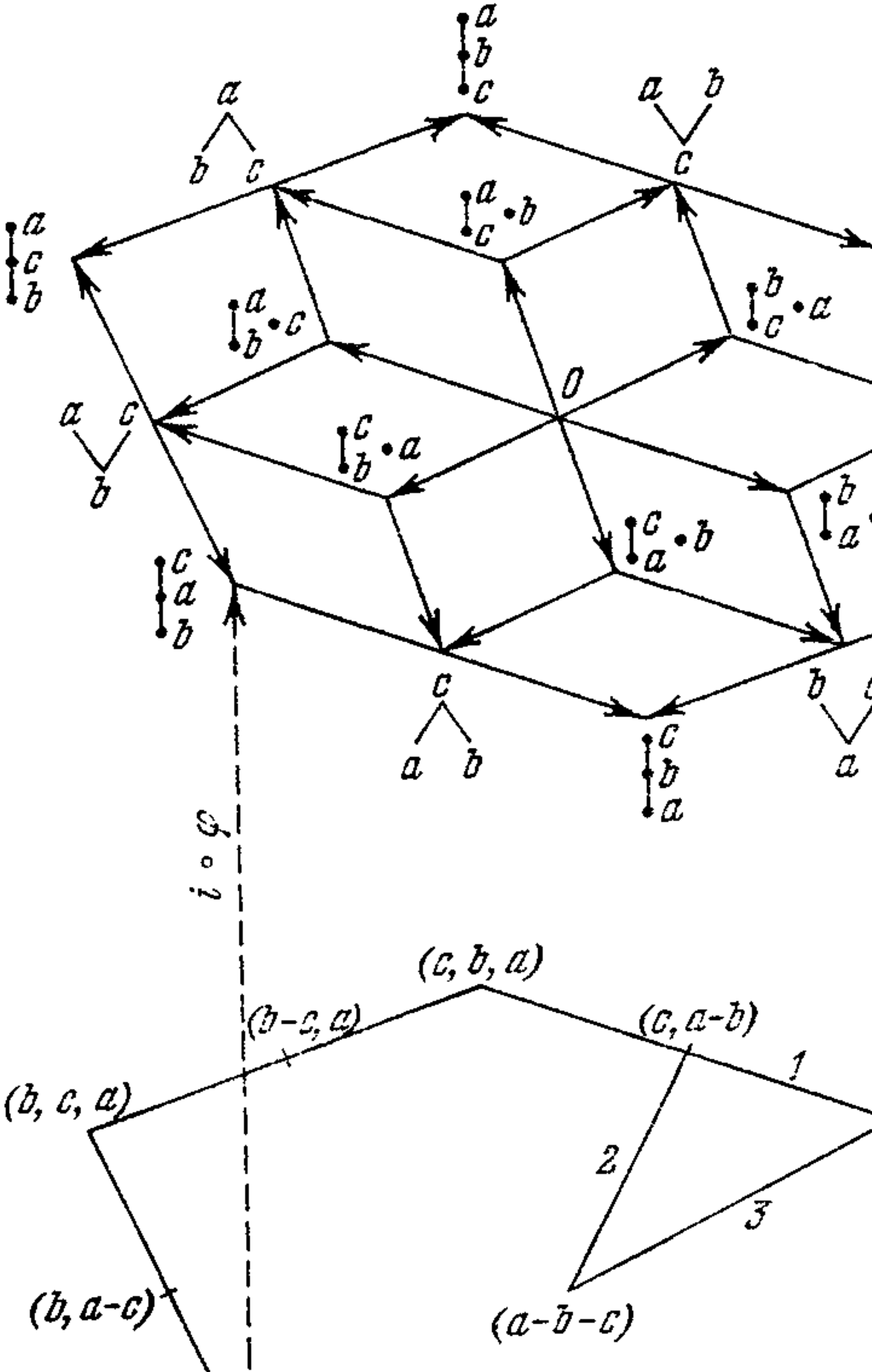
§ 6.4. Сравнение геометрического и метрического подходов

В атом параграфе мы проиллюстрируем работу алгоритма для построения ядра выпуклой оболочки на четырех парах простейших примеров и одновременно проведем обещанное сравнение результатов применения метрического и геометрического подходов к решению задач группового выбора. Пары примеров составлены следующим образом: в первом примере каждой пары исходные данные задаются в пространстве $\mathcal{QO}$ и в нем же ищется групповое решение, во втором примере в качестве исходных принимаются все точки выпуклой оболочки в пространстве $\mathcal{PO}$, полученной в первом примере пары, и групповое решение ищется также в $\mathcal{QO}$.

Все примеры проводятся для $\mathbf{A} = \{a, b, c\}$. Фрагменты про-
странств, изображаемые на рисунках, представляют собой выпуклые оболочки для исходных данных примера. Стрелками на рисунках указываются отношения включения. Для удобства представления о расположении исходных данных в пространствах

$\mathcal{PO}$ **и** $\mathcal{QO}$ на рис. 6.1 эти пространства изображены целиком. Исходные данные в первых примерах каждой пары для пространства $\mathcal{QO}$ будут обозначаться через $R_1, R_2, \dots,$ а соответствующие им точки в $\mathcal{PO}$ —через $P_1, P_2, \dots$ Начиная со второй пары примеров все исходные отношения будут сразу приводиться в одной таблице. Только в первом примере будет подробно описано формирование максимальных элементов P'_{i_2} ближайших кданным *P*^ В остальных же — они будут сразу отмечаться на рисунке.

Для сравнения условий, в которых приходится решать проблему группового выбора в пространствах $\mathcal{PO}$ и $\mathcal{QO}$, в каждом примере для каждого пространства будут определены точки ядра, точки, в которых достигается или минимум суммы расстояний


O
i ∘ φ
(c, b, a)
(b−c, a)
(c, a−b)
(b, c, a)
(b, a−c)
(a−b−c)
1
2
3

Ряс. 6.1.

до исходных, т. е. медианы, или минимум суммы квадратов таких расстояний, т. е. среднее *). Все эти данные будут сводиться

*) Мы предполагаем, что читатель уже знаком с основными положениями метрического подхода и способом построения группового решения в иен. Напомним здесь лишь то, что все расстояния между бинарными отношениями, вводимые при метрическом подходе аксиоматически, совпадают с расстоянием Хемминга между соответствующими булевыми матрицами отношений. в таблицы, непосредственно следующие за фрагментами пространств, являющихся выпуклыми оболочками для исходных данных примера.

Примеры подобраны таким образом, чтобы проиллюстрировать такие конфигурации расположения индивидуальных предпочтений, которые приводят к ядрам (множествам допустимых групповых решений) разной размерности — от минимального нулевого до максимального 3-мерного (подразумевая реализацию пространства $\mathcal{PO}$ в виде подмножества вершин куба размерности *nin* — D/2, где *п* — число объектов), а также показать различные случаи соотнесения точек ядра с точками, принимаемыми в качестве групповых решений в метрическом подходе, т. е. с медианами и средними.

Отметим очевидную из рис. 6.1 разницу в пространственном расположении точек в используемых пространствах. В пространстве $\mathcal{PO}$ все расстояния между соседними точками равны 1, т. е. в этом отношении пространство $\mathcal{PO}$ — «однородное», точки в нем расположены равномерно, без сгущений и разрежений. В пространстве $\mathcal{QO}$ дело обстоит не так. По периметру шестиугольника, на котором расположены 12 точек, расстояния между соседними точками равны 1, но между этими точками и тачкой, расположенной в центре пространства $\mathcal{QO}$. расстояние равно 2 или 3, как это показано на рис. 6.1.

Пример А1. Пусть исходное множество X состоит из двух линейных квазипорядков

$$R_1 = (b,\ a,\ c) \ \text{и}\ R_2 = (a - c,\ b).$$

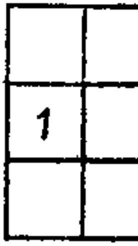
Выпишем матрицы отношений i?! и R_z:

$$(R_1) = \begin{array}{|c|c|c|}\hline 1 & & 1 \\ \hline 1 & 1 & 1 \\ \hline & & 1 \\ \hline \end{array}$$

$(R_2) =$ 

и их образы в пространстве $\mathcal{PO}$:

$$(P_1) = \begin{array}{|c|c|c|}\hline & 1 & \\ \hline & & \\ \hline 1 & 1 & \\ \hline \end{array}$$

$(P_2) =$ 

В данном примере исходные точки отстоят друг от друга на расстояние d = 5, т. е. почти на максимальное расстояние (йпшх^б). Построим ядро в $\mathcal{PO}$ в соответствии с процедурой, описанной в § 5.3. Находим

$$(P_{\min}) = (P_1 \cap P_2) = \begin{array}{|c|c|c|}\hline & & \\ \hline & & \\ \hline & & \\ \hline \end{array} \in \mathcal{PO}$$

$$(P) = (P_1 \cup P_2) = \begin{array}{|c|c|c|}\hline & 1 & \\ \hline 1 & & 1 \\ \hline 1 & 1 & \\ \hline \end{array} \notin \mathcal{PO}$$

Найдем теперь максимальные элементы P_t и $P_г$. Непосредственной проверкой убеждаемся, что какой бы элемент из разности $P\backslash Pi$ мы ни добавляли к Pi, мы не получим транзитивного отношения. Следовательно, P_4 уже само является максимальным элементом. В разности же $P\backslash P_2$ только добавление пары $\{(c,\ a)\}$ дает транзитивное отношение и, следовательно, $P_2' = P_2 \cup \{(c, a)\}$.

Отсюда получаем, что ядерное отношение $P_{\text{К}} = P_1' \cap P_2' = \{(c, a)\}$.

