Парадигма развития науки

Методологическое обеспечение

А. Е. Кононюк

ДИСКРЕТНО-НЕПРЕРЫВНАЯ МАТЕМАТИКА

Книга 2

Множества

Часть 1

Четкие

Киев
«**Освіта України**»
2012


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
К213

Рецензенты:

В. В. Довгай — к-т физ.-мат. наук, доц. (Национальный тех— нический университет «КПІ»);

В. В. Гавриленко — д-р физ.-мат. наук, проф.,

О. П. Будя — к-т техн. наук, доц. (Киевский университет эко— номики, туризма и права);

Н. К. Печурин — д-р техн. наук, проф. (Национальный ави— ационный университет).

Кононюк А. Е.
К213 Дискретно-непрерывная математика. (Множества (четкие)). — В 12-и кн. Кн 2, ч.1— К.: Освіта України. 2012. — 522 с.

ISBN 978-966-373-693-8 (многотомное издание)
ISBN 978-966-373-694-5 (книга 2)

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 2, ч.1)
© Кононюк А. Е., 2012
© Освіта України, 2012

Оглавление

Введение

Излагаемый ниже материал состоит из тех областей современной математики, которые имеют отношение к вычислениям, и, как следствие, обеспечивает пользователя средством для краткого и точного описания многих проблем разных направлений науки. Изложение материала в данном модуле и во всех следующих носит по возможности конструктивный характер. Везде, где возможно, в каждой новой теме используются понятия и термины из предыдущих тем; материал сопровождается многочисленными упражнениями и примерами. Следует подчеркнуть, что разбор примеров и решение упражнений является составной частью изучения предлагаемого материала.

Преследуя эту цель, мы должны з чего-то начать. Нашим исходным неопределяемым понятием является понятие **множества**, описываемое перечислениям свойств, которыми оно обладает. Исходя из этого, можно определить все следующие понятия конструктивным и математически приемлемым оьразом. Такой подход необходим, поскольку любую ошибку легко можно проследить, возвратившись назад к неправильному предложению где-то в цепочке соображений. Это также означает, что часть или же вся рассматриваемая теория может быть запрограммирована.

Множества обсуждаются в модуле 1 и используются дальше во всей теории. Терминология и обозначения обычно вводятся в соответствующем месте текста. Нам хотелось бы построить предлагаемую теорию довольно (математически) строго, однако при этом возникают обозначенные трудности с ее обоснованием. Вместо этого мы начнем из описания таких начальных понятий, как множество.

1. Основные понятия и операции над множествами

1.1. Основные понятия теории множеств

Основные определения

Понятие *множества* есть фундаментальным неопределяемым понятием. Интуитивно под множеством будем понимать ***совокупность определенных целиком предметных объектов, которые рассматриваются как единое целое.***

Можно говорить о множестве стульев в комнате, множества людей, которые живут в г. Киеве, множества студентов в группе, множества натуральных чисел, множества букв в алфавите, множества состояний системы и т.п. При этом о множестве можно вести речь только тогда, когда элементы множества различимы между собой. Например, нельзя говорить о множестве капель в стакане воды, так как невозможно четко и ясно указать каждую отдельную каплю.

Кстати, поскольку "множество" (set) в русском языке как бы намекает на "много", а понятие "много" (many) у каждого из нас свое, то, для предотвращение спора между русскоязычными людьми, мы будем слово "множество" использовать для любого количества элементов, как и англоязычный Запад. Даже для одного элемента. Даже в случаях, когда в множестве нет ни одного элемента - такое множество **называется пустым!** Это, в частности, разрешит рассказывать своим друзьям корректный, с точки зрения теории множеств, анекдот о "множестве ветеранов Полтавской битвы, которые сейчас живут"...

Кроме понятия множества есть еще лишь одно исходное базовое понятие - и все. Другие понятия в этой теории производные. ***Второе базовое понятие*** - это **ПРИНАДЛЕЖНОСТЬ (или "отношение принадлежности").** Т.е. "элемент принадлежит множеству". Здесь, тем более, нечего определять, имея ввиду, что слово "принадлежит" в повседневном языке можно заменять, с учетом контекста, многими синонимами, а именно:

- та березка "находится" в этом лесу,
- Петренко "числится" в студентах.

Будем предполагать, что мы всегда четко знаем, что принадлежит данному множеству, а что нет! Другое считаем несуществующим

вообще**!** Такое представление о множестве разрешает охарактеризовать его как четкое.

Но в 1972 г. американский ученый Л.А.Заде ввел понятие нечеткого (размытого) множества. Как отметил Л.А.Заде, **теория нечетких множеств - это, в сущности, шаг на пути к сближению точности классической математики и всепроникающей неточности реального мира, к сближению, порожденном человеческим стремлением, которое не прекращается, к лучшему пониманию процессов мышления и познания**.

Окружающую нас реальность мы познаем только с помощью созданных нами моделей, представлений и более или менее достоверных законов и приближений, приемлемых при данном состоянии наших знаний, однако, та самая модель того самого явления разными людьми воспринимается по-разному: формула может оставаться неизменной, но интерпретации ее могут быть разными. Мир воспринимается с помощью моделей, которые совершенствуют друг друга и взаимосвязанны между собой до тех пор, пока не происходит революция в идеях, которая прекращает такое объединение в моделях.

Наши модели - нечеткие; наши мысли, которые сформированы на основе более или менее независимых моделей, - нечеткие; **так мы отличаемся от компьютера**. **Человек,** кроме способности размышлять и логически мыслить, **владеет способностью принять во внимание параллельно понимания как общего, так и сопутствующего характера.** Эти общие и привходящие соображения в противоположность логическому соображению являются нечеткими и должны быть нечеткими. Живое существо, которое наделено инициативой, воспринимает и обрабатывает более или менее нечеткую информацию и своевременно приспосабливается к ней. Когда живое существо почти не имеет возможности проявить инициативу, когда ее ***функциональная энтропия почти равна нулю***, тогда нечеткость может исчезнуть и ее действия становятся определенными. **Биологическая клетка функционирует как маленький компьютер, управляющий маленькой фабрикой (слово «маленький» относится здесь к размеру, а не к сложности);** ***в этой системе почти нет энтропии. Человек, наоборот, владеет огромной функциональной энтропией***; он может **выбирать, принимать решение, эволюционировать, допускать и устранять ошибки, начинать сначала, понимать не все, формулировать свои знания в процессе научных исследований по формальной программе.**

Как концептуально объединить общие соображения с логическим соображением; как связать то, что является физической истиной, с тем, что представляет собой интерпретацию человеческой мысли? Как ввести нечеткость в математику, посколько в конечном счете именно в такой наиболее ясной математической форме следует выразить эту, на первый взгляд, странную взаимосвязь?

Что означает слово «**нечеткий**» или синонимические ему слова **для математика? Это значит, что какой-нибудь элемент принадлежит подмножеству, но только немного неопределенным образом**. Но, с другой стороны, мы знаем, что в математике есть только две приемлемые ситуации для элемента: он может или быть, или не быть элементом подмножества. Любая формальная логика, в том числе булева, основана на этом: элемент принадлежит или не принадлежит подмножеству данного множества.

Заслуга Л. А. Загде заключается в попытке выйти из этого тупика путем введения **понятия взвешенной принадлежности. Элемент может принадлежать подмножеству в большей или меньшей степени, и отсюда появляется основное понятие - понятие *нечеткого подмножества.***

Мы будем последовательно изучать теорию четких и нечетких множеств. Начнем из традиционных - четких множеств.

Отдельные объекты, из которых составляется множество, называются *элементами* множества. Так, число 3 — элемент множества натуральных чисел, а буква *б* — элемент множества букв русского алфавита.

Общим обозначением множества служат пары фигурных скобок { }, внутри которых пересчитываются элементы множества. Для обозначения конкретных множеств используются разные прописные буквы *A, S, X...* или прописные буквы с индексами A_1, A_2... Для обозначения элементов множества в общем виде используются разные строчные буквы *a,s, x...* или строчные буквы с индексами a_1, a_2...

Если множество содержит несколько элементов, то мы просто записываем все ее элементы. Например, если мы определим *A* как множество всех целых чисел строго между 6 и 10, то это можно записать таким образом:

$$A = \{7, 8, 9\}$$

и прочитать как

«*A* — множество, которое содержит 7, 8, 9».

Здесь символ «=» используется в определенном содержании: *A* равняется множеству... Дальше будет использоваться высказывание «или равное *A*.....»... Поэтому предложим процедуру установления

справедливости этого утверждения. Другими словами, **множество можно охарактеризовать определенными свойствами**, итак, множество A можно определить как

$$A = \{x: x \text{ — целое число и } 6 < x < 10\}$$

и прочитать как

«A есть множество всех x таких, что ...».

Множеству A принадлежат только те элементы, которые являются целыми числами, большими 6 и меньшими 10, т.е. 7, 8 и 9, итак, мы имеем 7, 8, 9, как и раньше.

Множества часто рассматривают как «неупорядоченные совокупности элементов», хотя иногда полезно подчеркнуть, что, например,

$$\{7, 8, 9\} = \{8, 9, 7\} = \{9, 8, 7\} = ...;$$

мы не делаем никакого предостережения о порядке, в котором рассматриваются элементы, поэтому было бы неправильным допускать какой-нибудь определенный порядок.

Если мы хотим сказать, что все березки, которые находятся в данном лесу, принадлежат и всему лесному богатству нашей страны, а все студенты, которые числятся в университете, числятся и студентами Украины, то для сокращения фраз используются строки

ПОДМНОЖЕСТВО или **ВКЛЮЧЕННАЯ.**

Для любого заданного объекта можно определить, принадлежит ли он множеству A. В частности, если число принадлежит множеству, то будем говорить, что «оно является элементом множества». Так, например, если 7 является элементом множества A, то это утверждение может быть записано таким образом:

«$7 \in A$».

Утверждение «6 не является элементом A» будем обозначать как

«$6 \notin A$».

До сих пор нам встречались символы $\{:\}$, $\{,,,\}$, $\in$ и $\notin$.

Их применение кажется довольно простым, однако оно требует достаточных навыков, которые будут проиллюстрированы на следующих примерах.

При использовании терминов **ПРИНАДЛЕЖИТ** или **ВКЛЮЧЕННЫЙ** могут быть очевидные синонимы. Но чтобы в них не запутаться и просто не перепутать с "принадлежит", нужно помнить одну простую вещь: "принадлежит" относится к случаю, когда "ЭЛЕМЕНТ принадлежит МНОЖЕСТВУ", а "включенный" - когда "МНОЖЕСТВО включено в МНОЖЕСТВО".

Например, множество студентов университета „включено" в множество студентов страны. Т.е. множество студентов университета „есть подмножеством" множества студентов страны.

У отношения включения есть ряд интересных свойств. Они могут быть выявлены любым исследователем, если он "поиграет" с этим отношением.

Например, можно сказать, что множество студентов группы К-001 включено в множество студентов университета, поскольку такая группа в университете числится. То, что из группы отчислены все студенты, для математики никакой роли не играет. Поскольку, НЕТ ни одного студента, который числится в этой группе, который бы не числился в университете. Такого рода соображения совсем корректно можно применить к любому пустому множеству и сделать обобщающий вывод, что пустое множество **есть включенным в любое множество, в том числе и в себя.**

Элемент, который принадлежит множеству, которое не содержит ни одного элемента, принадлежит и любому другому множеству, которое не содержит ни одного элемента.

Любое множество является собственным подмножеством.

Теория множеств оперирует со всеми множествми, кроме тех, которых нельзя создать. **Все эти множества, которые объединены в одно множество, называются <u>УНИВЕРСУМОМ</u>.**

Пример 1.1. Определить, какие из приведенных определений множеств являются правильными:

$A=\{1, 2, 3\}$,

$B=\{5, 6, 6, 7\}$,

$C=\{x: x \notin A\}$,

$D=\{A, C\}$,

$E=\{x: x=1$ или $x=\{y\}$ и $y \in E\}$,

$F=$ {множества, которые не является элементами самих себя}=

$=\{ x: x$ - множество и $x \notin x\}$?

Если число элементов множества A легко перечисляется и среди элементов множества нет повторений, то определение верно.

Множество B выглядит также правильно, за исключением лишь того, что число 6 встречается дважды. Мы можем проверить, принадлежит ли элемент множеству или нет. Таким образом, это наиболее важное требование в определении множества выполненно. Итак, мы можем рассматривать эту запись как верную и эквивалентную {5, 6, 7}. Однако в этой ситуации возникают следующие проблемы. Если мы рассмотрим первоначальное определение B и выбросим одно из чисел 6 из множества, то мы,

наверное, будем иметь $6 \in B$ и $6 \notin B$. Возникает противоречие. Поэтому мы будем рассматривать повторение символов в определении множеств как восприятие того самого символа, а его дублирование как недосмотр; удаление повторяемых символов образует основу для некоторых дальнейших математических соображений.

Множество A содержит числа; это может вызвать удивление, так как числа не существуют. Более точно, мы используем символы чисел; эти символы называются так же, как и числа. Поэтому B — множество имен, и мы обычно используем имена, чтобы представить объекты (элементы), на которые ссылаемся. В исчислениях имена имеют особое значение, особенно в изучении семантики программных языков (содержания программ). Мы здесь не будем входить в детальное обсуждение этих проблем; довольно указать ловушки и необходимость адекватных спецификаций рассмотренных объектов.

Возьмем, например, множество

X = {«Введение в теорию систем», «Основы структурных данных», «Введение в теорию систем»}.

Это - множество названий двух книг с одним элементом, который по невнимательности записан дважды, или же это - множество трех книг, две из которых имеют одно и то же название? Если правильно последнее, то две книги «Введение в теорию систем» следует разделить каким-нибудь способом. Из данной информации нельзя выяснить правильный ответ, поэтому в данном случае следует быть осторожным.

Определение C также справедливо как A, так как, если $x \in A$ то $x \notin C$ и, если $x \notin A$, то $x \in C$. Множество C очень большое: оно содержит «все», за исключением чисел 1, 2 и 3. Обозначение «все» выделено и, как мы вскоре увидим, «опасное» с математической точки зрения.

Так как определение D, A и C представляют множества, то отсюда получаем, что определение D также справедливо. Заметим, что это — множество множеств (ничего неверного в этом нет!) такое, что оно имеет только два элемента, в частности $1 \notin D$, даже если $1 \in 4$ и $A \in D$. Это легко проверить, так как $1 \neq A$, $1 \neq C$ и только A и C является элементами D.

Множество E является первым примером рекурсивно обусловленного множества; оно определяется (частично) в терминах самого себя. Конструктивный процесс продлевается бесконечно, поэтому мы должны иметь правило для определения элементов. Мы не можем записать их явно. Заметим, что E не определяется целиком в

терминах E, Мы должны знать о множестве что-то такое, что не зависит от определения; в данном случае это то, что $1 \in E$. Имеем:

$$1 \in E, \quad \text{поэтому} \quad \{1\} \in E,$$
$$\{1\} \in E, \quad \text{поэтому} \quad \{\{1\}\} \in E,$$
$$\{\{1\}\} \in E, \quad \text{поэтому} \quad \{\{\{1\}\}\} \in E \text{ и т.д.}$$

Хотя конструктивный процесс неограничен, беря любой элемент и располагая достаточным времени, можно определить, содержится ли этот элемент в E.