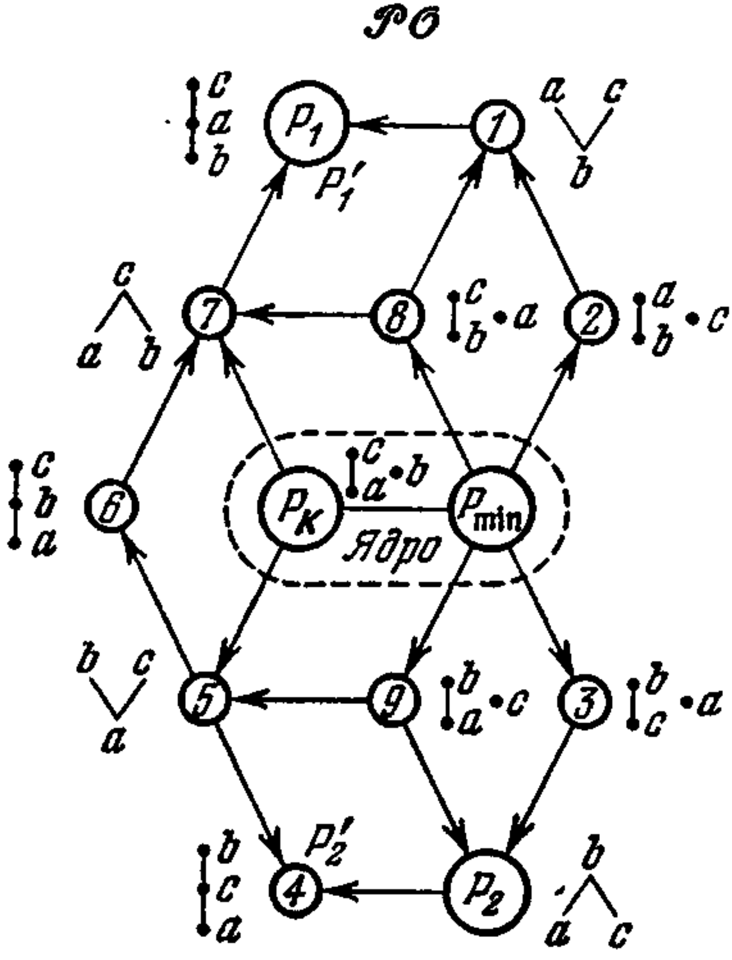
Таким образом, мы получили одномерное ядро, состоящее из двух точек: $P_{\min} = 0$ и $P_{\text{К}} = \{(c,\ a)\}$. Следовательно, в пространстве $\mathcal{PO}$ множество допустимых групповых решений состоит из двух отношений, первое из которых говорит о том, что эксперты «в среднем» не различают объекты или считают их несравнимыми, а второе — что эксперты могут «в среднем» провести различие только между двумя объектами а и с, а именно: U, с), а третий объект Ъ считают несравнимым с объектами а и с.

При переносе построенного ядра в пространство $\mathcal{QO}$ получаем только одно решение $R = (a - b - c).$, соответствующее Pmm, по-

скольку для *Рк* в пространстве $\mathcal{QO}$ яет соответствующего отношения. Полученное в $\mathcal{QO}$ отношение *R* содержательно означает, что «в среднем» эксперты считают все объекты эквивалентными.

Выпуклые оболочки для исходных данных представлены на рис. 6.2 и рис. 6.3 для пространств $\mathcal{PO}$ **и** $\mathcal{QO}$ соответственно.

Для рассматриваемых исходных данных все необходимые характеристики точек, составляющих выпуклую оболочку в обоих пространствах, сведены в таблицу 6.1. Эта таблица устроена следующим образом. В столбцах 3 и 6 помещены все точки, отмеченные на рис. 6.2 л 6.3 для пространств $\mathcal{PO}$ **и** $\mathcal{QO}$ соответственно. На одной и той же строке в этих столбцах указаны точки, соответствующие друг другу при переходе из одного пространства в другое, причем, если для точки из $\mathcal{PO}$ (например, точки ^к) в пространстве $\mathcal{QO}$ соответствующей точки нет, то в столбце 6 (напротив точки *Рк*) ставится прранстве точки распределены на три группы (см. столбцы 1 и 2): составляющие выпуклую оболочку, являющиеся исходными и образующие ядро выпуклой оболочки. В столбцах 4 и 7 для $\mathcal{PO}$ и $\mathcal{QO}$ соответственно звездочками отмечены точки, являющиеся


PO
P_1
P_1'
1
2
3
4
5
6
7
8
9
P_K
P_{min}
Ядро
P_2
P_2'

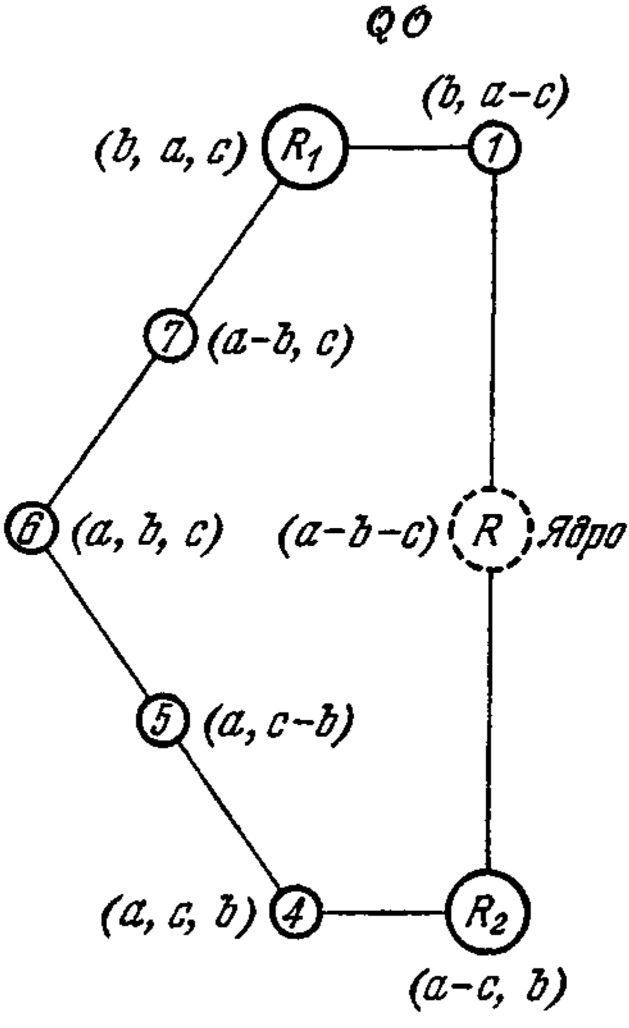


Рис. 6.2. Рис. 6.3.

медианами, а в столбцах 5 и 8 — являющиеся средними. Найти эти тачки можно непосредственным расчетом соответствующих сумм, подсчитывая расстояния от данной точки до всех исходных по рис. 6.1 или рис. 6.2 и 6.3. Заметим, что при расчете

Таблица 6.1

Точки			$\mathcal{PO}$		$\mathcal{QO}$		
			min Σd	min Σd^2	Точки	min Σd	min Σd^2
1	2	3	4	5	6	7	8
Выпуклой оболочки	Ядра	$P_{\min}$	*	*	R	*	*
		$P_{\text{К}}$	*	*	—		
	Исходные	P_1	*	—	R_1	*	
		P_2	*	—	R_2	*	
		1,4	*	—	1,4	*	
		2,8	*	*	—		
		3,9	*	—	—		
		5,6	*	*	5,6	*	*
		7	*	—	7	*	

медиан и средних во внимание принимались все точки пространств $\mathcal{PO}$ и $\mathcal{QO}$.

Таким образом, в этом примере, как это видно из таблицы 6.1, для двух исходных точек i?i и R_2 соответствующая им выпуклая оболочка в пространстве $\mathcal{QO}$ состоит из 8 точек, а в пространстве $\mathcal{PO}$ — из 13 точек. Ядро выпуклой оболочки в $\mathcal{QO}$ содержит одну точку Д, в $\mathcal{PO}$ —две точки: P_{mln} и Рк» Данные, характеризующие состояние проблемы группового выбора в рассматриваемом случае, представлены в таблице 6.2.

Таблица 6.2

		В пространств
Размерность ядра		1
Число точек в ядре		2
Возможные групповые решения при метрическом подходе (1-й уровень)	Медианы	Все 13 точек лой оболочки 2 ядерные)
	Средние	6 точек (из ядерные)
Допустимые групповые решения при геометрическом подходе (1-й уровень)		2 точки

Условия выбора группового решения в обоих пространствах при метрическом подходе можно прианать довольно сложными. Если групповое решение выбирать из числа медиан, то в QO таковым может служить любая из восьми точек выпуклой оболочки, а в $\mathcal{P}\mathcal{O}$ — любая из тринадцати точек выпуклой оболочки. Отметим при этом, что семь из восьми точек в $\mathcal{Q}\mathcal{O}$ и одиннадцать из тринадцати точек в $\mathcal{P}\mathcal{O}$ не попадают в ядро. Привлечение среднего только приблизительно наполовину уменьшает число возможных групповых решений. Все ядерные точки попали в обоих пространствах в число медиан и средних. При «комплексном» подходе, т. е. таком подходе, когда на* втором уровне выработки группового решения рассматриваются только те точки, которые на первом уровне удовлетворяют требованиям обоих подходов одновременно, в пространстве $\mathcal{Q}\mathcal{O}$ было бы только одно решение — точка й, а в $\mathcal{P}\mathcal{O}$ — только два решения: $P_{\min}$ и $P_{\text{К}}$. Пример А2. Будем теперь все точки выпуклой оболочки в пространстве $\mathcal{P}\mathcal{O}$, полученной в предыдущем примере А1, считать исходными, а поиск группового решения по-прежнему проводить в пространстве $\mathcal{Q}\mathcal{O}$. Очевидно, что при таком выборе исходных данных, выпуклые оболочки и их ядра в обоих пространствах (см. рис. 6.2 и 6.3) останутся без изменений.

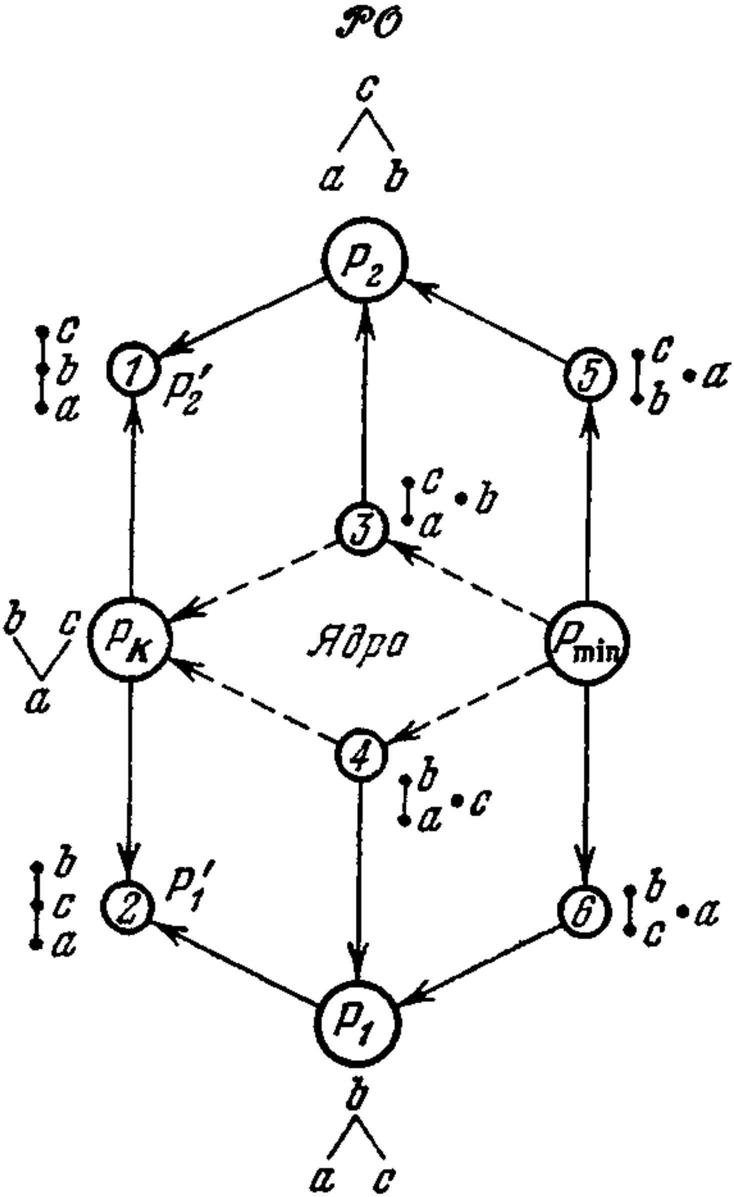
Таблица 6.3

Точки			$\mathcal{PO}$	
			min Σd	min
1	2	3	4	5
	Ядра́	$P_{\min}$ $P_{\text{К}}$	* *	*
Исходные и выпуклой оболочки		P_1 P_2 1, 4 2, 3, 8, 9 5, 7 6		

Таблица 6.4

		В пр
Размерность ядра		
Число точек в ядре		
Возможные групповые решения при метрическом подходе (1-й уровень)	Медианы	2 (я
	Средние	1 (я
Допустимые групповые решения при геометрическом подходе (1-й уровень)		2 то

Как видно из таблиц 6.3 и 6.4, условия группового выбора при метрическом подходе изменились: в пространстве $\mathcal{PO}$ только две точки являются медианами, а в $\mathcal{QO}$ — только три. При геометрическом подходе множество допустимых групповых решений остается неизменным, поскольку ядра выпуклых оболочек не изменились. Отметим две ицтересные особенности, выявляемые данной парой примеров. Первая — общая для всех пар рассматриваемых здесь примеров, состоит в том, что, как указывалось в § 5.2, выпуклая оболочка для каждой совокупности данных в общем


PO
c
a b
P_2
1
P_2'
c
b
a
5
c
b
a
3
c
a
b
P_K
b c
a
Ядро
P_{min}
4
b
a
c
2
P_1'
b
c
a
6
b
c
a
P_1
b
a c

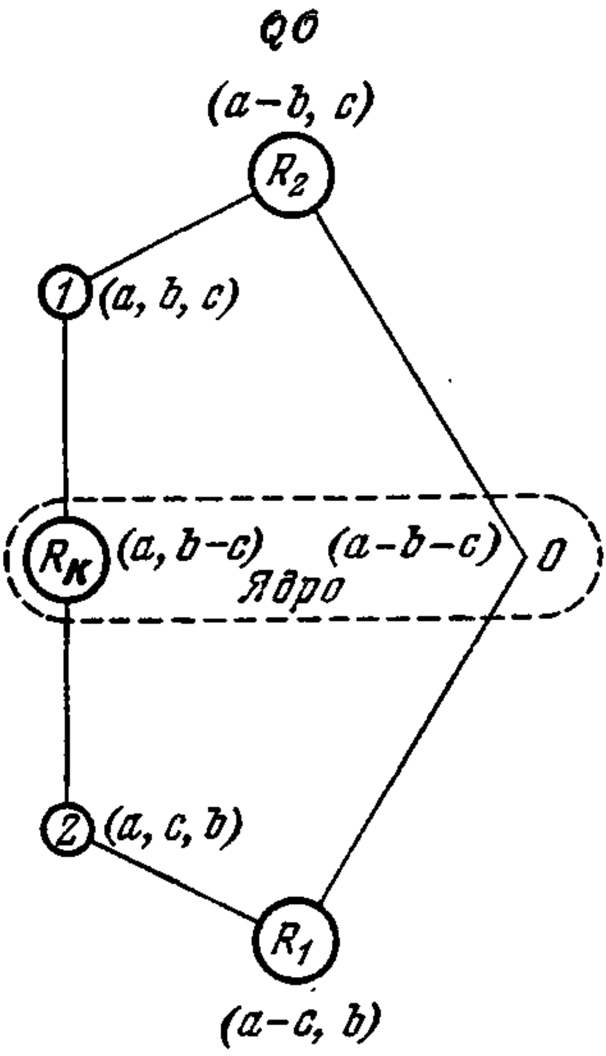


Рис. 6.4.   Рис. 6.5.

случае может порождаться несколькими базисами. Так, например, выпуклая оболочка (см. рис. 6.2) в $\mathcal{PO}$, кроме базисов *(Ри Рг)* и (1, 4) может быть образована базисами из точек

$(P_1, 6, P_2)$, $(P_1, 7, P_2)$, $(P_1, 5, P_2)$, $(P$

и т. д.

Выбирая эти наборы базисных точек в качестве исходных данных при метрическом подходе, мы будем в общем случае получать разные групповые решения при неизменном согласии всех наборов в смысле принципа Парето.

Вторая особенность состоит в следующем. В примере А1 при метрическом подходе можно было выбрать решения как принадлежащие, так и не принадлежащие ядру. Рассматриваемый же набор исходных данных дает нам пример того, что множество решений при метрическом и геометрическом подходах могут не пересекаться: ни одна из точек 5, 6, 7, составляющих в пространстве QO множество возможных групповых решений при метрическом подходе, не принадлежит ядру. Пример Б1. Исходные данные этого примера приведены в таблице 6.5. Выпуклые оболочки исходных точек приведены на рис. 6.4 и 6.5.

Возможные групповые решения при метрическом подходе и точки для данных примера Б1 представлены в таблице 6.6. Характеристики условий для выбора группового решения на первом уровне собраны в таблице 6.7.Таблица 6.5

Отметим, что расстояние между исходными точками $(d\{R_u\ R_2)$ == 4) и само расположение точек по сравнению с примером А1, как это видно из сравнения выпуклых оболочек на рис. 6.3 и 6.5, изменились незначительно, а размерность ядра в

t	1	2
R_i	$(a—c,\ b)$	$(a—b,\ c)$
P_i	$\{(b,\ a),\ (b,\ c)\}$	$\{(c,\ a),\ (c,\ b)\}$

обоих пространствах увеличилась. Условия выбора группового решения сохраняют двухуровневое^ и аналогичны условиям примера А1. При комплексном подходе количество решений на первом уровне может быть сокращено до двух в каждом пространстве.