Перейдем теперь к F; это довольно трудная задача. Чтобы увидеть, почему F не может существовать, мы сначала допустим существование, а потом продемонстрируем, что существует особый элемент (обозначим его через y) такой, что мы не можем определить, $y \in F$ или $y \notin F$. Вообще исследование «неудобного» примера, на котором мы можем показать логический изъян, проводить нелегко; однако в данном случае мы можем использовать само множество F. Чтобы прояснить дело, обозначим это множество через G. Если, как мы предполагаем, F — искомое множество, то или $G \in F$, или $G \notin F$. Рассмотрим два возможных случая:

а) $G \in F$. Тогда G удовлетворяет условию содержания, т.е. $G \notin G$, и, следовательно, $G \notin F$;

б) $G \notin F$ говорит о том, что G не удовлетворяет условию вхождения в F, и, следовательно, $G \in F$.

Таким образом, во всех случаях мы приходим к противоречию. Поэтому F не может существовать. Рассмотрим где была сделана ошибка. Множества множеств, вероятно, разрешаются, и бесконечно большие множества (например, рассмотреннок выше множество E) также разрешаются; однако с «множеством всех множеств» нельзя работать в обычной теории множеств — это требует другого рода математики. **Эта аномалия теории множеств известна как парадокс Рассела.** Если мы уже имеем множество H, то можно определить J:

$$J = \{x: x \in H \text{ и } x \in x\}.$$

Таким образом, мы будем использовать только множества, которые могут быть явно записаны или же построены путем хорошо определенных процессов. Поэтому множества не так тривиальны, как они могли вначале показаться. Однако, следуя приведенным выше правила, работа с ними не будет особо трудной.

Самое интересное в теории множеств то, что она рассматривает не только конечные множества - множества, которые содержат конечное число элементов, но и бесконечные, для которых даже

понятие числа не имеет значения. Т.е., теория множеств может рассматривать не только множество студентов в группе и множество березок в лесу, но и множество точек на прямой, и множество звезд на небе...

Таким образом, множества бывают **конечными и бесконечными.**

Множество называется *конечным*, если число его элементов конечно, т.е. если существует натуральное число N, которое является числом элементов множества. Множество называется *бесконечным*, если оно содержит бесконечное число элементов.

Основоположником теории множеств является Георг Кантор, который именно через бесконечности попортил себе много крови. Хотя с бесконечностью математики до него уже давно работали. Взять то же бесконечно большое множество точек на прямой или наоборот, бесконечно малые величины из высшей математики...

Но вся беда в том, что ни один живой человек не видел, не чувствовал, не щупал бесконечности! Поэтому до Кантора математики признавали и использовали так называемую ПОТЕНЦИАЛЬНУЮ бесконечность. Самый существенный пример - это понятие бесконечно большого числа в высшей математике. Бесконечно большое число это число, которое больше каждого наперед заданного. Какое бы вы не назвали число, ваш оппонент назовет, как минимум, на единицу больше, т.е. у нас с вами всегда в запасе число потенциально (!) большее, чем придумает оппонент.

Кантор же предложил ввести в математику АКТУАЛЬНУЮ бесконечность. **По Кантору бесконечность существует сразу вся. А раз бесконечные множества существуют, и сразу в целом, то с ними можно делать математические манипуляции. Их даже можно сравнивать на большее-меньшее.**

Поэтому Кантор начал задавать себе разные „каверзные" вопросы и искать на них математические ответы. Один из ключевых вопросов:

"БЕСКОНЕЧНО МНОГО - это всегда ОДИНАКОВО БЕСКОНЕЧНО МНОГО? Или, могут ли быть большие и меньшие бесконечности?" Чего больше, звезд на небе или точек на прямой?..

Кантор доказал теорему, из которой следует, что бесконечности могут быть разные по величине. Поскольку "число" и "количество" - слова в этом случае неуместны, то он ввел термин "мощность". ***Мощность*** **- это то, что остается, когда нас не интересует сущность элементов множества и порядок, в котором**

они располагаются. Т.е., он <u>определил понятие мощности строго</u>.

От множества студентов останется только мощность, если мы перестанем их различать и будем воспринимать их вне всякого порядка (в естественных условиях)**.**

К сожалению, приводить примеры множеств, которые имеют бесконечную мощность, используя березки и студентов, не получится вообще. Поэтому обратимся для наглядности к числам.

Пересчитывая что-то мы используем целые (положительные) числа 1, 2, 3 ... Их еще называют "натуральными". Американцы начинают этот ряд с нуля . При добавлении или вычитании нуля ничего не меняется. Главное, мы знаем, что чисел нам хватит для перечисления чего угодно. Мы также знаем, что это множество бесконечное. Кантор назвал это множество СЧЕТНЫМ и его **мощность - <u>мощностью счетного множества</u>**. Мощность этого множества Кантор взял за эталон и предложил сравнивать его с мощностями других множеств. Во-первых, он установил, что эта мощность больше мощности любого конечного множества (студентов, березок и т.п.). Во-вторых, он доказал, что много бесконечных множеств имеют ту же мощность (то же "количество" элементов), что и счетное. Один из примеров - множество целых положительных чисел имеет столько же элементов, сколько и множество целых парных положительных чисел. Т.е. они равномощны.

Действительно, запишем одно под одним числа:

1 2 3 4 ...

2 4 6 8 ...

Ясно, что обе последовательности имеют одинаковое количество элементов, поскольку любому числу первой, ВСЕГДА отвечает строго одно число второй последовательности, так что вторая последовательность не может исчерпаться раньше первой. И наоборот. Следовательно, эти множества равномощны. Следовательно, здесь ЧАСТЬ РАВНЯЕТСЯ ЦЕЛОМУ!

<u>Поскольку это доказано строго, то этому приходится верить</u>.

За свою длинную жизнь человек может столкнуться с чем угодно, <u>а с бесконечностью - никогда</u>! Поэтому, что может быть и чего не может быть в мире бесконечностей судить тяжело, основываясь лишь на жизненном опыте!

Из бесконечного множества звезд (мощность которых тоже счетная) мы видим лишь их ограниченное конечное множество. На нарисованном отрезке прямой, которая содержит бесконечное множество точек, мы видим конечное множество зерен грифеля,

которым отрезок нарисован. **Так что бесконечности вокруг нас существуют в "параллельном мире" по своим законам, которые теория множеств помогает изучать.**

Мы уже сказали "во-вторых", но есть еще и "в-третьих" - и это "в-третьих" - самое главное: **теорема Кантора, которая уже упоминалась.**

Дело в том, что если построить множество всех подмножеств конкретного множества, то всегда получите множество БОЛЬШЕЕ исходного.

Например, возьмем множество из 2-х элементов: один, два. Подмножествами этого множества будут 4 множества:

1) один, два - (любое множество подмножества самого себя)
2) один,
3) два,
4) пустое .

Другой пример: А І Б Подмножествами этого множества с трех элементов будет 8 множеств :

1) А, І, Б
2) А, І
3) А, Б
4) І, Б
5) А
6) І
7) Б
8) пустое

Из четырех элементов вышло бы 16 элементов. И этот ряд можно бесконечно продолжить, как ряд степеней числа 2.

Так вот, Кантор доказал, что если взять бесконечное множество счетной мощности, например, множество целых положительных чисел и построить (понятно, умозрительно) множество, которое содержит как элементы всего подмножества этого множества, то получим мощность БОЛЬШУЮ, чем счетная мощность. В принципе не существует способа пересчитать (пусть даже в бесконечности) такое множество. В нем всегда больше элементов.

Эта новая большая мощность называется мощностью КОНТИНУМА.

И снова парадокс. Мощность континума имеет, например, множество точек прямой или множество действительных чисел, которые то же самое. Больше того, любой отрезок числовой оси, даже такой малюсенький отрезок, как отрезок от 0 до 1, имеет мощность континума, т.е. **на нем больше чисел, чем найдется чисел в счетном**

множестве. А раз этот отрезок имеет мощность континума, как и вся (бесконечная) прямая и, естественно, любой ее отрезок, то можно сказать, что на отрезке от 0 до 1 ровно столько же точек, сколько на отрезке прямой от Земли до Юпитера.

Здесь тоже часть равняется целому, если и часть, и целое имеют мощность континума. И все они одинаково больше числа звезд на небе или числа всяческих алгоритмов...

Для бесконечностей существует очень простая арифметика, которая логически вытекает из предыдущих рассуждений. Суммирование двух счетных мощностей дает счетную мощность, а для континумов - мощность континума. При вычитании из мощности континума счетного - в остатке мощность континума. Но вот если вычитать из континума континум или из счетной мощности счетную - всякое может выйти в каждом конкретном случае.

Однако, не все так просто. **Бесконечность остается одной из ключевых категорий философии.** И здесь математика показывает все новые и новые грани этой проблемы. Тем более, если говорить не только о бесконечных, но и о бесконечных упорядоченных множествах.

Для того чтобы оперировать с конкретными множествами, как мы уже об этом говорили, нужно уметь задавать эти множества.

Способы задания множеств. Множество может быть задано перечислениям (списком своих элементов), порождающей процедурой или описанием характеристических свойств, которыми должны обладать его элементы.

Списком можно задавать лишь конечные множества. Задача типа N =1, 2, 3 ... — это не список, а условная пометка, допустимая лишь тогда, когда она сознательно не вызывает разночтений. Как мы уже отмечали, **список** обычно **заключают в фигурные скобки**. Например, $A=\{a, b, d, h\}$ означает, что множество A состоит из четырех элементов a, b, d и h.

Порождающая процедура описывает способ получения элементов множества из уже полученных элементов или из других объектов. Элементами множества считаются все объекты, которые могут быть построены с помощью такой процедуры. Примером служит описание множества M_4, где исходными объектами для построения являются натуральные числа, а породжуючею процедурой - высчисление, описанное формулой $\pi/2\pm k\pi$. Другой пример — множество M_{2^n} =1, 2, 4, 8, 16 ..., порождающая процедура для которого определяется следующими двумя правилами:

1) $1 \in M_{2^n}$;

2) если $m \in M_{2^n}$, то $2m \in M_{2^n}$.

(Правила, которые описаны таким образом, называются *индуктивными* или *рекурсивными;* о них речь будет идти дальше.) .

Третий пример — множество M_π, заданное следующим образом. Пусть имеется процедура вычисления цифр разложения числа π в бесконечную десятичную дробь: π=3,1415926536... По мере вычисления будем образовывать из последовательно стоящих цифр трехразрядные числа: 314, 159, 265 и т.д. Множество всех таких чисел обозначим M_π.

Весьма распространенной порождающей процедурой является образование множеств из других множеств с помощью операций над множествами, которые будут рассмотрены ниже.

Задание множества описанием свойств его элементов, пожалуй, наиболее обычно. В примере, который рассмотрен выше, так заданы множества M_2, M_3, M_5; да и задание множества M_4 можно интерпретировать, как описание свойств его элементов, заключающегося в возможности представить их в виде $\pi/2 \pm k\pi$. Множество M_{2^n} можно задать фразой «M_{2^n} — множество всех целых чисел, являющихся степенями двойки». В случае, когда свойство элементов M может быть описано коротким выражением $P(x)$ (означающим «x обладает свойство P»), M задается при помощи обозначения $M=\{x|P(x)\}$, которое читается так: «M — это множество x, обладающих свойство P». Например,

$$M_{2^n} = \{ x \mid x = 2^k \}, \quad \text{где } k \in N\},$$

$$M_4 = \{ x \mid x = \pi/2 \pm k\pi\}, \quad \text{где } k \in N\}.$$

К описанию свойств естественно предъявить требование точности и недвусмысленности. Например, множество всех хороших фильмов 2010 года разные люди зададут разными списками (быть может, пустыми); сами критерии, по которым производится отбор, при этом будут разные. Такое множество нельзя считать строго заданным. **Надежным способом точно описать свойство элементов данного множества служит задание распознающей (или, как говорят в математике, разрешающей) процедуры, которая для любого объекта устанавливает, обладает он данным свойством и, следовательно, является элементом данного множества или нет.** Например, для M_{2^n}, т.е. для свойства «быть степенью двойки»,

разрешающей процедурой может служить любой метод разложения целых чисел на простые множители.

Отметим, что в этом примере разрешающая процедура не является порождающей. Однако ее нетрудно сделать таковой: берем последовательно все натуральные числа и каждое из них раскладываем на простые множители; те числа, которые не содержат множителей, отличных от 2, включаем у M_{2^n}. С другой стороны, порождающая процедура может не быть разрешающей.

Важным понятием теории множеств есть понятие пустого множества, о котором мы уже упоминали раньше.

Пустым множеством называется множество, которое не содержит ни одного элемента. Пустое множество обозначается $\varnothing$. Например,

$$\{ x \in C \mid x^2-x+1=0\} = \varnothing.$$

Понятие пустого множества играет очень важную роль при задании множества с помощью описания. Так, без понятия пустого множества мы не могли бы говорить о множестве отличников группы или о множестве действительных корней квадратного уравнения, не убедившись предварительно, есть ли вообще в данной группе отличники или имеет ли данное уравнение действительные корни. Введение пустого множества позволяет соввершенно спокойно оперировать с множеством отличников группы, не убедившись предварительно, есть ли вообще в рассматриваемой группе отличники. **Пустое множество будем условно относить к конечным множествам.**

Рассмотрим теперь вопрос о равенстве множеств. Два множества называются *равными,* если они состоят из одних и тех же элементов, т.е. представляют собой одно и то же множество. Множества X и Y не равны $(X \neq Y)$, если либо в множестве X есть элементы, не принадлежащие Y, либо в множестве Y есть элементы, не принадлежащие X. Символ равенства множеств обладает свойствами:

$X=X$ - рефлексивность;

если $X=Y$, то $Y=X$ — симметричность;

если $X=Y$ и $Y=Z$, то $X=Z$ — транзитивность.

Из определения равенства множеств вытекает, что порядок элементов в множестве несуществен. Так, например, множества {3, 4, 5, 6} и {4, 5, 6, 3} представляют собой одно и то же множество.

Из определения множества следует, что в нем не должно быть неразличимых элементов. Поэтому в множестве не может быть одинаковых элементов. Запись {2, 2, 3, 5)}, как мы уже знаем,

следует рассматривать как некорректную и заменить ее на $\{2, 3, 5\}$. Так, множество простых делителей числа 60 равно $\{2, 3, 5\}$.

Рассмотрение способов задания множеств приводит к мысли о том, что самое понятие «точно задать множество» нуждается в уточнении. Такое уточнение совсем не просто, а его важность крайне велика и выходит далеко за пределы самой теории множеств.

Язык множеств — это универсальный язык математики. **Любое математическое утверждение можно сформулировать как утверждение о некотором соотношении между множествами**: о равенстве двух множеств, о непустоте некоторого множества («существует непрерывная нигде не дифференцируемая функция»), о непринадлежности элемента множества («с помощью циркуля и линейки нельзя построить круг, равновеликий данному квадрату»), и т.д. Поэтому анализ способов задания множеств связанн с анализом строгости математических утверждений вообще, т.е. с обсуждением самих оснований математики.

Понятие подмножества

Множество X является подмножеством множества Y, если любой элемент множества X принадлежит и множеству Y. Пусть Y — множество студентов группы, а X — множество отличников этой же группы. Так как каждый отличник группы есть в тот же время студентом этой группы, то множество X является подмножеством множества Y.

Многие определения теории множеств удобно давать в виде математических выражений, которые содержат некоторые логические символы. Для определения подмножества используем два таких символа:

$\forall$ — символ, который называют квантором и означающий «любой», «каков бы ни был», «для всех»;

$\rightarrow$ - символ следствия (импликации), что означает «влечет засобой».