Таблица 6.6

Точки			$\mathcal{PO}$	
			min Σd	min
1	2	3	4	5
Выпуклой оболочки	Ядра	$P_{\min}$ 3, 4 $P_{К}$	* * *	* *
	Ис-ход-ные	P_1 P_2	* *	
		1, 2 5, 6	* *	

Пример Б2. Пусть, как и в примере А2, точки выпуклой оболочки в пространстве $\mathcal{PO}$ будут исходными точками при поиске группового решения в пространстве $\mathcal{QO}$. Групповые решения при обоих подходах представлены в таблице 6.8. Из таб-$_{л}$ лицы 6.8 видно, что в обоих пространствах все медианы и все средние являются одновременно и точками ядра. В пространстве $\mathcal{QO}$ проблема выбора группового решения при комплексном под-* ходе решается однозначно выбором точки 7? к •Таблица 6.7

		В пространств
Размерность ядра		2
Число точек в ядре		4
Возможные групповые решения при метрическом подходе (1-й уровень)	Медианы	Все 10 точек лой оболочки 4 ядерные)
	Средние	2 (ядерные) точ
Допустимые групповые решения при геометрическом подходе (1-й уровень)		4 точки

Таблица 6.8

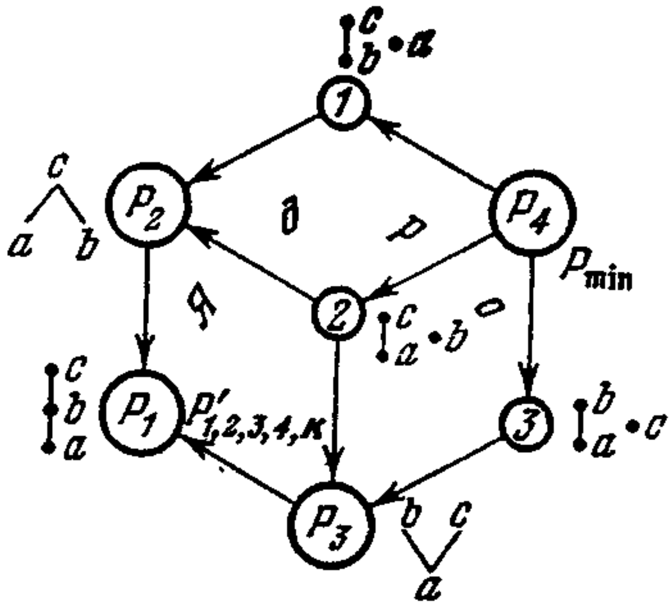
Точки			$\mathcal{PO}$	
			min Σd	min
1	2	3	4	
Исходные и выпуклой оболочки	Ядра́	$P_{\min}$ 3, 4 $P_{\text{К}}$	* * *	
		P_1 P_2	[illegible]	
		1, 2 5, 6		

Пример В1. Пусть исходные данные имеют вид
Таблица 6.9

i	1	2
R_i	(a, b, c)	$(a—b, c)$
P_i	$\{(c, b), (c, a), (b, a)\}$	$\{(c, a), (c, b)\}$

Выпуклые оболочки исходных данных приведены на рис. 6.6 и 6.7. Как видно из рис. 6.7, исходные данные в пространстве

$\mathscr{P}O$



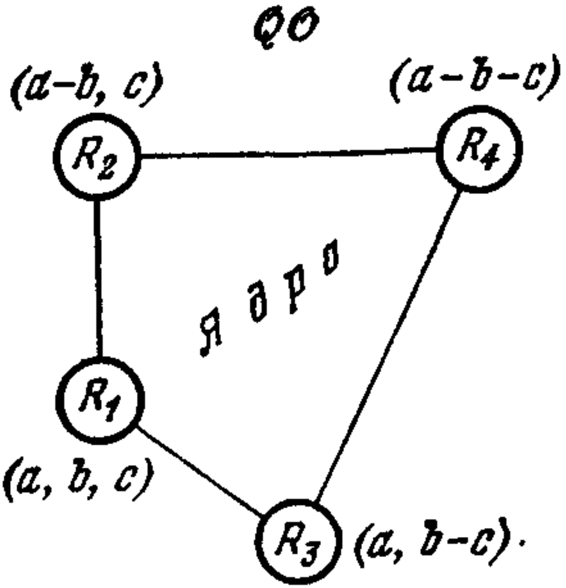


Рис. 6.6. Рис. 6.7.

QO составляют ядро и выпуклую оболочку одновременно. Размещение медиан и средних показано в таблице 6.10.

Таблица 6.1

Точки			$\mathcal{PO}$	
			min Σd	min
1	2	3	4	
Ядра́ и выпуклой оболочки	Исходные	1 P_1 P_2, P_3 P_4	 * *	
		1, 3 2	 *	

Таким образом, в пространстве $\mathcal{QO}$ любая исходная точка при геометрическом подходе может быть выбрана в качестве группового решения (см. табл. 6.11), а в $\mathcal{PO}$ —любая из семи точек выпуклой оболочки. Дополнительный учет критериев метрического подхода сокращает число решений на первом уровне в $\mathcal{QO}$ до двух-трех точек, а в $\mathcal{PO}$ —до четырех $(P_2,\ P_3,\ P_4,\ 2)$ или даже одной точки (2).

Пример В2. Размещение медиан и средних для случая, когда точки выпуклой оболочки в пространстве $\mathcal{PO}$ из примера В1 принимаются в качестве исходных для поиска решений в пространстве $\mathcal{QO}$, показано в таблице 6.12. Как видно из этой таблицы, на роль группового решения в пространстве $\mathcal{PO}$ при любом подходе может претендовать только одна точка 2. В пространстве же QO, естественно, условия выбора группового решения не изменились. Таблица 6.11

		В п
Размерность ядра		
Число точек в ядре		
Возможные групповые решения при метрическом подходе (1-й уровень)	**Медианы**	4 яд
	Средние	точ
Допустимые групповые решения при геометрическом подходе (1-й уровень)		Т

Таблица 6.12

Точки		$\mathcal{PO}$	
		min Σd	mi…
1, 2	3	4	5
Исходные, ядра́ и выпуклой оболочки	P_1		
	P_2, P_3		
	P_4		
	1, 3		
	2	*	*

Пример П. Пусть исходные данные имеют вид

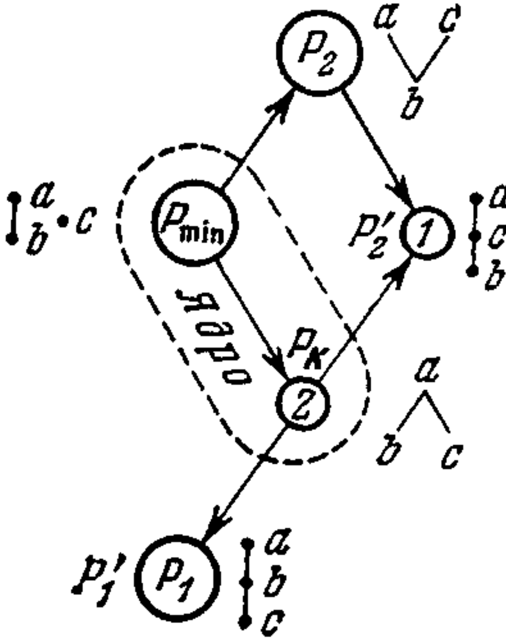
Таблица 6.13

,,,[1]

i	1	
R_i	(c, b, a)	$(b, a—c$
P_i	$\{(a, b), (a, c), (b, c)\}$	$\{(a, b),$

Выпуклые оболочки для исходных данных приведены на рис. 6.8 и 6.9. Обратим внимание на то, что в данном примере в пространстве QO (см. рис. 6.9) две точки (1 и 2) расположены одинаково относительно исходных точек. Однако точка 1 не попадает в число ядерных вследствие того, что в ней отношение на паре объектов (Ь, *c)* «тяготеет» к исходной точке 7?$_2$.

РО



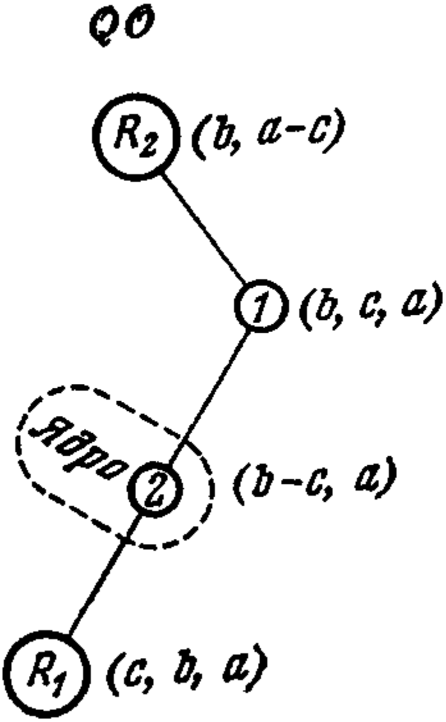


Рис. 6.8. Рис. 6.9.

Из таблицы 6.14 видно, что, подобно тому, как это было в примерах А1 и Б1, в обоих пространствах все точки выпуклой оболочки являются медианами, а точки ядра — к тому же и средними.

Таблица 6.14

Точки			$\mathcal{PO}$	
			min Σd	mi…
1	2	3	4	5
Выпуклой оболочки	Ядрá	$P_{\min}$	*	*
		2	*	*
	Исходные	P_1	*	
		P_2	*	
		1	*	*

Из приведенных в таблицах 6.14 и 6.15 данных следует, что при комплексном подходе число решений на первом уровне в $\mathcal{PO}$ ограничено двумя точками (Pmm и 2), а в $\mathcal{QO}$ — одной точкой (2).

Пример Г2. Размещение медиан и средних для последнего примера, когда точки выпуклой оболочки в пространстве $\mathcal{PO}$ изпримера Г1 являются исходными данными для поиска решений в пространстве $\mathcal{QO}$, показано в таблице 6.16.

В пространствах $\mathcal{PO}$ и $\mathcal{QO}$ проблема выбора группового ре-
шения на *1-й* уровне не вызывает никаких затруднений при комплексном подходе: таковым несомненно должна быть точка *2.*

Таблица 6.15

		В простран
Размерность ядра		1
Число точек в ядре		2
Возможные групповые решения при метрическом подходе (1-й уровень)	Медиана	Все 5 точе лой оболочк 2 ядерные)
	Средняя	3 точки 2 ядерные)
Допустимые групповые решения при геометрическом подходе (1-й уровень)		2 точки

Таблица 6.16

Точки			$\mathcal{PO}$	
			minΣd	**min**
1	2	3	4	5
Исходные и выпуклой оболочки	Ядра́	$P_{\min}$ 2	*	*
		P_1 P_2		
		1		

Эти примеры дают нам большинство возможных вариантов соотнесения ядерных точек, с одной стороны, и медиан и средних — с другой. Так, например, групповые решения при геометрическом подходе могут полностью содержаться в решениях метрического подхода (в $\mathcal{PO}$ —примеры А1, Б1, И; в $\mathcal{QO}$ — примеры А1, Б1, И, Г2) и наоборот (в $\mathcal{PO}$ — примеры В1, В2; в *Q0* — пример В2), могут полностью совпадать (в $\mathcal{PO}$ —примеры А2, Б2, Г2; в $\mathcal{QO}$ — пример Б2), могут не иметь общих решений (в $\dot{\mathcal{Q}}\mathcal{O}$ — пример А2).

Из непосредственного сравнения примеров в каждой паре видно, что при метрическом подходе разнообразие (увеличение числа) исходных мнений (точек) уменьшает число возможных групповых решений, в то время как при геометрическом подходе число допустимых групповых решений определяется, вообще говоря, расположением «самых крайних» мнений. Поясним это замечание на примере А2. В примере

А2 перечислены несколько базисов, образующих одну и ту же выпуклую оболочку (рис. 6.2). Первые два из них, а именно (Л, *Рг)* и (1, 4), являются минимальными по числу содержащих точек и составлены из таких «крайних» мнений: какие бы мнения из образуемой ими выпуклой оболочки (т. е. лежащие между этими крайними) мы не добавляли в эти базисы, ни выпуклая оболочка, ни, соответственно, ядро не вменятся.

Рассмотренные примеры являются также иллюстрацией к тому факту, что размерность ядра, и, следовательно, число точек в нем, связана не с максимальным расстоянием между исходными мнениями (точками), а с их взаимным расположением, определяющим конфигурацию выпуклой оболочки. Иллюстрацией этому служат примеры А1 и Б1: несмотря на то, что расстояние между исходными точками во втором примере меньше, чем в первом, размерность ядра во втором примере больше.

Суммируя сделанные замечания, можно сделать вывод о различной природе факторов, определяющих множества групповых решений при обоих подходах, и невозможности на данном уровне исследований установить строгие зависимости между результатами применения сравниваемых подходов.

Сравнение условий, характеризующих проблему группового выбора при обоих подходах, приводит нас к заключению, что комплексное применение геометрического и метрического подходов может, по-видимому, стать действенным средством для выделения хорошо обоснованных групповых решений, попадающих на второй уровень, где из множества допустимых групповых решений выбирается единственное решение. Глава VII

НЕКОТОРЫЕ ВОПРОСЫ ПРАКТИЧЕСКОГО ПРИМЕНЕНИЯ ГЕОМЕТРИЧЕСКОГО ПОДХОДА

§ 7.1. Анализ экспертиз НИР

Алгоритм «Ядро» был использован для анализа трех экспертиз НИР, проводившихся последовательно друг за другом на ежегодных научных конференциях. Экспертное оценивание НИР и обработка оценок производились по методике [27J, учитывающей специфику набора управляющих воздействий, применявшихся по результатам экспертиз. Качество НИР характеризовалось по двум аспектам, каждый из которых описывался 3—4 признаками.

Оценки по признакам выставлялись в специальным образом устроенных балльных шкалах. Полезно различать два вида балльных оценок: к первому виду относятся оценки, производимые при наличии объективного критерия, ко второму —когда не только нет общепринятых эталонов, но и сомнительно даже наличие некоего единственного объективного критерия, субъективным отражением которого являются оценки. Во втором случае оценки рассматриваются выполненными в шкале порядка. С этой точки зрения на основании балльных оценок составлялись совокупности упорядочений, которые затем обрабатывались программой «Ядро». Результаты этой обработки сравнивались с упорядочениями, соответствующими агрегированным показателям НИР или по аспектам, или в целом, или средним оценкам по одному из признаков.

Первая экспертиза НИР. По результатам первой экспертизы были составлены 18 совокупностей из пяти (по числу экспертов) упорядочений каждая, полученных на основе оценок НИР по трем признакам в шести научных секциях. В каждой секции докладывалось в среднем по 11 работ. Анализ этих совокупностей имел целью выяснить, имеется ли среди экспертов согласованность в смысле принципа Парето по использованным для оценки признакам. Единственное невырожденное групповое решениеимело вид 22221222222*), т. е. представляло собой разбиение 11 объектов на два упорядоченных класса, причем класс наиболее предпочтительных объектов состоит из одного объекта. Практически во всех комиссиях по всем трем признакам оказалось невозможным построить искомые групповые решения.

По результатам этой экспертизы исследовался также вопрос о том, как усреднение экспертных оценок влияет на согласованность экспертных суждений в смысле принципа Парето. По каждому из трех признаков в отдельности были рассчитаны средние (по числу экспертов в секции) арифметические балльные оценки для каждого объекта. На их основе были составлены шесть совокупностей из трех упорядочений каждая. Результаты обработки приведены в таблице 7,1. В этой таблице для каждой секции в первой строке указаны отношения из ядра, а во второй — отношение, соответствующее агрегированным показателям, принимавшееся за «истинное».

Таблица 7.1

Комиссии	Объек						
	1	2	3	4	5	6	7
I	3 3	5 8	5 6	1 1	3 4	9 9	5 6
II	1 1	3 6	3 8	3 9	3 7	3 5	3 4
III	2 4	2 5	2 11	2 12	2 3	2 10	1 1
IV	2 7	2 9	2 2	2 3	2 4	1 1	2 6
V	2 8	2 3	2 5	2 4	1 1	2 10	2 9
VI	2 6	2 2	2 4	2 10	2 3	1 1	2 9

Приведенные в таблице 7.1 допустимые групповые решения представляют собой в среднем разбиение оцениваемой совокупности объектов на 2-3 класса. Наиболее «тонко различающие» объекты групповое решение, составляющее ядро экспертных суждений первой комиссии, разбивает объекты на пять классов. Оно хорошо согласуется с упорядочением, полученным по агрегированным оценкам. Это может свидетельствовать или о том, что в дайной комиссии суждения экспертов по отдельным признакам

*) Здесь, и далее в этой главе, в такой записи цифра на £-м месте указывает ранг j-ro объекта. определялись совокупными (агрегированными) достоинствами работ, или связано с тем, что состав

работ сам по себе таков, что достоинства работ по агрегированному показателю монотонно связаны с частными признаками (а экспертная комиссия в целом довольно чутко отреагировала на это).

Отметим характер использования алгоритма в последнем случае: упорядочения, составившие ядро, представляют собой, по существу, новый фактор, агрегирующий осредненные экспертные предпочтения по трем «чистым» признакам. При этом упорядочения, соответствующие частным признакам, являются «вариациями» * упорядочения, составленного по агрегированным показателям.

Вторая экспертиза НИР. По результатам проведенного анализа в методику экспертного оценивания был внесен ряд изменений. Одно из них было связано со способом назначения балльных оценок и имело целью создать предпосылки для одинаковой «настройки» экспертов, т. е. для приведения их систем субъективных ценностей к некоторой, по возможности одинаково понимаемой системе. С этой целью во второй экспертизе на шкалах задавались оценки «эталонных» объектов, одним из которых, например, служило представление экспертов о типичной работе в определенной научной школе.

Оценивание НИР на этот раз производилось по четырем частным признакам и одному обобщенному признаку, т. е. по совокупному представлению о достоинствах НИР. Оценки по последнему признаку эксперты выражали сразу в виде ранжировки работ. Это позволило подвергнуть проверке гипотезу о том, что у

Таблица 7.2

Упорядочения по признакам	Объект			
	1	2	3	4
X	5	4	2	3
Y	3	3	2	2
W	4	3	3	2
Z	4	4	3	2
Ядро	5	5	2	2

эксперта в процессе работы складывается некоторое общее представление о совокупных достоинствах оцениваемых работ, которое находит отражение в ранжировке по обобщенному признаку и составляет ядро упорядочений по «чистым» признакам.

Для иллюстрации того, как проявляется этот феномен, приведем один пример. В верхней части таблицы 7.2 выписаны упоря-дочения по четырем частным признакам, а в нижней, отделенной чертой части,— соответствующее им ядро. Восьмой объект в этой комиссии был третьим эталонным объектом («типичная работа»). Как видно из самой таблицы, упорядочения объектов по частным признакам у этого эксперта варьируются вокруг упорядочения, представляющего собой разбиение оцениваемого набора объектов на три упорядоченных класса эквивалентных друг другу объектов.

Таблица 7.3

Упорядочения по признаку	Объе			
	1	2	3	4
X	5	4	2	3
Y	3	3	2	2
W	4	3	3	2
Z	4	4	3	2
Обобщенному	6	5	2	4
Ядро	5 6	5 5	2 2	2 2

Теперь обратимся к таблице 7.3. В ней по сравнению с данными таблицы 7.2 добавлено еще одно упорядочение оцениваемых работ по совокупному представлению об их достоинствах, полученное от того же эксперта. Ранги эталонного объекта (8) отсутствуют, поскольку он не принимался во внимание при оценивании по обобщенному признаку. Ядро теперь состоит из двух упорядочений. Первое совпадает с упорядочением, составляющим ядро при четырех признаках (табл. 7.2). Второе упорядочение как бы «уточняет» первое и представляет собой разбиение, уже довольно тонко «различающее» оцениваемые НИР: 7 работ разбиваются на 4 упорядоченных класса. Из непосредственного оценивания видно, что это упорядочение очень «близко» к упорядочению по обобщенному признаку.