Определение подмножества, которое может быть сформулировано в виде: для любого x утверждение «x принадлежит X» влечет за собой утверждение «x принадлежит Y», запишется так:

$$\forall x[x \in X \rightarrow x \in Y].$$

Более краткой записью выражения «X является подмножеством Y» будет запись

$$X \subseteq Y,$$

что читается как «Y содержит X». Используемый здесь символ $\subseteq$ означает включение. Если желают подчеркнуть, что Y содержит и другие элементы, кроме элементов из X, то используют символ строгого включения $\subset$:

$$X \subset Y.$$

Связь между символами $\subset$ и $\subseteq$ дается выражением

$$X \subset Y \Leftrightarrow X \subseteq Y \quad X \neq Y.$$

Здесь использован символ $\Leftrightarrow$, который означает эквивалентность (в содержании «то же самое, что»).

Отметим некоторые свойства подмножества, которые вытекают из его определение:

$$X \subseteq Y \quad \text{(рефлективность);}$$

$$[X \subseteq Y \text{ и } Y \subseteq Z] \rightarrow X \subseteq Z \quad \text{(транзитивность).}$$

Несколько труднее видеть, что для любого множества M

$$\varnothing \subseteq M.$$

Действительно, пустое множество $\varnothing$ не содержит элементов. Следовательно, добавляя к M пустое множество, мы фактически ничего не добавляем. Поэтому всегда можно считать, что любое множество M содержит в себе пустое множество в качестве подмножества.

Верхняя и нижняя границы множества

Имея дело с множеством действительных чисел, можно сравнивать элементы этого множества по их значению и, в частности, находить наибольший и наименьший элемент множества. Для конечных множеств, заданных перечислением, эта задача не представляет трудностей. Так, для множества $T=\{4, 3, 5, 6\}$ имеем $\max T=6$, $\min T=3$. Однако если множество задано описательным способом, например, указано лишь правило вычисления числовых значений его элементов, то задача определения наибольшего и наименьшего элементов становится довольно трудной.

Несколько более легкой задачей является нахождение лишь области, внутри которой лежат все элементы множества.

При решении этой задачи очень полезными являются **понятия верхней и нижней границ множества**.

Пусть S — множество действительных чисел. Верхней границей S является число C, такое, что для любого $x \in S$ имеет место $x \leq C$. Чисел, которые могут рассматриваться в качестве верхней границы

множества, может быть бесконечно много, а может и не быть вообще. Так, в множестве $m<S<M$ любое $C\geq M$ является верхней границей. **Множество всех целых чисел не имеет верхней границы.**

Точной верхней границей или *супремумом* множества S, обозначаемой $\sup S$, называется верхняя граница, которая не превосходит любую другую верхнюю границу. В приведенном выше примере

$$\sup S=M.$$

Множество может иметь только одну точную верхнюю границу, так как если C_1 и C_2 — две такие границы, то $C_1\leq C_2$ и $C_2\leq C_1$ и, следовательно, $C_1=C_2$.

Нижней границей множества S является число c, такое, что для любого $x\in S$ имеет место $x\geq c$. Точной нижней границей или *инфинумом* множества S, обозначаемой $\inf S$, называется нижняя граница, не меньшая любой другой нижней границы. В приводимом примере

$$\inf S=m.$$

Теорема (теорема о верхней и нижней границах подмножества). Если $B\subseteq A$, то

$$\inf B\geq\inf A\ ;\ \sup B\leq\sup A.$$

Доказательство. Обозначим через b' элемент множества B, имеющий наименьшее значение, т.е. $b'\in B$ и $b'=\inf B$. Но $B\subseteq A$, т.е. $b'\in A$. Пусть a' - элемент множества A, имеющий наименьшее значение, т.е. $a'\in A$ и $a'=\inf A$. При этом если $b'=a'$, то $b'=\inf A$, если $b'\neq a'$, то $b'>a'=\inf A$. Таким образом, $b'\geq\inf A$ или $\inf B\geq\inf A$.

Вторая часть теоремы доказывается аналогично.

1.2. Операции над множествами

Предварительные замечания

Говорят операции НАД множествами не потому, что они расположены "над" множествами, а просто так принято.

Основных операций всего три. Это меньше, чем в школьной арифметике. Хотя даже это множество операций немного избыточное.

Операции называются: **ОБЪЕДИНЕНИЕ, ПЕРЕСЕЧЕНИЕ** и **ДОПОЛНЕНИЕ.** Они чем-то напоминают операции элементарной алгебры: добавление, умножение и изменение знака. Но эта аналогия приблизительна и опасна.

Если рассмотреть внимательно студенческую группу В-004, то объединение множества отличников и спортсменов даст множество под названием "слава группы В-004". Принципиальное отличие объединения множеств от школьного добавления не только в том, что студенты - это не числа и мы их не пересчитываем (!), но и в том, что студенты, которые одновременно отличники и спортсмены, будут учтены один раз. Так что може оказаться, что отличников четыре, а спортсменов двадцать, но их объединение по названию "слава группы В-004" будет содержать всего двадцать два студенты. Ясно, что пересечение этих множеств даст двух студентов, которые одновременно и отличники и спортсмены.

Когда у нас появляются в руках объекты, а мы можем брать любые объекты, и операции - а мы основную тройку этих операций тоже определили, то будем говорить об АЛГЕБРЕ.

Как мы уже говорили, над множествами можно делать действия, которые во многом напоминают действия добавления и умножение в элементарной алгебре. Чтобы лучше разобраться в действиях над множествами, необходимо вспомнить законы, которые существуют в элементарной алгебре.

Алгебра множеств очень отличается от элементарной, хотя есть некоторые аналогии. В алгебре множеств есть те же названия законов: КОММУТАТИВНЫЙ, АССОЦИАТИВНЫЙ и ДИСТРИБУТИВНЫЙ (переместительний, соединительный и распределительный). Первые два сходны с элементарной алгеброй. А вот дистрибутивный закон имеет и аналог в элементарной алгебре (выражаясь "по-школьному": произведение суммы есть сумма произведений), но имеет и отличительную версию. В теории множеств пересечение с объединением равно объединению пересечений и (!) объединение с пересечением равно пересечению объединений. Второе не имеет аналога в элементарной алгебре: "Сумма с произведением не равна произведению сумм".

Проиллюстрируем сказанное сначала словесно:

Коммутативный закон: объединение (пересечение) отличников и спортсменов равно объединению (пересечению) спортсменов и отличников.

Ассоциативный закон: от изменения порядка объединения (пересечение) спортсменов, отличников и девушек результат не меняется.

Дистрибутивный закон (только оригинальная версия): объединение девушек с пересечением спортсменов и отличников равно множеству, в котором пересекаются объединение девушек и спортсменов с объединением девушек с отличниками. (В условных обозначениях это было бы намного короче и нагляднее, но об этом мы будем говорить ниже).

Несколько тяжелее воспринимается на слух закон поглощения, который, однако, в ряде случаев разрешает упрощать теоретико-множественные конструкции. Пересечение отличников с объединением отличников и спортсменов дает множество отличников.

Или второй вариант.

Объединение отличников с пересечением отличников и спортсменов дает множество отличников.

Если тщательно обдумать сказанное, то справедливость результатов очевидна.

Есть еще закон, который следует отнести к важнейшим законам (свойствам). Это закон **ИДЕНПОТЕНТНОСТИ**, который звучит так (на примере спортсменов): *объединение (пересечение) множества спортсменов с множеством спортсменов дает множество спортсменов.*

Приведем еще один очень важный закон - **ЗАКОН Де Моргана**: дополнение объединения отличников со спортсменами равно пересечению дополнения множества спортсменов с дополнением множества отличников.

И второй вариант.

Дополнение пересечения отличников со спортсменами равно объединению дополнения множества спортсменов с дополнением множества отличников. В качестве универсума (для дополнения) можно взять множество студентов группы (или университета, или мира - роли не играет).

Очень простой закон **ДВОЙНОГО ДОПОЛНЕНИЯ.** Дополнение дополнения множества спортсменов есть само множество спортсменов.

Следствие из этого закона.

Дополнение дополнения дополнения множества спортсменов есть дополнения множества спортсменов.

Важными есть еще такие два закона:

ПРОТИВОРЕЧИЕ и ИСКЛЮЧЕННОГО ТРЕТЬЕГО.

Противоречие: Пересечение множества спортсменов с дополнением множества спортсменов несовместимы. Так как в дополнение множества спортсменов входят все другие студенты

неспортсмены, то у этого пересечения не может быть общих элементов.

Исключенного третьего: Объединение множества спортсменов с дополнением множества спортсменов совпадает с рассмотренным универсумом. Действительно, так как в дополнение множества спортсменов входят все другие студенты неспортсмены с универсума, то это объединение именно и составляет весь универсум.

Проиллюстрируем сказанное выше в математической форме

Пусть a и b — некоторые числа, $a+b$ — их сумма и ab — их произведение. Сумма и произведение чисел имеют следующие свойства, которые называются законами алгебры.

1. $a+b=b+a;\ ab=ba$ — коммутативный или переместительний закон;
2. $(a+b)+c=a+(b+c);\ (ab)c=a(bc)$ - ассоциативный или соединительный закон;
3. $(a+b)c=ac+bc$ — дистрибутивный или распределительный закон.

Заметим, что в ассоциативном и коммутативном законах можно заменить действие сложения умножением, а действие умножения сложением. При этом получим другой закон, который будет так же справедлив, как и первый. Однако в дистрибутивном законе подобной симметрии нет. Если в этом законе заменить сложение умножением, а умножение сложением, то придем к абсурду:

$$(ab)+c=(a+c)(b+c).$$

Спрашивается, всегда ли это так? Не существует ли алгебры, в которой дистрибутивный закон был бы так же симметричен относительно добавления и умножения, как коммутативный и ассоциативный законы? Оказывается, существует алгебра, а именно **алгебра множеств**, в которой все три закона симметричны относительно действий сложения и умножения. Сходство между действиями сложения и умножения оказывается также в существовании двух замечательных чисел 0 и 1, таких, что сложение первого и умножение на второе не изменяют ни одного числа:

$$a+0=a,\ \ a\cdot 1=a.$$

Заметим, что второе соотношение выходит из первого заменой (+) на (·) и 0 на 1.

Однако и здесь сходство между действиями сложения и умножения не простирается особенно далеко. Так, число 0 играет немного особую роль в сравнении со всеми другими числами, в том числе и единицей.

Эта особая роль числа 0 следует из соотношения $a \cdot 0=0$. Если мы в этом выражении заменим (•) на (+) и 0 на 1, то приходим к соотношению $a+1=1$, что почти никогда не будет верным.

Как мы увидим дальше, сходство между нулем и единицей в алгебре множеств будет значительно большим, чем в обычной алгебре.

После этих предварительных замечаний можно приступить к детальному рассмотрению операций над множествами.

Объединение множеств

Объединением множеств X и Y называется множество, которое состоит из всех тех и только тех элементов, которые принадлежат хотя бы одному из множеств X, Y, т.е. принадлежат X или принадлежат Y. Объединение X и Y обозначается через $X \cup Y$. Формальное определение

$$X \cup Y=\{ x \mid x \in X \text{ или } x \in Y \},$$

Объединение множеств иногда называют суммой множеств и обозначают $X+Y$. Однако свойства объединения множеств несколько отличаются от свойств суммы при обычном арифметическом понимании. Поэтому этим термином мы пользоваться не будем.

Пример 1. Если $X=\{1, 2, 3, 4, 5\}$ и $Y=\{2, 4, 6, 7\}$, то

$$X \cup Y= \{1,2,3,4,5,6,7\}.$$

Пример 2. Если X — множество отличников в группе, а Y — множество студентов, которые живут в общежитии, то $X \cup Y$ — множество студентов, которые или учатся на отлично или проживают в общежитии.

Пример 3. Рассмотрим два круга, приведенных на рис 1. Если X — множество точек левого круга, а Y— множество точек правого круга, то $X \cup Y$ представляет собой заштрихованную область, которая ограничена обеими кругами.

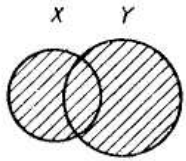


Рис. 1. Объединение множеств.

Понятие объединения можно распространить и на большее число множеств. Обозначим через $\Re = \{X_1, \ldots, X_n\}$ совокупность n множеств $X_1, \ldots, X_n$, которую называют иногда **системой множеств.** Объединение этих множеств

$$\bigcup_{i=1}^{n} X_i = \bigcup_{X \in \Re} X = X_1 \bigsqcup \ldots \bigcup X_n \qquad (1)$$

представляет собой множество, которое состоит из всех тех и только тех элементов, которые принадлежат хотя бы одному из множеств системы $\Re$.

Для объединения множеств справедливы коммутативный и ассоциативный законы

$$X \bigcup Y = Y \bigcup X; \qquad (2)$$

$$(X \bigcup Y) \bigcup Z = X \bigcup (Y \bigcup Z) = X \bigcup Y \bigcup Z, \qquad (3)$$

справедливость которых вытекает из того, что левая и правая части равенств состоят из тех самых элементов. Далее,

$$X \bigcup \varnothing = X. \qquad (4)$$

Это также очевидное соотношение, так как пустое множество не содержит элементов, а значит, X и $X \cup \varnothing$ состоят из одних и тех же элементов. Из (4) видно, что пустое множество $\varnothing$ играет роль нуля в алгебре множеств. Здесь имеет место аналогия с выражением $a+0=a$ в обычной алгебре.

Пересечение множеств

Пересечением множеств X и Y называется множество, которое состоит из всех тех и только тех элементов, которые принадлежат как множеству X, так и множеству Y. Пересечение множеств X и Y обозначается через $X \cap Y$. Формальное определение

$$X \cap Y = \{ x \mid x \in X \text{ и } x \in Y \} \qquad (5)$$

Пересечение множеств иногда называют произведением множеств и обозначают XY. Однако свойства пересечения множеств несколько отличаются от свойств произведения в обычном арифметическом понимании. Поэтому этим термином мы пользоваться не будем.

Пример 4. Для множеств X и Y в примере 1 $X \cap Y = \{2, 4\}$.

Пример 5. Для множеств X и Y в примере 2 $X \cap Y$ — множество отличников группы, которые живут в общежитии.

Пример 6. Рассмотрим два круга, которые приведены на рис. 2.

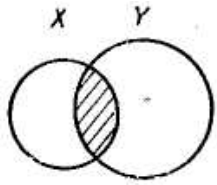


Рис. 2. Пересечение множеств

Если X — множество точек левого круга, а Y — множество точек правого круга, то $X \cap Y$ представляет собой заштрихованную область, являющуюся общей частью обоих кругов.

Операция пересечения позволяет установить ряд соотношений между двумя множествами.

Множества X и Y называются непересекающимися, если они не имеют общих элементов, т.е. если

$$X \cap Y = \varnothing. \qquad (6)$$

Пример 7. Непересекающимися множествами есть:

1) множества {1, 2, 3} и {4, 5, 6};

2) множество отличников и множество неуспевающих студентов в группе;

3) множество точек кругов X и Y на рис. 1.3.

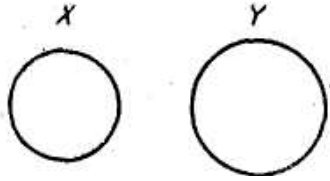


Рис. 3. Непересекающиеся множества

Говорят, что множества X и Y находятся *в общем положении,* если выполняются три условия:

- существует элемент множества X, который не принадлежит Y;
- существует элемент множества Y, который не принадлежит X;
- существует элемент, который принадлежит как X, так и Y.

Укажем одно отличие алгебры множеств от алгебры чисел. Если a и b — два числа, то между ними могут быть три соотношения или три возможности:

$$a<b,\ a=b,\ b<a. \qquad (7)$$

Для двух множеств X и Y, однако, может не выполняться ни одно из соотношений:

$$X \subset Y, \quad X = Y, \quad Y \subset X. \tag{8}$$

Так, если X — множество отличников, а Y — множество студентов, которые живут в общежитии, то три ранее приведенных соотношения будут означать:

$X \subset Y$ - каждый отличник обязательно живет в общежития;

$X = Y$ — в общежитии живут все отличники и только они;

$Y \subset Z$— все студенты, которые живут в общежитии, есть отличниками.