Обратимся теперь к таблице 7.4. В этой таблице приведены невырожденные ядра по результатам обработки экспертных оценок в комиссии № 3. В ней выписаны четыре упорядочения, каждое из которых составило ядро соответствующей совокупности упорядочений. Первые три упорядочения получены по оценкам, которые три эксперта проставили НИР по четырем частным признакам.

Четвертое упорядочение составляет ядро по оценкам НИР всеми экспертами комиссии по одному частному признаку.

Как видно из этой таблицы, в «середине» исследованных совокупностей экспертных суждений лежит довольно «грубое» разбиение оцениваемого набора объектов на два упорядоченныхкласса эквивалентных друг другу объектов. Аналогичные результаты получены по данным комиссии № 2.

Проведенный анализ экспертных оценок экспертизы НИР показал общее повышение примеров согласия в смысле принципа Парето, и, хотя не все исследованные совокупности упорядочений

Таблица 7.4

Ядра	1	2	3	4
Эксперт № 2	3	1	3	1
Эксперт № 3	3	1	3	1
Эксперт № 4	2	2	3	1
Признак *X*	3	1	3	1

дали интересные, невырожденные решения, однако даже в экспертизе, по своей методике не ориентированной на получение групповых решений, удовлетворяющих принципу Парето, мы наблюдаем интересные примеры согласованности в исследуемом здесь смысле.

Третья экспертиза НИР. Если на данных второй экспертизы геометрический подход применялся для проверки гипотезы о механизме многопараметрического индивидуального оценивания, то данные последней экспертизы были использованы для дальнейшего исследования механизма «группового» оценивания.

На основе агрегированных показателей научно-исследовательских работ, подсчитанных для восьми комиссий, было составлено 16 групп (по 8 для каждого из двух аспектов) — в среднем из семи упорядочений каждая. Только одна группа упорядочений дала искомое решение. Таким образом, совокупности агрегированных показателей, как и в предыдущих экспертизах, оказались непригодными для построений коллегиальных решений, удовлетворяющих принципу единогласия Парето.

Другая картина наблюдается при анализе упорядочений, соответствующих средним оценкам признаков для каждого аспекта в отдельности. Как видно из таблицы 7.5, в большинстве случаев были получены коллективные решения.

Близкие к приведенным в таблице 7.5 результаты были получены (табл. 7.6) при анализе пар упорядочений, полученных по агрегированным показателям с учетом стандартных отклонений для каждой работы, рассчитанных по формуле

$$\sigma(n, j) = \sqrt{\frac{1}{m_0} \sum_{m=1}^{m_0} [\omega(n, j, m) - \ldots}$$

где $\omega(n, j, m)$, _ — согласованная по эталону оценка тг-й работы по /-му признаку у *т-то* эксперта, $\omega(n, j)$— коллективная оценка по признаку, m_0 — число экспертов в данной комиссии. Эти данные свидетельствуют о том, что учет стандартных отклонений

Таблица 7.5

Комиссия	Решения, получе
	Аспекта I
1	1 1 1 1 1 1 1 1 9
2	1 2 2 2 2 2 2 2
3	—
4	1 1 1 1 1 1 1 1 1 1 1 12
5	1 2 3 3 5 5
6	—
7	—
8	1 1 3 3 3 3 7

Таблица 7.6

Комиссия	Решения, полученные по агре
	Аспект I
1	1 1 3 4 5 6 7 7 9
2	1 2 3 3 4 4 4 8
3	1 2 3 4 4 4 4 4 4 4 4 4 4 4
4	1 1 1 4 4 4 4 4 4 4 4 4
5	1 2 2 2 2 6
6	
7	2 1 1 1 1 1 1 1 1 10 10
8	1 1 3 3 3 3 7

дает возможность получать искомые групповые решения, хорошо согласующиеся с упорядочениями, соответствующими агрегированным (по аспектам) показателям.

§ 7.2. Процедура выработки группового решения

Практика и теория экспертного метода имеют весьма обширную библиографию. Достаточно указать, что только один краткий обзор [28] содержит список 313 работ, явно не охватывающий всей относящейся к этому вопросу литературы.

Выбор конкретного метода экспертных оценок зависит от характера решаемой практической задачи и ряда условий, которые при этом нужно соблюдать. Общие вопросы, связанные с такимвыбором детально обсуждаются в упомянутом обзоре. Здесь мы кратко остановимся на одном аспекте, связанном с методикой организации экспертного оценивания.

Методики проведения экспертиз, как правило, не содержат в себе операций (тестов, процедур и т. п.), выполнение которых гарантировало бы (хотя бы с достаточной степенью уверенности), что групповое решение будет удовлетворять тем или иным, заранее оговоренным требованиям, обладать заранее обусловленными свойствами. Обстоятельство это, как представляется, вырастает в большую и самостоятельную проблему. В практике экспертного метода нередко встречаются ситуации, когда практически ни один из имеющихся в распоряжении исследователя методов обработки не дает удовлетворительных результатов. Например, результаты применения методов факторного анализа или многомерного шкалирования не поддаются разумной интерпретации; мажоритарные правила не дают транзитивных отношений; распределение частот рангов близко к равномерному; селекция экспертов с целью формирования «согласованной» группы не позволяет отобрать нужного минимального количества экспертов; ядра состоят только из пустых отношений и т. д. Только «механическое» правило нахождения группового, решения в метрическом подходе всегда работает безотказно: любая из просмотренных тем или иным способом точек, дающая минимум суммы расстояний или квадратов расстояний до исходных точек, может быть принята в качестве группового решения.

По своему существу рассматриваемая ниже идея использования эталонных объектов очень близка к механизму, обусловливающему существование невырожденных решений. Последнее означает, что во множестве рассматриваемых объектов существует такое его подмножество, которое мы будем называть *эталонным,* что предпочтения экспертов на нем полностью совпадают. При этом все эксперты согласны так же и с тем, что остальные, не входящие в это множество элементы, одинаковым образом группируются в классы эквивалентности, а последние одинаковым образом располагаются (упорядочиваются) между элементами эталонного множества.

В плане дальнейшего развития геометрического подхода эти соображения были использованы для построения процедуры экспертного оценивания, которая способствовала бы выработке группового решения, удовлетворяющего принципу Парето.

Процедурой выработки решений будем называть конечную последовательность операций (операций в широком смысле этого слова), ориентированных на получение решений, обладающих наперед заданными свойствами. В терминах геометрического подхода процедура выработки решения должна быть ориентирована на получение в общем случае невырожденных групповых отно-шений, удовлетворяющих принципу Парето на согласование индивидуальных предпочтений.

Рассмотрим одну из возможных процедур. Мы будем различать в этой процедуре итерации и шаги. Каждая итерация состоит из двух шагов. Номер итерации будет обозначаться индексом *s* над символами отношений и множеств.

5-я итерация. На первом шаге экспертам предъявляется все множество объектов $\mathbf{A} = \{a_1, a_2, \ldots, a_n\}$_, _. , для выработки инди-
видуальных предпочтений, .и в результате этого предъявления в пространстве QO получается некоторое множество точек X^e — множество отношений индивидуального предпочтения. Ядро этого множества X^e обозначим через $\mathbf{K}(\mathbf{X}^s)$, а через $R^s_{\mathbf{max}}$ — макси-
мальное отношение из K(X). Отношение $R^s_{\mathbf{max}}$ представляет собой
упорядоченное разбиение оцениваемой совокупности объектов на несколько классов эквивалентности, в том числе $R^s_{\mathbf{max}}$ может быть вырожденным отношением, т. е. полной эквивалентностью. Пусть теперь мы имеем произвольное невырожденное отношение линейного квазипорядка R^s: $R^s \subseteq R^s_{\mathbf{max}}$. Отношение R* будем называть *опорным* отношением. Рассмотрим отношение безразличия для R^s, которое будем обозначать через $I_{\mathbf{R}^s}$. Этому
отношению соответствует разбиение множества объектов А на классы: $\mathbf{A} = \mathbf{A}^s_1 \cup \mathbf{A}^s_2 \cup \ldots \cup \mathbf{A}^s_l$, где l — число классов.

На втором шаге 5-й итерации процедуры каждому эксперту классы объектов $\mathbf{A}^s_1, \ldots, \mathbf{A}^s_l$ предъявляются раздельно, и на
каждом из них эксперты определяют свои предпочтения. Таким образом совокупность предпочтений 1-го эксперта на 5-м шаге может быть представлена набором предпочтений на каждом из классов разбиения множества $\mathbf{A}$: $(R_j)^s_i = (R_1, \ldots, R_l)^s_i$.

Далее возможны два пути. Первый, наиболее трудоемкий по объему вычислений, состоит в том, что дальнейшему анализу подвергаются отношения, представляющие собой комбинацию предпочтений различных экспертов на данных *I* классах:

$$R_{i_1 i_2 \ldots i_l} = R_{i_1} \cup R_{i_2} \cup \ldots \cup R_{i_l} \cup (R^s \setminus$$

Очевидно, что $R_{i_1 i_2 \ldots i l} \subseteq$

$\subseteq R^s.$ Всего таких отношений может быть составлено то[1],'где *т* — число экспертов. Полученный набор отношений принимается за исходное множество $\mathbf{X}^{s+1}$ для первого шага последующей итерации.