Очевидно, что эти соотношения не исчерпывают всех возможностей. На самом деле, как вытекает из предыдущих определений, между двумя множествами X и Y может быть одно из пяти отношений:

$$X \subset Y; \quad X = Y; \quad Y \subset X; \quad X \cap Y = \varnothing;$$

X и Y находятся в общем положении.

Понятие пересечения можно распространить и на большее, чем два, число множеств. Рассмотрим систему множеств $\Re = \{X_1..., X_n\}$. Пересечение этих множеств записывается в виде

$$\bigcap_{X \in \Re} X = \bigcap_{i=1}^{n} X_i = X_i \cap ... \cap X_n \tag{9}$$

и представляет собой множество, элементы которого принадлежат каждому из множеств системы $\Re$.

Нетрудно видеть, что пересечение множеств обладает коммутативным свойством

$$X \cap Y = Y \cap X \tag{10}$$

и ассоциативным

$$(X \cap Y) \cap Z = X \cap (Y \cap Z) = X \cap Y \cap Z. \tag{11}$$

Заметим также, что имеет место соотношения

$$X \cap \varnothing = \varnothing, \tag{12}$$

аналогичное соотношению $a \cdot 0 = 0$ в обычной алгебре.

Соотношение (12) вместе с соотношением (4) показывает, что пустое множество играет роль нуля в алгебре множеств.

Разность множеств

Данная операция в отличие от операций объединения и пересечение ***определяется только для двух множеств***. Разностью множеств X и Y называется множество, которое состоит из всех тех

и только тех элементов, которые принадлежат X и не принадлежат Y. Разность множеств X и Y обозначается через $X\backslash Y$. Таким образом,

$$X\backslash Y = \{ x \mid x \in X \text{ и } x \notin Y \}. \qquad (13)$$

Пример 8. Для множеств X и Y примера 1 $X\backslash Y=\{1, 3, 5\}$, $X\backslash Y= \{6, 7\}$. Если X и Y — множества из примера 2, то $X\backslash Y$— множество отличников, которые не проживают в общежитии. Для множества X и Y примера 3 $X \backslash Y$ — заштрихованная фигура на рис. 4.

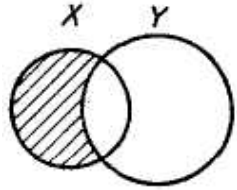


Рис. 4. Разность множеств

1.3. Универсальное множество и дополнения множеств

Универсальное множество

Как мы видели, роль нуля в алгебре множеств играет пустое множество. Спрашивается, не существует ли множество U, которое будет играть роль единицы, т.е. удовлетворять условию

$$X \cap U = X , \qquad (14)$$

аналогичному условию $a \cdot 1 = a$ в обычной алгебре. Соотношение (14) означает, что пересечение или «общая часть» множества U и множества X для любого множества X совпадает с самим этим множеством. Но это возможно лишь в том случае, если множество U содержит все элементы, из которых может состоять множество X, так что любое множество X полностью содержится в множестве U . Множество U, удовлетворяющее этому условию, называется *полным* или *универсальным*.

Исходя из сказанного, можно дать следующее определение универсального множества.

Определение. Если в некотором рассмотрении участвуют только подмножества некоторого фиксированного множества U, то это самое большое множество U называется *универсальным множеством*.

Следует отметить, что в разных конкретных рассмотрениях роль универсального множества могут играть разные множества. Так, при рассмотрении множеств студентов в группе (отличники; студенты, которые получают стипендию; студенты, которые проживают в общежитии, и т.п.) роль универсального множества играет множество студентов в группе.

Универсальное множество удобно изображать графически в виде множеств точек прямоугольника.

Отдельные области внутри этого прямоугольника будут означать разные подмножества универсального множества. Изображение множеств в виде областей в прямоугольнике, который представляет универсальное множество, называется *диаграммой Эйлера-Венна*.

Универсальное множество обладает интересным свойством, которое не имеет аналогии в обычной алгебре, а именно, для любого множества X справедливо соотношение

$$X \cup U = U. \tag{15}$$

Действительно, объединение $X \cup U$ представляет собой множество, в которое входят как все элементы множества X, так и все элементы множества U. Но множество U уже содержит в себе все элементы множества X, так что $X \cup U$ будет состоять из тех же элементов, которые и U, т.е. представляет собой само универсальное множество U.

Дополнение множества

Множество $\overline{X}$, которое определяется из соотношение

$$\overline{X} = U \setminus X, \tag{16}$$

называется дополнением множества X (к универсальному множеству U). На диаграмме рис. 5 множество $\overline{X}$ представляет собой незаштрихованную область.

Формальное определение

$$\overline{X} = \{ x \mid x \in U \text{ и } x \notin X \}.$$

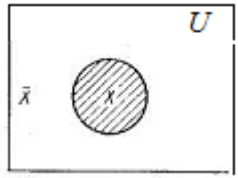


Рис. 5. Дополнение множества

Пример 9. Если $U = \{1, 2, 3, 4, 5, 6, 7\}$ и $X = \{3, 5, 7\}$, то $\overline{X} = \{1, 2, 4, 6\}$.

Из (16) следукт, что X и $\overline{X}$ не имеют общих элементов, так что

$$X \cap \overline{X} = \varnothing. \tag{17}$$

Кроме того, не имеется элементов U, которые не принадлежали бы ни X, нет $\overline{X}$, так как те элементы, которые не принадлежат X, принадлежат $\overline{X}$. Следовательно,

$$X \cup \overline{X} = U. \tag{18}$$

Из симметрии формулы (18) относительно X и $\overline{X}$ следует не только то, что $\overline{X}$ является дополнением X, но и то, что X является дополнением $\overline{X}$. Но дополнение $\overline{X}$ есть $\overline{\overline{X}}$. Таки образом

$$\overline{\overline{X}} = X. \tag{19}$$

С помощью операции дополнения можно в удобном виде представить разность множеств

$$X \setminus Y = \{ x \mid x \in X \text{ і } x \notin Y \} = \{ x \mid x \in X \text{ и } x \in \overline{Y} \},$$

т.е.

$$X \setminus Y = X \cap \overline{Y} \tag{20}$$

Разбиение множества

Одной из операций над множествами, которая более всего часто встречается, есть операция разбиения множества на систему подмножеств. Так, система курсов данного факультета является разбиением множества студентов факультета; система групп данного курса является разбиением множества студентов курса. Если N — множество натуральных чисел, а A_0 и A_1 - множество

парных и непарных чисел, то система $\{A_0, A_1\}$ будет разбиением множества N. Множество натуральных чисел может быть разбито иначе: на множество чисел, которые делятся на 3 без остатка, с остатком 1 и с остатком 2.

Продукция предприятия разбивается на систему множеств, которые состоят из продукции первого сорта, второго сорта и брака. Подобные примеры можно было бы продолжать бесконечно.

Для того, чтобы дать понятию разбиения строгое определение, рассмотрим некоторое множество M и систему множеств $\Re = \{X_1,..., X_n\}$. Система множеств $\Re$ называется разбиением множества M, если она удовлетворяет следующим условиям:

1)любое множество X из $\Re$ является подмножеством множества M:

$$\forall X \in \Re[X \subseteq M]; \tag{21}$$

2) любые два множества X и Y из $\Re$ являются непересекающимися:

$$\forall X, Y \in \Re[X \neq Y \rightarrow X \cap Y = \varnothing]; \tag{22}$$

3) объединение всех множеств, которые входят в разбиение, дает множество M:

$$\bigcup_{X \in \Re} X = M\ . \tag{23}$$

К понятию разбиения мы вернемся при рассмотрении отношения эквивалентности, с которым оно очень тесно связано.

Тождества алгебры множеств

С помощью операций объединения, пересечение и дополнение из множеств можно составлять различные алгебраические выражения. Обозначим через $\aleph(X,Y,Z)$ некоторое алгебраическое выражение, которое составлено из множеств X, Y и Z. Оно само представляет собой некоторое множество. Пусть $\Im(X, Y, Z)$ — другое алгебраическое выражение, которое составлено с тех же множеств. Если оба алгебраических выражения представляют собой одно и то же множество, то их можно приравнять друг к другу, получая алгебраическое тождество вида

$$\aleph(X,Y,Z) = \Im(X,Y,Z). \tag{24}$$

Такие тождества бывают очень полезны при преобразованиях алгебраических выражений над множествами, и некоторые из них мы рассмотрим в данном разделе.

1) На рис. 6 приведена диаграммы Эйлера-Венна для выражений

$$(X \cup Y) \cap Z \; i \; (X \cap Z) \cup (Y \cap Z).$$

Из этих диаграмм видно, что оба выражения определяют одно и то же множество, так что в алгебре множеств имеет место тождество

$$(X \cup Y) \cap Z = (X \cap Z) \cup (Y \cap Z). \qquad (25)$$

алогичное дистрибутивному закону $(a+b)c=ac+bc$ в обычной алгебре.

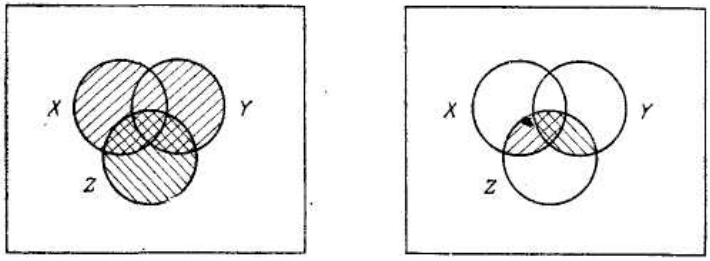


Рис. 6. Геометрическая иллюстрация тождества

$$(X \cup Y) \cap Z = (X \cap Z) \cup (Y \cap Z).$$

2. В обычной алгебре мы не можем заменить в дистрибутивном законе действие сложения умножением, а действие умножения сложением, так как это приводит к абсурдному выражению *(ab)+c=(a+c)(b+c)*. Иначе отстоит дело в алгебре множеств. На рис. 7 приведена диаграммы Эйлера-Венна для алгебраических выражений

$$(X \cap Y) \cup Z \; i \; (X \cup Z) \cap (Y \cup Z).$$

Оба эти выражения дают одно и то же множество, так что имеет место тождество

$$(X \cap Y) \cup Z = (X \cup Z) \cap (Y \cup Z). \qquad (26)$$

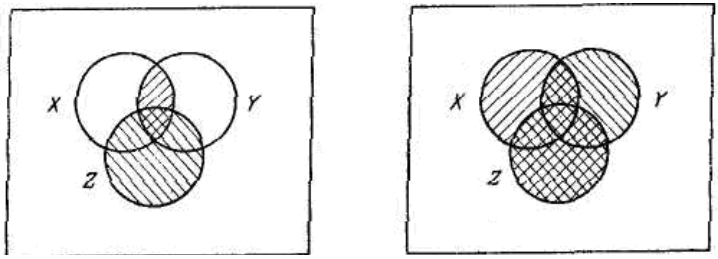


Рис. 7. Геометрическая иллюстрация тождества

$$(X \cap Y) \cup Z = (X \cup Z) \cap (Y \cup Z).$$

3. Легко убедиться, что если $Y \subseteq X$, то

$$X \cap Y = Y, X \cup Y = X. \qquad (27)$$

Действительно, все элементы множества Y являются в тот же время и элементами множества X. Значит, пересечение этих множеств, т.е. общая часть множеств X и Y, совпадает с Y. В объединение множеств X и Y множество Y не внесет ни одного элемента, который уже не входил бы в него, будучи элементом множества X. Следовательно, $X \cup Y$ совпадает с X.

4. Полагая в (27) $Y=X$ и учитывая, что $X \subseteq X$, находим

$$X \cap X = X, X \cup X = X. \qquad (28)$$

Установление тождеств алгебры множеств с помощью диаграммы Ейлера-Венна в ряде случаев оказывается неудобным. Есть более общий способ установления тождественности двух алгебраических выражений.

Пусть, как и раньше, через $\aleph(X,Y,Z)$ и $\Im(X, Y, Z)$ обозначены два алгебраических выражения, которые получены путем применения операций объединения, пересечения и дополнения к множествам X, Y и Z. Для того, чтобы доказать, что $\aleph = \Im$, достаточно показать, что $\aleph \subseteq \Im$ и что $\Im \subseteq \aleph$. В свою очередь, что бы показать, что $\aleph \subseteq \Im$, нужно убедиться в том, что из $x \in \aleph$ следует $x \in \Im$. Аналогично, что бы показать, что $\Im \subseteq \aleph$ нужно убедиться, что из $x \in \Im$ следует $x \in \aleph$.

Воспользуемся этим методом, чтобы доказать еще несколько тождеств.

5. Докажем тождество

$$\overline{X \cup Y} = \overline{X} \cap \overline{Y}. \qquad (29)$$

Предположим, что $x \in \overline{X \cup Y}$, т.е., что $x \notin X \cup Y$. Это значит, что $x \notin X$ и $x \notin Y$, т.е. $x \in \overline{X}$ и $x \in \overline{Y}$. Следовательно, $x \in \overline{X} \cap \overline{Y}$. Предположим теперь, что $y \in \overline{X} \cap \overline{Y}$, т.е. $y \notin \overline{X}$ и $y \notin \overline{Y}$. Это значит, что $y \notin X$ и $y \notin Y$, т.е., что $y \notin X \cup Y$. Следовательно, $y \in \overline{X \cup Y}$.

6. Тождество

$$\overline{X \cap Y} = \overline{X} \cup \overline{Y} \qquad (30)$$

докажем, приведя обе его части к одинаковому виду.

Выполняя операцию дополнения над обеими частями (30), получим

$$\overline{\overline{X \cap Y}} = \overline{\overline{X} \cup \overline{Y}}.$$

Левая часть этого выражения дает $X \cap Y$. То же самое получим, преобразовывая правую часть по правилу (29).

В литературе тождества (29) и (30) по обыкновению называются тождествами де-Моргана**.**

1.4. Свойства операций над множествами

Операции над множествами, которые сформулированы выше, как и операции над числами, обладают некоторыми свойствами (табл. 1). Эти свойства выражаются совокупностью тождеств, справедливых независимо от конкретного содержания входящих в них множеств, являющихся подмножествами некоторого универсума U.

Тождества (1а)—(3а) выражают соответственно *коммутативный, ассоциативный* и *дистрибутивный* законы для объединения, а тождества (1б)—(3б) — те же законы для пересечения. Соотношение (1а) - (7а) определяют свойства пустого множества Ø и универсума U относительно объединения, а соотношение (4 б)— (7 6) - относительно пересечения.

Выражения (8а) и (8б), которые называютья *законами идемпотентности,* позволяют записывать формулы с множествами без коэффициентов и показателей степени. Зависимости (9а) и (9б) представляют *законы поглощения,* а (10а) и (10б) — *теоремы де Моргана.*

Соотношение (11)-(20) отражают свойства дополнения, разности, дизъюнктивной суммы, включение и равенства.

Таблица 1

1а) $A \cup B = B \cup A$	1б) $A \cap B = B \cap A$
2а) $A \cup (B \cup C) = (A \cup B) \cup C$	2б) $A \cap (B \cap C) = (A \cap B) \cap C$
3а) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	3б) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
4а) $A \cup Ø = A$	4б) $A \cap U = A$
5а) $A \cup \overline{A} = U$	5б) $A \cap \overline{A} = Ø$
6а) $A \cup U = U$	6б) $A \cap Ø = Ø$
7а) $\overline{Ø} = U$	7б) $\overline{U} = Ø$
8а) $A \cup A = A$	8б) $A \cap A = A$
9а) $A \cup (A \cap B) = A$	9б) $A \cap (A \cup B) = A$
10а) $\overline{A \cup B} = \overline{A} \cap \overline{B}$	10б) $\overline{A \cap B} = \overline{A} \cup \overline{B}$

11) если $A \cup B = U$ и $A \cap B = Ø$, то $B = \overline{A}$

12) $\overline{A} = U \setminus A$

13) $\overline{\overline{A}} = A$

14) $A \setminus B = A \cap \overline{B}$

15) $A + B = (A \cap \overline{B}) \cup (\overline{A} \cap B)$

16) $A + B = B + A$

17) $(A + B) + C = A + (B + C)$

18) $A+\varnothing=\varnothing+A=A$

19) $A \subset B$, если и только если $A \cap B = A$ или $A \cup B = B$ или $A \cap \overline{B} = \varnothing$

20) $A = B$, если и только если $(A \cap \overline{B}) \cup (\overline{A} \cap B) = \varnothing$

Принцип двойственности. Первые десять свойств в табл. 1 представленные парами *двойственных (дуальных)* соотношений, одно из которых получается заменой в другом символов: $\square$ на $\cap$ и $\cap$ на $\cup$, а также $\varnothing$ на U и U на $\varnothing$. Соответствующие пары символов $\cup$, $\cap$ и $\varnothing$, U называются *двойственными (дуальными) символами.*

При замене в любой теореме символов, которые в нее входят, дуальными получим новое предложение, которое также является теоремой *(принцип двойственности или дуальности).* Тождества (11) и (12) не изменяются при замене символов дуальными, поэтому их называют *самодвойственными.*

Принцип дуальности можно распространить на разность и дизъюнктивную сумму, если использовать тождества (11) и (15). Аналогично в соответствии с соотношению (19) можно заменить $A \subset B$ на $A \cap B = A$ или $A \cup B = B$. Но поскольку дуальным $A \cap B = A$ есть $A \cup B = A$, то дуальным $A \subset B$ следует считать $B \subset A$. Поэтому, расширяя принцип дуальности на выражения, в которые входит символ включения, необходимо при переходе к дуальному выражению все знаки $\subset$ заменить на $\supset$ и обратно.

Метод доказательств. Доказательство тождеств (табл. 1) основано на отношении принадлежности. Чтобы убедиться, например, в справедливости тождестbf (3а), положим

$$x \in A \cup (B \cap C),$$

тогда $x \in A$ или $x \in (B \cap C)$. Если $x \in A$, то x принадлежит объединению A с любым множеством, т.е. $x \in A \cup B$ и $x \in A \cup C$ следовательно, x есть элемент пересечения множеств $A \cup B$ и $A \cup C$, т.е. $(A \cup B) \cap (A \cup C)$. Если $x \in B \cap C$, то $x \in B$ и $x \in C$,

следовательно $x \in A \cup B$ и $x \in C$ т.е. и в этом случае x есть элемент пересечения тех же множеств. Таким образом, доказано

$$A \cup (B \cap C) \subset (A \cup B) \cap (A \cup C).$$

Аналогично доказывается и соотношение

$$A \cup (B \cap C) \supset (A \cup B) \cap (A \cup C).$$

В соответствии с определению равенства множеств приходим к требуемому тождеству

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) .$$

Важно отметить, что любая теорема алгебры множеств и, в частности, соотношение, которые приведены в табл. 1.1, выводимы из первых пяти свойств, которые в свою очередь доказываются только в терминах отношения принадлежности. Это можно рассматривать как иллюстрацию аксиоматического подхода к алгебре множеств. Например, соотношение (8а*)* доказывается следующими преобразованиями с использованием тождеств (4б), (5а), (3а), (5б) и (4а):

$$A \cup A = (A \cup A) \cap U = (A \cup A) \cap (A \cup \overline{A}) = A \cup (A \cap A) = A \cup \varnothing = A.$$

Диаграммы Венна. Графические методы алгебры множеств основаны также на *диаграммах Венна.* Построение диаграммы начинается с разбиения плоскости на 2^n ячеек с помощью n фигур (замкнутых линий), где n — число различных множеств, которые принимают участие в данной совокупности соотношений. При этом каждая последующая фигура должна иметь одну и только одну общую область с каждой из ранее построенных фигур. Такое разбиение называют *символом Венна.* На рис. 8 показан символ Вєнна для n=3, который разбивает плоскость на 8 ячеек (внешняя область также считается ячейкой). Для определенного количества n переменных символ Венна имеет стандартный вид.

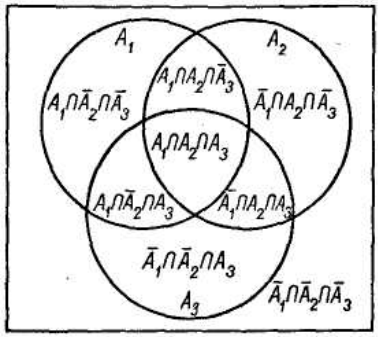


Рис. 8. Символ Венна при n=3.

Замкнутые области символа Венна, как и круги Ейлера, соответствуют переменным (множествам $A_1, A_2, \ldots, A_n$), а каждая ее область - пересечению $\prod_{i=1}^{n}\tilde{A}_i$, где символ ~ указывает, что под знаком пересечения стоит соответствующая переменная A_i, или ее дополнение $\overline{A}_i$: $(i=1, 2, \ldots, n)$. При этом внешняя область соответствует пересечению дополнений всех переменных

$$\prod_{i=1}^{n}\overline{A}_i = \overline{A}_1 \cap \overline{A}_2 \cap \ldots \cap \overline{A}_n .$$

Универсум отождествляется с плоскостью, которая может ограничиваться замкнутой линией, образующей какую-нибудь фигуру (прямоугольник, круг или овал).

Система теоретико-множественных соотношений отображается на символ Венна выделением (штриховкой) тех ячеек, которые отвечают пустым подмножествам. В результате и получаем диаграмму Венна. Объединение любой совокупности заштрихованных ячеек соответствует пустому множеству $\varnothing$, а объединение всех незаштрихованных ячеек дает универсум U.

Для отображения уравнения с правой частью, равной $\varnothing$, достаточно заштриховать области, соответствующие левой части уравнения. Уравнение $A=B$ преобразуется согласно формуле (20) (табл. 1) к виду

$$(A \cap \overline{B}) \cup (\overline{A} \cap B) = \varnothing .$$

Это значит, что следует заштриховать все те области в B, которые не входят в A, и те области в A, которые не входят в B. Включению $A \subset B$ на основании свойства 19 (табл. 1) соответствует уравнение $A \cap \overline{B} = \varnothing$. Его отображение на диаграмме осуществляется штриховкам ячейки, соответствующей пересечению A с дополнением B.

Применение диаграмм Венна. Построим, например, диаграмму Венна (рис. 9) для системы уравнений:

$$A = B \cup C; \quad B = \overline{C} \cup \overline{D}; \quad \overline{C} \cap \overline{D} = \varnothing; \quad A \cap D = B \cap C \cap D .$$

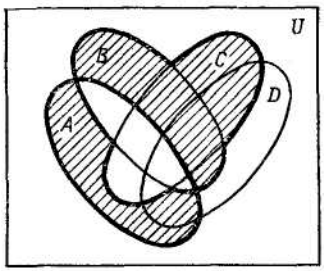
Отображение $A = B \cup C$ осуществляется штриховкам всех тех ячееек в B и C, которые не входят в A, а также всех ячеек в A, которые не входят ни в B, ни в C. Так как $B = \overline{C} \cup \overline{D}$, то в B следует заштриховать все, что входит одновременно в C и D, а в $\overline{C}$ *и* $\overline{D}$ - все, что не входит в B.

В силу $\overline{C} \cap \overline{D} = \emptyset$ штрихуется вся общая часть $\overline{C}$ *и* $\overline{D}$. Наконец, из $A \cap D = B \cap C \cap D$ следует, что $A \cap D$ есть общая часть B, C и D, и значит $A \cap D \subset B$ и $A \cap D \subset C$; $A \cap D$ входит и в B ив C, поэтому ячейки $A \cap \overline{B} \cap D$ и $A \cap \overline{C} \cap D$ должны быть заштрихованы. Кроме того, поскольку $B \cap C \cap D$ есть общая часть A и D, то $B \cap C \cap D \subset A$ и, следовательно, должна быть заштрихована ячейка $A \cap B \cap C \cap D$.

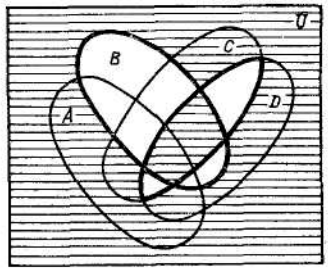
Последовательные этапы построения диаграммы Венна на рис. 9 отмечены штриховкой, наклоненной вправо, горизонтальной, вертикальной и наклоненной влево.

Как видно, единственная незаштрихованная ячейка соответствует $A \cap B \cap C \cap \overline{D} = U$, так как ею исчерпывается универсум U. Это возможно только в случае, когда $A = B = C = U$ и $D = \emptyset$, что и представляет собой решение системы уравнений.

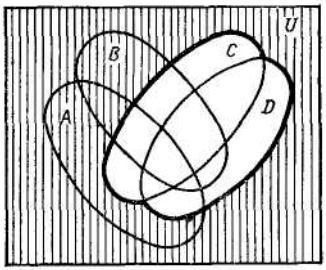
Покажем также, как решается уравнение $X \cup C = D$. Его отображение на диаграмме Венна осуществляется штриховкой ячеек в X и C, которые не входят в D, а также всех ячеек в D, которые не входят ни в X, ни в C (рис. 10). Из диаграммы Венна (заштрихованную часть не учитываем) получаем $D \setminus C \subset X \subset D$.



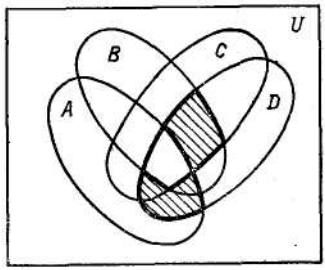
$A = BUC$



$B = \bar{C} U \bar{D}$



$\bar{C} \cap \bar{D} = \emptyset$



$A \cap D = B \cap C \cap D$

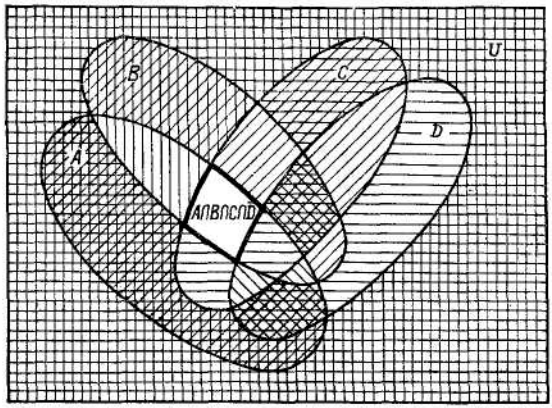


Рис. 9. Диаграмма Венна для системы уравнений

Важно отметить, что этот результат содержится в диаграмме Венна и не нужно никакой дополнительной информации, чтобы судить о свойствах решения уравнения.

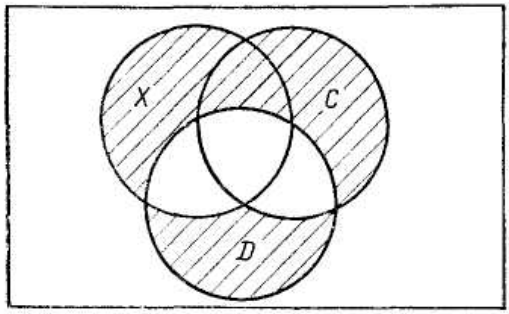


Рис. 10. Диаграмма Венна для уравнения $X \cup C=D$

Этот пример наглядно иллюстрирует различие между диаграммами Венна и кругами Ейлера, которые в литературе часто необоснованно называют диаграммами Ейлера-Венна. Иногда особенность диаграмм Венна видят только в том, что в них используются не круги, а овалы или другие замкнуты фигуры. Однако главное отличие не в этом (Ейлер тоже допускал применение фигур, которые не являются кругами). **Диаграмма Венна отображает систему соотношений на стандартном символе для *n* переменных путем «деформации» этого символа выделением (штриховкой) области пустых подмножеств.**

1.5. Упорядочение элементов и прямое произведение множеств

Упорядоченное множество

Наряду с понятием множества как совокупности элементов важным понятием являются понятия упорядоченного множества или кортежа.

Кортежем называется последовательность элементов, т.е. совокупность элементов, в которой каждый элемент занимает определенное место. Сами элементы при этом называются ***компонентами*** кортежа (первая компонента, вторая компонента и т.д.). Примеры кортежей: множество людей, которые стоят в очереди; множество слов в фразе; числа, которые выражают долготу и широту точки на местности, и т.п. Во всех этих множествах место каждого элемента является вполне определенным и не может быть произвольно измененно.

В технических задачах эта определенность часто является просто предметом договоренности. Так, только договоренностью можно объяснить, чему долготу ставят на первое место, а широту на второе. **Состояние кибернетической системы часто описывают множеством параметров, которые принимают числовые значения. При этом состояние системы — просто некоторое множество чисел.** Чтобы не оговаривать каждый раз, какое число что означает, устанавливают заранее, какой параметр считать первым, какой вторым и т.д., т.е. совокупность параметров представляют в виде упорядоченного множества. Так, если обозначить через h высоту самолета, а через v - его скорость, то кортеж $x=(h, v)$ будет описывать состояние самолета.

Число элементов кортежа называется его ***длиной***. Для обозначение кортежа будем использовать круглые скобки. Так, множество

$$a = (a_1, a_2, \ldots, a_n) \tag{31}$$

является кортежем длины n с элементами $a_1, a_2, \ldots, a_n$. Кортежи длины 2 называются парами или упорядоченными парами, кортеже длины 3-тройками, 4 — четверками и т.д. В общем случае кортежи длины n называются n-ками. Частным случаем кортежа является кортеж (a) длины 1 и пустой кортеж длины 0, обозначаемый ( ) или Λ. В отличие от обычного множества в кортеже могут быть и одинаковые элементы: два одинаковых слова в фразе,

одинаковые числовые значения долготы и широты точки на местности и т.п. В дальнейшем будем рассматривать упорядоченные множества, элементами которых являются вещественные числа. Такие упорядоченные множества называют точками пространства или векторами. Так, кортеж (a_1,a_2) может рассматриваться как точка на плоскости или вектор, проведенный из начала координат в данную точку (рис. 11, *а*). Компоненты a_1 и a_2 будут проекциями вектора на оси 1 и 2

$$\text{Пр}_1(a_1,a_2)=a_1; \quad \text{Пр}_2(a_1,a_2)=a_2.$$

Кортеж (a_1, a_2, a_3) может рассматриваться как точка в трехмерном пространстве или как трехмерный вектор, проведенный из начала координат в эту точку (рис. 11,*б*). Проекции вектора на оси координат

$$\text{Пр}_i(a_1, a_2, a_3)=a_i, \ i=1, 2, 3.$$

Однако в данном случае можно говорить о проекциях кортежа сразу на две оси, например 1 и 2, т.е. на координатную плоскость. Нетрудно видеть, что эта проекция представляет собой двухэлементный кортеж

$$\text{Пр}_{12}(a_1, a_2, a_3)=(a_1,a_2) .$$

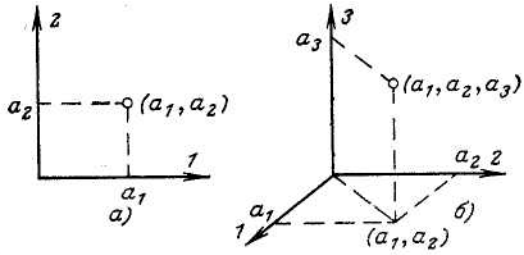


Рис. 11. Проекции двух- и трехэлементного кортежа.