Второй путь состоит в том, что в качестве элементов множества $\mathbf{X}^{s+1}$ принимаются то упорядочений $(R_j)^s,\ j = 1,\ 2,\ \ldots,\ l,$ каждое из которых составлено из предпочтений одного из экспертов на всех классах разбиения множества A^s.

(5+1)-я итерация начинается с построения ядра множества отношений индивидуального предпочтения X^{s+1}, полученных на втором шаге предыдущей итерации, а затем повторяются всеизложенные выше операции. Формально процедура заканчивается при выполнении условия $R^{s+1} = R^s_{\max}.$

Остановимся более подробно на понятии опорного отношения. Требование наличия невырожденного решения означает или совпадение экспертных предпочтений относительно, как минимум, одной пары объектов из А, или то, что «в среднем» эксперты считают два объекта неразличимыми. Использование опорного отношения призвано способствовать выполнению этого условия. Выбор этого отношения может производиться с учетом специфики решаемой задачи и степени согласованности экспертов на первом шаге каждой итерации. Кроме того, опорное отношение можно задавать извне, т. е. на объектах, которые в общем случае непосредственно не входят во множество А подлежащих экспертизе объектов. Упорядочение объектов опорного отношения в этом случае может определяться, например, предпочтениями лица, принимающего решение; объективными данными, возможно даже теми, которые не учитываются в самой экспертизе; результатами подобных же экспертиз, в которых согласованность экспертов оказалась удовлетворительной для организаторов экспертизы; вспомогательной операцией, как, например, предваряющее экспертизу разбиение оцениваемых объектов на 2-3 упорядоченных класса, и последующим формированием опорного отношения из 2-3 объектов, попавших в пересечение (по разбиениям всех экспертов), и т. п.

Нужно отметить, что наличие тем или иным образом сформированного опорного отношения полностью не гарантирует появление невырожденных групповых решений. Это обстоятельство может явиться, например, следствием того факта, что объекты опорного отношения, будучи строго упорядочены между собой, оцениваются экспертами как неразличимые в сравнении с частью объектов из оцениваемой совокупности, или предпочтения некоторых экспертов на парах, содержащих объекты из R^s и из $\mathbf{A} \backslash R^s$ диаметрально противоположны. В этом случае, учитывая специфику задачи, можно или сменить объекты в опорном отношении, или, оставив те же объекты, подкорректировать порядок на них, или, наконец, заменить часть экспертов.

В практических приложениях встречаются такие ситуации, когда до экспертизы или уже после нее становится очевидным, что множество оцениваемых объектов или состав экспертов неоднородны или и то и другое одновременно. Иллюстрацией этого может служить, например, экспертиза НИР в одной тематической комиссии, когда, скажем, одна часть работ оказалась чисто теоретического характера, а другая — прикладного. К тому же по пристрастиям к работам того или иного характера можно, очевидно, разделить и экспертов в комиссии. Не останавливаясь на методах выявления этих факторов, предположим, что они намиустановлены. В этом случае представляется разумным на классах разбиения всей совокупности оцениваемых объектов искать, «под-ядра», т. е. ядра индивидуальных предпочтений для каждой в отдельности группы экспертов-единомышленников. Групповые решения каждой группы (подядра) могут быть объединены опорным отношением. Возможность реализации этой идеи определения подядер заложена в первом из путей комплектования наборов предпочтений, рассматриваемых на втором шаге процедуры выработки решения.

Ниже описываются две конкретные процедуры выработки решения, первая из которых носит экспериментальный характер, а вторая была применена при определении признаков для оценки методов прогнозирования выходного потока на обогатительной фабрике и выбора самого метода прогнозирования.

Процедура I. Оценка художественных произведений. Объектами оценки в первой процедуре выработки решений служили литературные произведения детективного жанра. Цель экспертизы состояла в том, чтобы от данной группы экспертов получить групповое упорядочение оцениваемых объектов по общему впечатлению об их достоинствах. В группу оцениваемых объектов было подобрано 8 произведений, группа экспертов состояла из 7 человек. Предварительная экспертиза отобранных произведений дала набор упорядочений (таблица 7.7), на основе которого оказалось невозможным получить невырожденное групповое решение
с нужными свойствами.

Таблица 7.7

Эксперты	Объекты				
	1	2	3	4	5
1	1	2	3	4	4
2	1	2	4	3	4
3	1	3	1	3	2
4	1	2	5	3	4
5	2	1	4	3	3
6	2	2	1	4	3
7	2	1	1	4	5

Для построения опорного отношения был составлен набор из 7 произведений с таким расчетом, чтобы в нем были как явные «лидеры», так и явные «аутсайдеры». Объекты этой группы упорядочивались также по общему впечатлению. В таблице 7.8 приведены экспертные оценки дополнительной группы объектов и их ядро.

На основе групповых решений, полученных на объектах дополнительной группы, было составлено опорное отношение из четырех объектов: $a > c > e > g$. На первом шаге процедуры экспертам предъявлялись все 8 объектов оцениваемой совокупности А

= {1, 2, 3, 4, 5, 6, 7, 8), и сообщалось опорное отношение, между объектами которого эксперты должны были расставлять объекты из А. В таблице 7.9

Таблица   7.8

Эксперты	Объе			
	a	*b*	*c*	*d*
1	1	2	2	2
2	1	2	2	2
3	1	2	2	2
4	1	3	3	2
5	1	3	3	2
6	1	2	3	4
7	1	4	2	3
Ядро	1	2	2	2
	1	2	2	2

приведены экспертные оценки, полученные после этого предъявления, и их ядро (буквами в таблице обозначаются объекты опорного отношения). Как видно из этой таблицы, наиболее «тонко различающее» оцениваемые объекты упорядочение, помещенное

Таблица   7.9

Эксперты	Объ…					
	a	1	2	c	3	4
1	1	2	3	4	5	6
2	1	1	4	3	5	6
3	1	2	4	3	5	7
4	1	2	4	5	3	7
5	1	2	4	5	3	6
6	1	2	3	5	4	6
7	1	1	4	5	3	7
Ядро	1 1	1 2	3 3	3 3	3 3	6 6

в последней строке таблицы 7.9, представляет собой разбиение 12 объектов на 5 классов:

(•)

$$(a) \succ (1) \succ (2,\ 3,\ c) \succ (4,\ 5,\ 6,\ e) \succ (7$$

Без объектов опорного отношения оба решения из ядра имеют вид:

$$(1) \succ (2,\ 3) \succ (4,\ 5,\ 6) \succ (7,\ 8).$$

На классах (2, 3, *c)* и (4, 5, ,6, *e)* разбиения (*) была проведена следующая итерация. Поскольку внутри каждого класса объекты оказались неразличимыми по общему впечатлению, то для каждого класса были введены дополнительные признаки: объекты первого класса оценивались по «занимательности сюжета», а второго— «по убедительности метода раскрытия преступления». Ядро предпочтений

для первого класса «выделило» на первое место объект 2, а для второго имело вид $4 \sim 5 > 6 \sim e.$ Окончательное групповое решение имело вид $1 > 2 > 3 > 4 \sim 5 > 6 > 7 \sim 8.$
Таким образом, к характерной особенности этой процедуры можно отнести использование дополнительных объектов при построении опорного отношения и введение дополнительных признаков на второй итерации для увеличения «разрешающей способности» группового решения, получаемого на второй итерации.
Процедура II. Выбор характеристик. Объектами оценки во второй процедуре служили восемь характеристик методов прогнозирования. Учитывая относительно большое количество характеристик, а, следовательно, и трудность сравнения алгоритмов по стольким показателям, а также различную природу самих характеристик (от чисто количественных, объективных, таких, как «точность прогноза», до качественных, таких, как «эксплуатационные характеристики») была проведена экспертная оценка важности характеристик с целью выделения наиболее важных из них. Оценка производилась следующим образом. Экспертам предлагалось разбить все характеристики на две группы, помещая в первую из них наиболее важные характеристики, а во вторую — наименее важные. Результаты этого этапа экспертизы приведены
в таблице 7.10.
Таблица 7.1

Эксперты №№	Характерис	
	Наиболее важные (группа 1)	
1	3, 5	1
2	1, 2, 3, 5, 6	4
3	3, 4, 5	1
4	3, 5, 6	1
5	2, 3, 4, 5, 6, 8	1
6	1, 3, 5, 6, 8	4
7	2, 3, 5, 6	1
8	3, 4, 5, 6	1

Непосредственный анализ экспертных оценок таблицы 7.9 показывает, что данный набор характеристик можно разбить на три группы. К первой относятся характеристики *5* и 5, признаваемые всеми без исключения экспертами как наиболее важные, к тре-тьей — характеристика 7, отнесенная всеми экспертами в группу менее важных, и ко второй группе — характеристики 1, *2, 4, 6,* S, относительно которых мнения экспертов разошлись. В силу этого из каждой группы было выбрано по одной характеристике, из которых составилась эталонная группа из трех упорядоченных по важности характеристик $(5 > 4 > 7)$. На втором этапе экспертизы каждый эксперт ранжировал характеристики, расставляя их между эталонными. Результаты этого этапа экспертизы приведены в таблице 7.11.

Таблица 7.11

Эксперты №№	Характери			
	3	1	2	4
1	1	5	6	4
2	1	6	5	3
3	1	4	5	2
4	1	6	7	3
5	1	5	6	2
6	1	5	6	3
7	1	4	5	2
8	1	4	5	2
Ядро	1 1	4 4	4 4	3 3

Эта группа экспертных оценок была обработана с целью выделения ядра. В результате получено два упорядочения, помещенных в последних строках таблицы. Таким образом, в числе наиболее важных оказались характеристики 5, 5, *4* и *6*.