Обобщая эти понятия, будем рассматривать упорядоченное *n*-элементное множество вещественных чисел $(a_1,..., a_n)$ как точку в воображаемом *n*-мерном пространстве, которое называют иногда гиперпространством, или как *n*-мерный вектор. При этом компоненты *n*-элементного кортежа a будем рассматривать как проекции этого кортежа на соответствующие оси

$$\text{Пр}_i\, a=a_i, \ i=1, 2, ...,n.$$

Если $i, j,...,l$ — номера осей, причем $1\le i<j<,...l\le n$, то проекция кортежа a на оси $i, j,...,l$ равна

$$\text{Пр}_{i, j,...,l}\, a=(a_i, a_j,..., a_l). \qquad (32)$$

Прямое произведение множеств

Прямым произведением множеств X и Y называется множество, обозначаемое $X \times Y$ и состоящее из всех тех и только тех упорядоченных пар, первая компонента которых принадлежит множеству X, а вторая принадлежит множеству Y. Таким образом, элементами прямого произведения являются двухэлементные кортежи вида (x, y). Формальное определение

$$X \times Y = \{(x, y) \mid x \in X, y \in Y\} \qquad (33)$$

Пример 10. Пусть X={1, 2}, Y={1, 3, 4}.

Тогда $X \times Y$ = {(1, 1), (1, 3), (1, 4), (2, 1), (2, 3), (2, 4)}.

Геометрическое представление этого множества приведено на рис. 12, *а*.

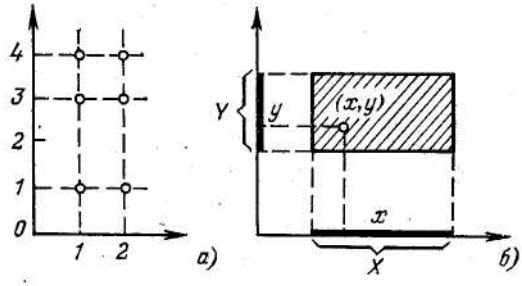


Рис. 12. Геометрическая иллюстрация прямого произведения множеств.

Пример 11. Пусть X и Y — отрезки вещественной оси. Прямое произведение $X \times Y$ изобразится заштрихованным прямоугольником, который показано на рис. 12,б. Из этого рисунка следует, что свойства прямого произведения отличаются от свойств обычного произведения в арифметическом смысле. В частности, прямое произведение изменяется при изменении порядка сомножителей, т.е.

$$X \times Y \neq Y \times X . \qquad (34)$$

Операция прямого произведения легко распространяется и на большее число множеств. Прямым произведением множеств $X_1, X_2, \ldots, X_r$ называется множество, которое обозначается $X_1 \times X_2 \times \ldots \times X_r$ и которое состоит из всех тех и только тех кортежей длины r, первая компонента которых принадлежит X_1, вторая X_2 и т.д.

Частным случаем операции прямого произведения является понятия *степеней* множества. Пусть M — произвольное множество.

Назовем s-й степенью множества M и обозначим через M^s прямое произведение s одинаковых множеств, равных M:

$$M^s = \underbrace{M \times M \times \cdots \times M}_{s \text{ разів}} \tag{35}$$

Это определение пригодно для s=2,3... Его можно расширить на любое целое неотрицательное s, если специальными определениями положить

$$M^1 = M, \quad M^0 = \{\Lambda\}. \tag{36}$$

Если R — множество вещественных чисел, то $R^2 = R \times R$ представляет собой вещественную плоскость, а $R^3 = R \times R \times R$ представляет собой трехмерное вещественное пространство.

Проекция множества

Операция проектирования множеств тесно связанна с операцией проектирования кортежа и может применяться лишь к таким множествам, элементами которых являются кортежи одинаковой длины.

Пусть M — множество, которое состоит из кортежей длины s. Тогда проекцией множества M будем называть множество проекций кортежей из M.

Пример 12. Пусть M={(1, 2, 3, 4, 5), (2, 1, 3, 5, 5), (3, 3, 3, 3, 3), (3, 2, 3, 4, 3)}.

Тогда Пр_2M={2, 1, 3}; $\text{Пр}_{2,4}M$ = {(2, 4), (1.5), (3, 3)}.

Легко проверить, что если $M = X \times Y$, то

$$\text{Пр}_1M = X; \quad \text{Пр}_2M = Y, \tag{37}$$

а если $Q \subseteq X \times Y$,

то

$$Пр_1Q \subseteq X; \quad Пр_2Q \subseteq Y. \tag{38}$$

1.6. Соответствия

В математике, как и в жизни, разные объекты могут чему-то отвечать или не отвечать. Находиться между собой в определенных отношениях или наоборот - не находиться. И основой формализации (математизации) здесь также служат множества.

Т.е. между множествами могут устанавливаться разные СООТВЕТСТВИЯ и ОТНОШЕНИЯ. Больше того (а серьезные математики может быть даже сказали бы "прежде всего"), множества нередко могут ОТОБРАЖАТЬСЯ одно в другое и даже в самих себя.

Человек может соответствовать профессии, зарплата соответствовать должности, наказание - преступлению, оценка - знаниям.

Глядя на многочисленные примеры вокруг мы замечаем, что для определение конкретного соответствия необходимо определить два множества: множество (область) определений и множество (область) значений. А также определить "пары соответствий". Например, область определения - группа В-005, которая сдает экзамен; область значений - отлично, хорошо, удовлетворительно, неудовлетворительно - множество оценок. И множество пар: Иванов - отличник, Петров - хорошист, Сидоров - отличник. А Петренко - не явился. Вот вам и готовое соответствие.

Соответствия имеют свойства.

1. В данном случае соответствие НЕ-ВСЮДУ-ОПРЕДЕЛЕННО, поскольку для Петренко в этом соответствии нет пары. (Даже если бы мы написали в ведомости Петренко - н/з, то это все равно бы не попало в соответствие, поскольку "н/з" нет в множестве допустимых значений!). Если бы деканат своевременно исключил из ведомости Петренка, как отчисленного, то это соответствие стало бы ВСЮДУ-ОПРЕДЕЛЕННЫМ.

2. Соответствие ФУНКЦИОНАЛЬНОЕ, поскольку каждому студенту отвечает не больше одной оценки. Такое соответствие называют по-простому, ФУНКЦИЕЙ. В данном случае через Петренка это не всюду определенная функция. Никакой разности со школьной функцией кроме той принципиальной, что здесь аргументами и значениями могут быть не только числа, а любые объекты.

Если бы за один экзамен студенты могли получать несколько оценок, то соответствие было бы НЕФУНКЦИОНАЛЬНЬНЫМ. Т.е. не было бы функцией. (Оно было бы "многозначной [недетерминированнной] функцией", но это уже другая математика).

3. Данное соответствие НЕИНЪЕКТИВНО, поскольку отлично получил больше, чем один студент. Если бы Сидоров, через фатальную склонность к несчастьям, получил не отлично, а удовлетворительно (или не удовлетворительно), то соответствие было бы ИНЪЕКТИВНО. Получение студентами олимпийских медалей за победу в беге на 100 метров было бы примером инъективного соответствия.

4. Данное соответствие НЕСЮРЪЕКТИВНО, поскольку на экзамене были использованные не все возможные оценки. На реальных экзаменах обычно бывает задействован весь возможный спектр оценок, поэтому это соответствие бывает "по жизни" СЮРЪЕКТИВНО.

5. Соответствие, которое одновременно ВСЮДУ-ОПРЕДЕЛЕНО, ФУНКЦИОНАЛЬНО, ІНЪЕКТИВНО и СЮРЪЕКТИВНО называется БИЕКТИВНЫМ. Еще его называют ВЗАИМНО-ОДНОЗНАЧНЫМ. Довольно убедительный пример биективного соответствия голова на плечах. Возьмите множество голов, множество плеч и убедитесь во всех четырех свойствах.

Выделение соответствий в отдельную категорию предложили европейцы, а точнее французы, а еще точнее, Николя Бурбаки (это французский Козьма Прутков, который состоял из математиков-интеллектуалов). Американская школа считает соответствие частным случаем отношений. У нас разговор об отношении будет идти отдельно - так легче разложить все по полочкам.

Формальное определение соответствия

Рассмотрим два множества X и Y. Элементы этих двух множеств могут каким-либо образом сопоставляться друг с другом, образуя пары (x, y). Если способ такого сопоставления определен, т.е. для каждого элемента $x \in X$ указан элемент $y \in Y$, с которым сопоставляется элемент x, то говорят, что между множествами X и Y установлено соответствие. При этом совсем необязательно, чтобы в сопоставлении участвовали все элементы множеств X и Y.

Для того чтобы задать соответствие, необходимо указать:

1) множество X, элементы которого сопоставляются с элементами другого множества;

2) множество Y, с элементами которого сопоставляются элементы первого множества;

3) множество $Q \subseteq X \times Y$, определяющее закон, в соответствии с которым осуществляется соответствие, т.е. перечисляющее все пары

(x,y), участвующие в сопоставлении. Таким образом, **соответствие**, обозначаемое *q*, **представляет собой тройку множеств**

$$q = (X,\ Y, Q), \qquad (39)$$

в которой $Q \subseteq X \times Y$. В этом выражении первая компонента X называется ***областью отправления соответствия***, вторая компонента Y называется ***областью прибытия соответствия***, третья компонента Q называется ***графиком соответствия***. Термин «график» будет более подробно разъяснен при рассмотрении частного вида соответствия, называемого ***функцией***.

Кроме трех рассмотренных множеств X, Y, Q, с каждым соответствием неразрывно связаны еще два множества: это множество $\text{Пр}_1 Q$, которое называется ***областью определения соответствия***, в которое входят элементы множества X, участвующие в сопоставлении, и множество $\text{Пр}_2 Q$, которое называется ***областью значений соответствия***, в которое входят элементы множества Y, участвующие в сопоставлении.

Если $(x, y) \in Q$, то говорят, что элемент y соответствует элементу x. Геометрически это удобно изображать стрелкой, направленной от x к y.

Пример 13. Пусть $X = \{1, 2\}$, $Y = \{3, 5\}$, так что

$$X \times Y = \{(1, 3),\ (1, 5), (2, 3),\ (2, 5)\}.$$

Это множество дает возможность получить 16 различных соответствий. Приведем некоторые из них:

$$Q_1 = \{(1, 3)\};\ Пр_1 Q_1 = \{1\};\ Пр_2 Q_1 = \{3\};$$

$$Q_2 = \{(1, 3),(1, 5)\};\ Пр_1 Q_2 = \{1\};\ Пр_2 Q_2 = \{3, 5\} = Y.$$

Пример 14. На предприятии есть три автомашины: две грузовые α и β, котрые работают в две смены, и автобус γ, который используется редко. Машина β находится в ремонте. В штате имеется три шофера a, b, c, из которых c находится в отпуске. Распределение шоферов по машинам представляет собой соответствие. Одним из возможных соответствий будет следующее:

$$q = (\{a, b, c\},\ \{\alpha, \beta, \gamma\},\ \{(a, \alpha), (a, \gamma), (b, \alpha)\}).$$

Геометрически это соответствие изображено на рис. 13, *а*. В нем элемент α соответствует элементам a и b, а элемент γ соответствует элементу a. Соответствие q определено на a и b, но не определено на c, следовательно, областью определения соответствия есть множество $\{a, b\}$. Областью значений соответствия есть множество $\{\alpha, \gamma\}$.

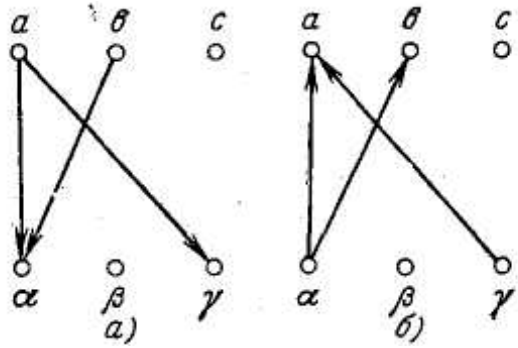


Рис. 13. Геометрическое представление прямого и обратного соответствий

Если $\Pr_1 Q = A$, то соответствие называется *всюду определенным* (в противном случае соответствие называется частным); если $\Pr_2 Q = Y$, то соответствие называют *сюръективным*.

Множество всех $b \in B$, соответствующих элементу $a \in A$, называется *образом* a в B при соответствии Q. Множество всех a, которым соответствует b, называется *прообразом* b в A при соответствии Q. Если $C \subseteq \Pr_1 Q$, то образом множества C называется объединение образов всех элементов C. Аналогично определяется прообраз множества D для любого $D \subseteq \Pr_2 Q$.

Соответствие Q называется *функциональным* (или однозначным), если образом любого элемента из $\Pr_1 Q$ является единственный элемент из $\Pr_2 Q$. Соответствие Q между A и B называется *взаимно-однозначным* (иногда пишут «1-1-соответствие»), если оно всюду определено, сюръективно, функционально и, кроме того, прообразом любого элемента из $\Pr_2 Q$ является единственный элемент из $\Pr_1 Q$.

Пример 15. Круг Q радиуса 1 с центром в точке (3, 2) (рис. 14), т.е. множество пар действительных чисел (x, y), которые удовлетворяют соотношению $(x-3)^2+(y-2)^2 \le 1$, задает соответствие между R и R (осью абсцисс и осью ординат). Образом числа 4 при этом соответствии является единственное число 2, образом числа 3 является отрезок [1, 3] осы ординат; этот же отрезок [1,3] является образом отрезка [2, 4] осы абсцисс, который, в свою очередь, служит прообразом числа 2. Данное соответствие не является функциональным. Примером функционального соответствия между действительными числами на том же рис. 14 служит дуга ABC.

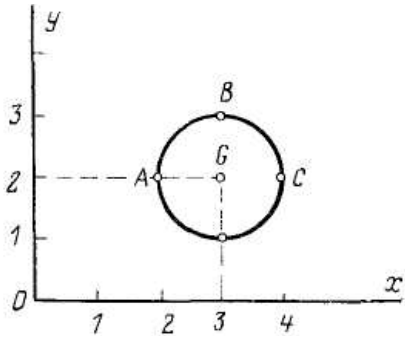


Рис. 14.

Еще раз напомним, что **для задания соответствия необходимо указать не только множество *Q*, но и множества *A* и *B***, т.е. указать, подмножеством какого именно прямого произведения является Q. В данном примере тот же круг Q_1 задает и другое соответствие: между отрезком [2, 4] и отрезком [1, 3]. При этом по некоторым свойствам соответствия

$$Q_1 \subseteq R^2 \text{ и } Q_1 \subseteq [2, 4]\times[1, 3]$$

отличаются: например, второе соответствие в отличие от первого всюду определено и сюръективно. Учитывая эти соотношения, следовало бы определять соответствие как тройку множеств (Q, A, B). Тогда не пришлось бы оговариваться, что один круг может задавать два соответствия; это и так было бы ясно из различия троек (Q_1, R, R) и $(Q_1, [2, 4], [1, 3])$. Однако такие оговорки приходится делать редко: либо множества A и B ясны из контекста, либо различия в их выборе не влияют на исследуемые свойства соответствия. Поэтому «определение через тройку множеств» здесь использоваться не будет.

2) Англо-русский словарь устанавливает соответствие между множеством английских и русских слов. Это соответствие не является функциональным (так как одному английскому слову, как правило, ставится в соответствие несколько русских слов); кроме того, оно практически никогда не является полностью определенным: всегда можно найти английское слово, которое не содержится в данном словаре.

(При этом остается в стороне вопрос (вообще говоря, законный), является ли множество английских слов (равно как и русских) точно заданным множеством).

3) Позиция на шахматной доске представляет собой взаимно-однозначное соответствие между множеством фигур, которые остались на доске, и множеством занятых ими полей.

4) Множество векторов вида $(n,\ 2^n)$ где $n \in N$, задает взаимно-однозначное соответствие между множеством натуральных чисел и множеством M_{2^n} степеней двойки.

Обратное соответствие

Для каждого соответствия

$$q = (X, Y, Q), \quad Q \subseteq X \times Y$$

существует обратное соответствие, которое получается, если данное соответствие рассматривать в обратном направлении, т.е. определять элементы $x \in X$, с которыми сопоставляются элементы $y \in Y$. Соответствие, обратное соответствию q, будем обозначать

$$q^{-1} = (Y, X, Q^{-1}), \tag{40}$$

где $Q^{-1} \subseteq Y \times X$.

Пример 16. Обратным соответствием для примера 14 будет закрепление автомашин за шоферами

$$(\{\alpha, \beta, \gamma\}, \{a, b, c\}, \{(\alpha, a), (\alpha, b), (\gamma, a)\}),$$

что геометрически показано на рис. 13,б.

Из приведенного примера видно, что геометрическое представление обратного соответствия получается путем изменения напрявлення стрелок в геометрическом представлении прямого соответствия. Отсюда следует, что обратным соответствием обратного соответствия будет прямое соответствие

$$(q^{-1})^{-1} = q. \tag{41}$$

Композиция соответствий

Композицией соответствий называется последовательное применение двух соответствий.

Композиция соответствий есть операция с тремя множествами X, Y и Z, на которых определены два соответствия

$$\left.\begin{aligned} q &= (X, Y, Q), \quad Q \subseteq X \times Y; \\ p &= (Y, Z, P), \quad P \subseteq Y \times Z, \end{aligned}\right\} \tag{42}$$

причем область значений первого соответствия совпадает с областью определения второго соответствия:

$$\text{Пр}_2 Q = \text{Пр}_1 \Gamma. \tag{43}$$

Первое соответствие определяет для любого $x \in \Pr_1 Q$ некоторый, возможно и не один, элемент $y \in Y$. Согласно определению операции композиции соответствий теперь нужно для найденного $y \in Y$ определить $z \in Z$, воспользовавшись вторым соответствием. Таким образом, композиция соответствий сопоставляет с каждым элементом x из области определения первого соответствия $\Pr_1 Q$ один или несколько элементов z из области значений второго соответствия $\Pr_2 Q$.

Композицию соответствий q и p будем обозначать $q(p)$, а график композиции соответствий — через $Q \circ P$. При этом композиция соответствий (1.42) запишется в виде

$$q(p) = (X,\ Z,\ Q \circ P),\ Q \circ P \subseteq X \times Z. \qquad (44)$$

Пример 17. Если q — соответствие, которое определяет распределение шоферов по автомашинам, а p — соответствие, которое определяет распределение автомашин по маршрутам, то соответствие $q(p)$ есть соответствие, которое определяет распределение шоферов по маршрутам.

Естественно, что операцию композиции можно распространить и на больше, чем два числа соответствий.

Взаимно-однозначные соответствия и мощности множеств.

Если между конечными множествами A и B существует взаимно-однозначное соответствие, то $|A|=|B|$. Действительно, если это не так, то либо $|A|>|B|$ и тогда, поскольку соответствие всюду определено, в A найдутся два элемента, которым соответствует один и тот же элемент $b \in B$ (нарушится единственность образа), либо $|A|<|B|$ и тогда, поскольку отображение сюръективно, в B найдутся два элемента, соответствующих одному и тому же самому $a \in B$ (нарушится единственность прообраза). (Обращаем внимание на то, что в этом простом рассуждении оказываются существенными все четыре свойства взаимно-однозначного соответствия).

Этот факт, во-первых, позволяет установить равенство мощностей двух множеств, не вычисляя этих мощностей, а во-вторых, часто дает возможность вычислить мощность множества, установив его взаимно-однозначное соответствие с множеством, мощность которого известна или легко вычисляется. В качестве иллюстрации этого приема

приведем доказательство важной теоремы о числе подмножеств конечного множества.

Теорема. Если для конечного множества A $|A|=n$, то число всех подмножеств A равно 2^n, т.е. $2^{|A|}$.

Доказательство. Занумеруем элементы A номерами от 1 до n: $A=\{a_1, a_2,..., a_n\}$ и рассмотрим множество B_n двоичных векторов из нулей и единиц длины n. Каждому подмножеству $A^* \subseteq A$ поставим в соответствие вектор $v = (v_1, v_2, ..., v_n) \in B_n$ следующим образом:

$$v_i = \begin{cases} 0, & \text{если} \quad a_i \notin A^*; \\ 1, & \text{если} \quad a_i \in A^*. \end{cases}$$

Например, если $A=\{a, b, c, d, e\}$, то подмножеству $\{a, c, d\}$ соответствует вектор (1, 0, 1, 1, 0), а подмножеству $\{b\}$ соответствует вектор (0, 1, 0, 0, 0). Пустому подмножеству любого A соответствует вектор из одних нулей, а самому A — вектор из одних единиц. Очевидно, что установленное соответствие между множеством всех подмножеств A и двоичными векторами длины n является взаимно-однозначным и число подмножеств A равно $|B_n|$. А так как B_n является прямым произведением n двухэлементных множеств $\{0, 1\}$, то в силу следствия, что $|A^n|=|A|^n$ имеем $|B_n| = 2^n$.

Множества *равномощны*, если между их элементами можно установить взаимно-однозначное соответствие. Для конечных множеств это утверждение доказывается, что и было сделано ранее. Для бесконечных множеств оно является определением равномощности. Множества, равномощные N, называются *счетными*. Соответствие, которое установлено в примере 16, п. 4, показывает, что множество M_{2^n} счетно. Вообще любое бесконечное подмножество N счетно. Действительно, пусть $N' \subset N$. Выберем в N' наименьший элемент и обозначим его n_1; в $N'\setminus\{n_1\}$ выберем наименьший элемент и обозначим его n_2; наименьший элемент $N'\setminus\{n_1, n_2\}$ обозначим n_3 и т.д. Поскольку для всякого натурального числа имеется лишь конечное множество меньших натуральных чисел, то любой элемент N' рано или поздно получит свой номер. Эта нумерация, т.е. соответствие (n_i, i), и есть взаимно-однозначное соответствие между N' и N.

Множество N^2 счетно. Нумерацию N^2 можно устроить таким образом. Разобьем N^2 на классы. К первому классу N^2_1 отнесем все пары чисел с минимальной суммой. Таких пар всего одна: (1, 1). Ко

второму классу N^2_2 отнесем все пары чисел с суммой 3: $N^2_2 = \{(1, 2), (2, 1)\}$. В общем случае $N^2_i = \{(a, b) | a+b=i+1\}$. Каждый класс N^2_i содержит ровно i пар. Упорядочим теперь классы по возрастанию индексов i, а пары внутри класса — по возрастанию первого элемента и занумеруем получившуюся последовательность пар номерами 1, 2, 3 ... Легко видеть, что если $a+ b=i+1$, то пара $\{a, b\}$ получит номер $1+2 + ... + (i— 1)+a$. Эта нумерация и доказывает счетность N^2, из которой, в свою очередь, непосредственно следует счетность множества P положительных рациональных чисел, т.е. дробей вида a/b, где a и b — натуральные числа. Аналогично находится счетность N^3 и вообще N^k для любого натурального k.

(На примере множества P видно, что нумерация числового множества может не иметь ничего общего с упорядочением его элементов по величине. В множестве P нет ни наименьшего элемента, ни двух соседних по величине элементов (для любых двух дробей p_1 и p_2 всегда найдется дробь, которая лежит между ними, например $(p_1+ p_2)/2$, однако существует элемент с наименьшим номером и элементы с соседними номерами.)

Нетрудно понять, что объединение конечного числа счетных множеств M_1, M_2, ..., M_k счетно. Действительно, перенумеруем сначала все первые элементы множеств, потом все вторые и т.д. Объединение счетного множества конечных множеств также счетно (сначала нумеруем все элементы первого множества, потом все элементы второго множества и т.д.). Из последнего утверждения следует, что множество всех слов в любом конечном алфавите счетно. Менее очевидно, что счетно и объединение счетного множества счетных множеств. Примером такого объединения есть множество

$$\bigcup_{i \in N} N^i$$

всех векторов с натуральными компонентами.

Множество всех действительных чисел отрезка $[0, 1]$ не является счетным *(теорема Кантора)*. Действительно, предположим, что оно счетно и существует его нумерация. Расположим все числа, которые изображены бесконечными десятичными дробями, в порядке этой нумерации:

$$0, a_{11}\, a_{12}\, a_{13} ...$$
$$0, a_{21}\, a_{22}\, a_{23} ...$$
$$0, a_{31}\, a_{32}\, a_{33} ...$$
$$.....$$

Рассмотрим любую бесконечную десятичную дробь $0, b_1, b_2, b_3 \ldots$, такую, что $b_1 \neq a_{11}$, $b_2 \neq a_{22}$, $b_3 \neq a_{33}$ и т.д. Эта дробь не может войти в указанную последовательность, так как от первого числа она отличается первой цифрой, от второго числа — второй цифрой и т.д. Следовательно, все числа из отрезка $[0, 1]$ не могут быть пронумерованы и множество всех действительных чисел отрезка $[0, 1]$ *несчетно*. Его мощность называется, как мы уже говорили, *континум;* множества такой мощности называются *континуальными.* Метод, который использован при доказательстве, называется *диагональным методом* Кантора.

Множество всех подмножеств счетного множества континуально. Это становится ясным, если воспользоваться, как и в приведенной раньше теореме, представлением подмножества в виде последовательности (но теперь уже бесконечной!) нулей и единиц: на i-м месте стоит 1, если i-й элемент множества входит в данное подмножество, и 0 в противном случае. Получаем взаимно-однозначное соответствие между подмножествами счетного множества и правильными двоичными дробями, которые в свою очередь, однозначно соответствуют множеству чисел отрезка $[0, 1]$. Как показывается в теории множеств (с помощью метода, аналогичного диагональному), для множества любой мощности множество его подмножеств имеет более высокую мощность. Поэтому **не существует множества максимальной мощности**. Парадокс Кантора именно и заключается в том, что «множество всех множеств» должно содержать все множества и, следовательно, иметь максимальную мощность, которая противоречит результатам теории множеств.

Примеры решения типовых задач

Операции над множествами

Объединение множеств

1.1. a) $A=\{a,b,d\}$, $B=\{b,d,c,h\}$, $A \sqcup B=\{a,b,d,c,h\}$.

b) $M_3 \cup M_4 = M_3 = M_4$ (так как M_3 и M_4 равны)

c) Обозначим футбольные команды высшей лиги Φ_i*:*

$$M_7=\{\Phi_1, \Phi_2, \ldots, \Phi_{18}\}.$$

Тогда $\bigcup_{i=1}^{18} \Phi_i$ - множество всех футболистов (но не команд!) высшей лиги.

d) Обозначим через N_k множество всех натуральных чисел, которые делятся на k и не равны k, а через P – множество всех простых чисел (принято считать, что $1 \notin P$). Тогда $\bigcup_{i \in P} N_i$ – множество всех составных, т.е. непростых чисел.

Пересечение множеств

1.2. a) $A=\{a,b,d\}$, $B=\{b,d,c,h\}$, $A \cap B=\{b,d\}$.

b) $M_3 \cap M_4 = M_3 = M_4$

c) $\bigcap_{i=1}^{18} \Phi_i = \varnothing$; более того, для любых i и j $\Phi_i \cap \Phi_j = \varnothing$.

d) $\bigcap_{i \in P} N_i = \varnothing$ (обозначение те же, что и в примере 1.1,d), так как элемент такого множества должен делиться на все простые числа; ввиду бесконечности множества простых чисел это невозможно.

Разность множеств

1.3. a) $A=\{a,b,d\}$, $B=\{b,d,c,h\}$, $A \setminus B=\{a\}$, $A \setminus B=\{c,h\}$,

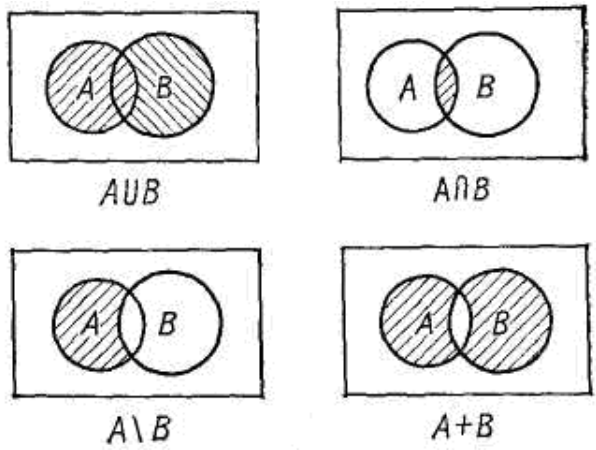
b) $M_3 \setminus M_4 = M_4 \setminus M_3 = \varnothing$.

c) $M_7 \setminus M_6$ – множество всех команд высшей лиги, за исключением „Ворсклы”.

d) Дизъюнктивная сумма (симметричная различность)
$A+B$:$\{1,2,3\}+\{2,3,4\}=\{1,4\}$.

Применение кругов Эйлера

Для наглядного изображения соотношений между подмножествами которого-небудь универсума U используют круги Эйлера (см. приведенний ниже рисунок). Множества, которые получают по результатам операций над множествами A и B, изображены на рисунке заштрихованными областями



Прямое произведение множеств

1.4.1) Множество $R \times R = R^2$ — это множество точек плоскости, точнее, пар вида (a, b), где $a, b \in R$ и являются координатами точек плоскости.

Координатное представление точек плоскости, предложенное французским математиком и философом Декартом, исторически первый пример прямого произведения. Поэтому иногда прямое произведение называют декартовым.

2) $A = \{a, b, c, d, e, f, g, h\}$, $B = \{1, 2, ... 8\}$. Тогда $A \times B = \{a1, a2, a3, ..., h7, h8\}$ — множество, которое содержит обозначение всех 64 клеток шахматной доски.

3) Рассмотрим множество числовых матриц 3×4, т.е. матриц вида

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{vmatrix}$$

где a_{ij} принадлежат множеству R действительных чисел. Строки матрицы — это элементы множества R^4. Сама матрица, которая рассматривается как упорядоченный набор (т.е. вектор) строк, — это элемент множества $(R^4)^3=R^4\times R^4\times R^4$. Компоненты матрицы, заданной таким образом, — строки, а не числа. Поэтому $(R^4)^3 \neq R^{12}$. Содержательный смысл этого неравенства состоит в том, что в векторе из R^{12} не содержится никакой информации о строении матрицы; тот же вектор из R^{12} мог бы перечислять элементы матриц 4×3 или 2×6,

которые как математические объекты вовсе не совпадают с матрицами 3×4.

Приведенный пример показывает, в частности, что компонентами векторов могут быть также векторы.