§ 7.3. Обсуждение

Анализ экспертных оценок НИР дал нам примеры существования искомых решений. Однако, многие из исследованных совокупностей экспертных суждений оказались таковы, что построить на их основе групповое решение с заданными свойствами не оказалось возможным. В практических задачах, когда выбрано правило построения группового решения, такая ситуация часто может признаваться недопустимой: решение, удовлетворяющее данному принципу согласования экспертных суждений, должно быть получено обязательно. Поэтому и был поставлен вопрос о поиске путей, которые если бы и не гарантировали полностью, то в достаточной мере способствовали бы получению искомого решения.

В специальной литературе подробно обсуждаются естественные, «разумные» требования, характеризующие конкретные прин-ципы согласования индивидуальных предпочтений, и новое теоретическое исследование этих требований не является целью данной работы. Обращение к этим требованиям продиктовано следующим обстоятельством.

Общая теория аксиоматического подхода к проблеме согласования отношений дает мало рекомендаций для практического построения группового предпочтения. Так, рассмотрение некоторых наборов естественных требований приводит или к выводу о невозможности существования принципов согласования для этих требований (теорема Эрроу о «невозможности» [3D, или к тому, что им могут удовлетворять только тривиальные решения.

Существует несколько возможностей для выхода из этого положения. Один из них, например, указывается в [3, стр. 432J при разборе парадокса Эрроу: сохранить формулировку задачи и отказаться от одного из условий как слишком ограничительного. В этом плане считается наиболее целесообразным отказаться от так называемого требования «независимости объектов» (требование «несвязанных/альтернатив»). Вместо требования «независимости» предполагается возможным ввести в задачу разумно выбранные гипотетические объекты (альтернативы) в качестве базиса для оценки сравнительных степеней предпочтений.

ОСНОВНОЙ элемент предложенной процедуры выработки решений — опорное отношение, по-видимому, той же природы, что и «разумно выбранные гипотетические объекты». Отметим, однако, что использованные в двух практических задачах опорные отношения по своему характеру несколько отличаются друг от друга. Эти отличия порождены как спецификой оцениваемых объектов и характером предпочтений на них, так и различием целей экспертиз.

В первой задаче высказываемые экспертами предпочтения носят, если можно так сказать, крайне субъективный, «вкусовой» характер. Поэтому в этом случае естественно стремление получить такое групповое предпочтение, которое было бы по возможности максимально независимым от привнесения новых объектов, а также сохранить первоначальное максимальное согласие, которое было у экспертов до введения опорного отношения. Собственно говоря, удовлетворение последнего требования и может служить в этом случае некоторой проверкой корректности использованных в процедуре операций и основанием для того, чтобы рекомендовать полученное

групповое предпочтение для дальнейшего рассмотрения и формирования окончательного решения.
Непосредственное сравнение общих для всех экспертов предпочтений, полученных до (рис. 7.1) и после (рис. 7.2) процедуры, свидетельствует о том, что введение опорного отношения в данном случае не исказило тех предпочтений, которые совпадали у всех экспертов. Задача оценивания признаков для методов прогнозирования отличалась от первой тем, что в ней, во-первых, решение должно было быть обязательно получено, во-вторых, априори предполагалось, что индивидуальные предпочтения не должны быть сильно «окрашены» субъективными пристрастиями специалистов, и, наконец, нужно было минимизировать число обращений к экспертам.

	1	2	3	4	5	6	7	8
1	×			1	1	1	1	1
2		×		1		1	1	1
3			×				1	1
4				×			1	1
5					×		1	1
6						×	1	1
7							×	
8								×

	1	2	3	4	5	6	7	8
1	×	1	1	1	1	1	1	1
2		×		1	1	1	1	1
3			×	1	1	1	1	1
4				×			1	1
5					×		1	1
6						×	1	1
7							×	
8								×

Рис. 7.1. Рис. 7.2.

До начала экспертизы признаков было сделано предположение о невозможности получения согласованного решения с первого предъявления всех признаков. Поэтому, с учетом перечисленных выше условий задачи, использованная операция для построения опорного отношения — дихотомия исходного набора признаков — преследовала цель получить информацию об объектах, на которых предпочтения экспертов совпадают или разнятся, с тем чтобы полученные данные можно было использовать как основу для уточнения и согласования предпочтений экспертов на остальных объектах. Таким образом, опорное отношение было построено на подмножестве оцениваемой совокупности объектов и входило составной частью в суждение всех экспертов при последующих шагах процедуры.

Конкретные операции, используемые для построения опорных отношений, должны быть, конечно, практически оправданы и целесообразны с точки зрения особенностей решаемой задачи и данного правила нахождения группового решения. Очевидно, что их разнообразие не исчерпывается приведенными примерами.

В этой связи представляется интересным дальнейшее исследование различных операций для построения опорных отношений с точки зрения тех нарушений и модификаций, которые эти операции привносят в известные требования группового выбора. Приведем два примера возможных интерпретаций привносимых изменений. 1. Введение опорных отношений можно интерпретировать как сужение пространства предпочтений за счет отбрасывания отношений, в которых, для первой задачи — дополнительно вводимые объекты, а во второй — часть подлежащих оценке объектов упорядочены так же, как в опорном отношении.

2. Вводимые опорные отношения нарушают требование «не-навязанности» группового решения: в первой задаче — на множестве объектов, расширенном за счет объектов опорного отношения, во второй —- на исходном множестве. Правда, в последнем случае, опорное отношение можно считать в некоторой сте* пени «самонавязанным» на части оцениваемых объектов.

Теоретическое исследование подобных интерпретаций может способствовать появлению и обоснованию новых практических приемов введения опорных отношений.

Указатель обозначений

{ ... } — множество

$\in$ — принадлежность множеству

$\notin$ — непринадлежность множеству

$\subseteq$ —символ включения

$\subset$ —символ строгого включения

$\cup$ — объединение множеств

I — пересечение множеств
\ — разность множеств
× — прямое произведение множеств
$\overline{X}$ — дополнение множества X
X^s — степень множества
$\varnothing$ — пустое множество
I—универсальное множество
R—множество вещественных чисел
sup M — верхняя граница множества M
inf M — нижняя граница множества M
$\text{Пр}_i\ M$ — проекция множества на ось i
$(a_1 \ldots a_n)$ — упорядоченное множество (кортеж, вектор)
Λ — пустой кортеж
$\forall$ — квантор общности
$\exists$ — квантор существования
$\rightarrow$ — следствие, отображение
$f \circ g$ — композиция функций f и g по связке $\circ$
$\equiv$ — символ отношения эквивалентности, тождества
$\leq$ — символ отношения порядка
$<$ — символ отношения строгого порядка
? — символ отношения доминирования
$d(x, y)$ — расстояние между элементами множества
$\| \mathbf{x} \|$ — норма величины $\mathbf{x}$
E_n — евклидово n-мерное пространство

Литература

1. Горбатов В.А. Основы дискретной математики. - М.: Высшая школа, 1986.

2. Коршунов Ю.М. Математические основы кибернетики. - М.:Энергия, 1980.

3. Кузнецов О.П., Адельсон-Вельский Г.М. Дискретная математика для инженера. - М.:Энергия, 1980.

4. Кук Д., Бейз Г. Компьютерная математика.- М.: Наука, 1990.

5. Сигорский В.П. Математический аппарат инженера.-К.: Техника, 1977.

6. Кононюк А.Ю. Дискретна математика. В 2 ч. Ч.1, 2- К: 2009

7. Кононюк А.Ю. Вища математика. В 2 ч. Ч.1, 2- К: 2007.

8. Аверкин А.Н., Батыршин И.З. и др. Нечеткие множества в моделях управления и икусственного интелекта.- М.: Наука, 1986.

9. Кофман А. Введение в теорию нечетких множеств.- М.: Радио и связь, 1982.

10. Кофман А. Введение в прикладную комбинаторику.- М.: Наука, 1975.

11. Згуровский М.З. Интегрированные системы оптимального управления и проектирования.- К: Высшая школа, 1990.

13. Минский М. Фреймы для представления знаемый. - М.:Энергия, 1979.

14. Вильсон А. Дж. Энтропийные методы моделирования сложных систем.- М.: Наука, 1978.

15. В.Ю. Юрков, О.В. Лукина /Прикладная геометрия, вып. 8, N 18 (2006), стр. 9-36

16. Камени Д., Снелл Д., Томпсон Д. Введение в конечную математику.— М.: Изд-во иностр. лит., 1963.

17. Шиханович Ю. А. Введение в современную математику. — М.: Наука, 1965.

18. Клинни С. К. Введение в математику. — М.: Изд-во иностр. лит., 1957.

19. Карр Ч., Хоув Ч. Количественные методы принятия решений в управлении и экономике. — М.: Мир, 1966.

20. Феликс Л. Элементарная математика в современном изложении.— М.: Просвещение, 1967.

21. Берж К. Теория графов и ее применение. — М.: Изд-во иностр. лит., 1962.

22. Форд Л. Р., Фалкерсон Л. Р. Потоки в сетях. — М.: Мир, 1966.

16. Колмогоров А. Н., Фомин С. В. Элементы теории функций и функционального анализа. — М.: Наука, 1968.

17. Шрейдер Ю. А. Что такое расстояние?—М.: Физматгиз, 1963.

18. Шрейдер Ю. А. Равенство, сходство, порядок—М.: Наука, 1971.

19. Бронштейн И. Н., Семендяев К. А. Справочник по математике.— М.: Наука, 1964.

20. Справочник по системотехнике/Под ред. Р. Макола. — М.: Советское радио, 1970.