4) Пусть A— конечное множество, элементами которого являются символы (буквы, цифры, разделительные знаки, знаки операций и т.д.). Такие множества обычно называют *алфавитами*. Элементы множества A^n называют *словами* длины n в алфавите A. Множество всех слов в алфавите A — это множество

$$\bigcup_{i\in N} A^i = A^1 \cup A^2 \cup A^3 \cup \ldots$$

При написанные слов (которые по нашему определению являются векторами) не принято пользоваться ни запятыми, ни скобками как разделителями; зато они могут оказаться символами самого алфавита. Поэтому слово в алфавите A — это просто конечная последовательность символов алфавита A. Например, десятичное целое число - это слово в алфавите цифр {0, 1,.., 9}. Текст, напечатанный на принтере, является словом в алфавите, которое определяется клавиатурой данного принтера (включая разделительные знаки и пробел!).

Проекция множества

6.1) Проекция точки плоскости на 1-ю ось — это ее абсцисса (первая координата); проекция на 2-ю ось — ордината.

2) $V = \{(a, b, d), (c, b, d), (d, b, b)\}$, $\text{пр}_1 V = \{a, c, d\}$, $\text{пр}_2 V = \{b\}$, $\text{пр}_{2,3} V = \{(b, d), (b, b)\}$.

Индивидуальные тестовые задачи

Упражнение 1.1.

1. Рассмотреть по крайней мере две возможные интерпретации множества

{Петренко, Петренко, Симоненко).

Определить каждую из них настолько однозначно, насколько это возможно.

2. Рассмотреть следующие четыре множеств. Выяснить, какие записи этих множеств могут быть упрощены и какие из них эквивалентны. Предложить все возможные интерпретации, которые удовлетворяют описанным выше предположениям:

а) {1, 2, 3, 4}; б) {I, II, III, IV, V};

в) {1, один, one, uno, ein}; г) {5, V, пять, five}.

3. Проверить справедливость утверждений

«Это утверждение неверно» и «я лгун».

Какие слова в утверждении требуют собственного (т.е. математически точного) определение для того, чтобы сделать ответ математически строгим?

4. Пусть X - множество $\{1, 2\}$, а Y — множество $\{x: x=y+z; y, z\in X\}$. Определить в явном виде множество Y.

Какие это множества:

$$\{y: y=x+z; x,z\in X\} \text{ и } \{y: x=y+z; z\in X\}?$$

5. Предположим, что x является элементом, а не множеством. Тогда $y\notin x$ для каждого y, и отсюда следует, что $x\notin x$. Можно ли упростить множество $\{x, \{x\}, \{\{x\}\}\}$? Что можно сказать относительно $\{x, y, \{x, y\}\}$?

6. Пусть A — множество всех целых чисел. Описать словами множество

$$X=\{x: x\in A \text{ и } x=1 \text{ или } (x-2)\in X\}.$$

Упражнение 1.2. 1. Пусть

$$U=\{1,2,3,4\},\quad X=\{1,5\},\ Y=\{1,2,4\},\quad Z=\{2,5\}.$$

Найти множества:

а) $X\square Y'$; *б)* $(X\cap Z)\cup Y'$; *в)* $X\cup(Y\cap Z)$;

г) $(X\cup Y)\cap(X\cup Z)$; *д)* $(X\cup Y)'$; *е)* $X'\cap Y'$; *же)* $(X\cap Y)'$;

з) $(X\cup Y)\cup Z$; *и)* $X\cup(Y\cup Z)$; *к)* $X\backslash Z$; *л)* $(X\backslash Z)\cup(Y\backslash Z)$,

(Объяснение: Y'- *дополнение множества)*

2. Пусть

$$U=\{a, b, c, d, e, f\},\ A=\{a, b, c\},\ B=\{f,e,c,a\},\ C=\{d,e,f\}.$$

Найти множества:

а) $A\backslash C$; б) $B\backslash C$; в) $C\backslash B$; г) $A\backslash B$; д) $A'\cup B$; е) $B\cap A'$; ж) $A\cap C$;

3. Даны два произвольных множества A и B такие, что $A\cap B=\varnothing$. Что представляют собой множества $A\backslash B$ и $B\backslash A$?

4. Даны два произвольных множества C и D такие, что $C\cap D'=\varnothing$. Что можно сказать о $C\cap D$ и $C\cup D$?

5. Дано произвольное множество X. Найти множества:

а) $X\cap X'$; б) $X\cup X'$; в) $X\backslash X'$.

6. Какие из следующих утверждений справедливы:

а) $0\in\varnothing$; б) $\varnothing=\{0\}$; в) $|\{\varnothing\}|=1$;

г) $\{\{\varnothing\}\}\in\{\{\{\varnothing\}\}\}$; д) $|\{\{\varnothing\}\}|=2$?

Этот вопрос коварный. Хотя это может показаться простым или надуманным, пустое множество и его свойства являются достаточно важными. Если вы не совсем уверены в ответе, проработайте вопрос,

используя аналогию портфеля вместо множества. Таким образом, {{}, {}} - портфель, который содержит два пустых портфеля, и, следовательно, |{{}, {}}|=2 и т.д.

7. Пусть M и N — два конечных компьютера с фиксированными программами. Дальше пусть A — множество значений данных, доступных M и таких, что если $x \in A$ и машина M работает с входным словом x, то M останавливается и выдает результат. Аналогично пусть B — множество значений данных, которые приводят N к остановке и выдаче результата. Если любой элемент A доступен M и N, что мы может сказать об элементах B' ? Объяснить эту ситуацию с помощью символов и объяснить бесполезность этой информации.

8. При определении операции объединения подчеркивалось, что мы использовали включение «или». Как в терминах множеств можно выразить исключающее «или»?

9. Часто в вычислениях будут использоваться арифметические операции для образования новых множеств. Так, если A и B — множества чисел, то

$$A+B=\{x: x=a+b,\ a\in A,\ b\in B\}.$$

Аналогично определяются операции •, -, / между множествами чисел. Найти следующие множества:

а) {1,2}+{1,3};
б) {1,2} □ {1,3};
в) {1,2}• {1,3};
г) {1, 2} I {1,3};
д) {1,2}-{1,3};
е) {1, 2}\{1, 3};
ж) {2,4}/{ 2};
з) {2, 4}\{2};
и) {2, 4}-{2}.

Упражнение 1.3.

1. Начертить диаграмму, которая иллюстрирует построение множеств, рассмотренных в задаче 1 упражнения 1.2.

2. Как можно представить следующие множества, используя диаграммы Венна:

$$\{A, \{A\}\}, \{\{a\}, \{b\}, \{X, Y, Z\},$$

где

$$X = \{x: x = 1 \text{ или } (x-2)\in X\},$$
$$Y = x: x = 3 \text{ или } (x-3)\in Y\},$$
$$Z = \{x: x = 2 \text{ или } (x-2)\in Z?$$

Упражнение 1.4.

1. Доказать, что

$$A \text{ I } (B \text{ I } C) = (A \text{ I } B) \text{ I } C.$$

2. Пусть даны множества A, B и C: $C \subseteq B$. Доказать, что:

а) $A \square C \subseteq A \cap B$; б) $A \cup C \subseteq A \cup B$; в) $A \backslash B \subseteq A \backslash C$;

г) $C \backslash A \subseteq B \backslash A$; д) $B' \backslash A \subseteq C' \backslash A$.

3. Показать справедливость равенства

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

4. Доказать, что $(A \cup B)' = A' \cap B'$. (Указание: показать, что $(A \cup B) \cup (A' \cap B') = U$ и $(A \cup B) \cap A' \cap B' = \varnothing$.)

5. Доказать эквивалентность следующих утверждений, т.е. что из каждого следует другое:

а) $A \cup B = U$; б) $A' \subseteq B$; в) $A' \cap B' = \varnothing$.

6. Какие из следующих утверждений справедливы:

а) $0 \in \varnothing$; б) $\{\varnothing\} \subseteq \varnothing$; в) $\varnothing \subseteq \{\varnothing\}$;

г) $\varnothing \subseteq U$; д) $\{\varnothing\} \subseteq \{\{\varnothing\}\}$?

Сравните ответы на этот вопрос с ответами к упражнению 1.2,6. Существует связь между символами $\in$ и $\subseteq$, однако это не одно и то же. Как аналогия «портфеля» связана с символом $\subseteq$?

7. Показать, что для конечного множества A

$$|2^A| = 2^{|A|}.$$

(Указание: выписать множество $A=\{a_1, \ldots, a_n\}$ и рассмотреть его подмножества.)

Упражнение 1.5.

1. Пусть $X=\{a, b, c\}$ и $Y = \{a, b, e, f\}$. Найти $X \times Y$ и Y^2.

2. Доказать: при $A \subseteq X$ и $B \subseteq Y$, $A \times B \subseteq X \times Y$.

3. Доказать, что $A \times (B \cap C) = (A \times B) \cap (A \times C)$.

4. Доказать, что для любых непустых конечных множеств A и B выполняются соотношения:

а) $\varnothing \times A = \varnothing$; б) $U \times A \neq A$; в) $A \subseteq A \times A$;

г) $|A \times \{x\}| = |A|$;

д) $A \times B = B \times A$ тогда и только тогда, когда $A = B$.

2. Отображения и функции

Возвратимся к функциональному соответствию (т.е. к функции). Если это соответствие и вдобавок еще и всюду-определенное, то оно называется ОТОБРАЖЕНИЕМ.

Если отобразить множество студентов в группе, множество фамилий в группе, то это скорее всего будет ОТОБРАЖЕНИЕ множества студентов НА множество фамилий. Т.е. сюръективное соответствие. Если же отобразить множество студентов группы на множество фамилий студентов университета, то говорят, что имеет место ОТОБРАЖЕНИЕ множества студентов В множество фамилий. Т.е., в области значений будут и "незадействованные фамилии".

Мы подошли до одного из самых фундаментальных, может поэтому и неблагозвучных, понятий в теории множеств, и математики вообще, мы подошли к ГОМОМОРФИЗМУ.

Пример. Отобразим множество точек участка земной поверхности на множество точек карты. Сейчас оставим в стороне то, что какое-то множество точек земной поверхности отобразится в одну точку на карте, в таких случаях неинъективность - обычное дело. Для нас существенным образом важно то, что чем выше точки земной поверхности над уровнем моря, тем в более коричневые точки карты они отображаются.

Таким образом, мы рассматриваем не просто множества элементов. В первом случае здесь между элементами множества существует отношение "выше", а во втором - " более коричневые ". Где выше в первому - там более коричневые во втором. "Выше" и " более коричневые " - это отношения, которые заданы на своих множествах.

Отображение земной поверхности НА карту не просто ставит всем элементам одного множества элементы другого. Но, кроме того, если между двумя элементами первого множества существует отношение "выше", то между их образами во втором множестве имеет место отношения " более коричневые ". Очевидно, если точки земной поверхности лежат на одной высоте, то они отобразятся в точки карты с одинаковой коричневостью. Такое отображение называется ГОМОМОРФНЫМ. Или говорят, что между этими множествами существует ГОМОМОРФИЗМ.

Обратим внимание на то, что слово это не очень благозвучное, а по американским меркам и громоздкое. Поэтому по обыкновению используется более короткий (усеченный) термин - МОРФИЗМ.

Морфизмы играют в математике исключительную роль. Так как мамематику часто отождествляют с математическим моделированием, то приведем афоризм из одной умной философской книжки: КРАСИВАЯ МОДЕЛЬ ВСЕГДА ГОМОМОРФНА.

2.1. Формальное определение отображения и его свойств

Пусть X и Y — некоторые множества и $\Gamma \subseteq Y \times X$, причем $\Pr_1\Gamma = X$.

Тройка множеств (X, Y, Γ) определяет некоторое соответствие, которое обладает, однако, тем свойством, что его область определения $\Pr_1\Gamma$ совпадает с областью значений, т.е. X, и, следовательно, это соответствие определено всюду на X. Другими словами, для каждого $x \in X$ существует $y \in Y$, так что $(x,y) \in \Gamma$. Такое всюду определенное соответствие называется *отображением* X в Y, и записывается как

$$\Gamma : X \to Y \tag{45}$$

Под словом «отображение» часто понимают однозначное отображение. Однако мы не будем придерживаться этого правила и будем считать, что каждому элементу $x \in X$ отображение Γ ставит в соответствие некоторое подмножество

$$\Gamma x \subseteq Y, \tag{46}$$

которое называют образом элемента x. Закон, в соответствии с которым осуществляется соответствие, определяется множеством Γ.

Пример 18. Если в примере 14 исключить из рассмотрения шофера c, то получим отображение Γ: $X \to Y$, в котором $X=\{a, b\}$ множество шоферов; $Y=\{\alpha, \beta, \gamma\}$ - множество автомашин;

$\Gamma=\{(a, \alpha), (a, \gamma), (b, \alpha)\}$ - распределение шоферов по автомашинам. Геометрическое представление этого отображения приведено на рис. 15.

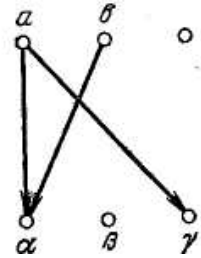


Рис. 15. Геометрическое представление отображения

Обратимся теперь к рассмотрению некоторых свойств отображения. Пусть $A \subseteq X$. Для любого $x \in A$ образом x будет множество $\Gamma x \subseteq Y$. Совокупность всех элементов Y, которые являются образами Γx для всех $x \in A$, назовем образом множества A и будем обозначать ΓA. Согласно этому определению

$$\Gamma A = \bigcup_{x \in A} \Gamma x \cdot \tag{47}$$

Если A_1 и A_2 — подмножества X, то

$$\Gamma(A_1 \cup A_2) = \Gamma A_1 \cup \Gamma A_2 . \tag{48}$$

Однако соотношение

$$\Gamma(A_1 \cap A_2) = \Gamma A_1 \cap \Gamma A_2 \tag{49}$$

справедливо только в том случае, если отображение является однозначным. В общем же случае

$$\Gamma(A_1 \cap A_2) \subseteq \Gamma A_1 \cap \Gamma A_2 . \tag{50}$$

Полученные соотношения легко обобщаются и на большее число подмножеств A_i. Так, если $A_1.., A_n$ — подмножества X, то

$$\Gamma\left(\bigcup_{i=1}^{n} A_i\right) = \bigcup_{i=1}^{n} \Gamma A_i ; \tag{51}$$

$$\Gamma\left(\bigcup_{i=1}^{n} A_i\right) \subseteq \bigcup_{i=1}^{n} \Gamma A_i \cdot \tag{52}$$

Поскольку **отображение является частным случаем соответствия,** для отображения имеют место введенные при рассмотрении соответствий понятия обратного отображения и композиции отображений.

Типы отображений

При *отображении* X в Y каждый элемент x из X имеет один и только один образ $y=\Gamma(x)$ из Y. Однако совсем не обязательно, чтобы и всякий элемент из Y был образом некоторого элемента из X (рис. 16, *а).* Если же любой элемент из Y есть образ, по крайней мере, одного элемента из X (рис. 16, б), то говорят, что имеет место *отображение X на Y (сюръекция* или *накрытие).*

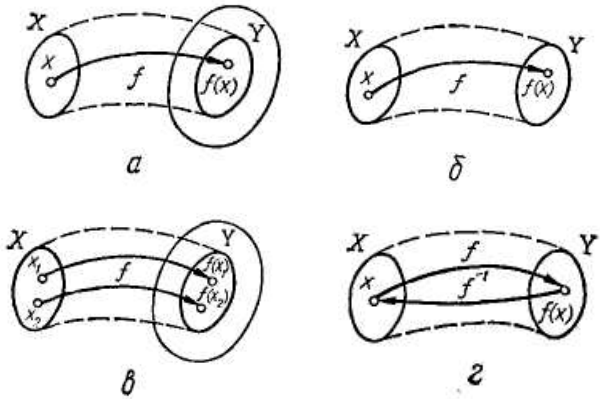


Рис. 16. Типы отображений:
а — отображение X в Y;
б — отображение X на Y (сюръекция);
в — взаимно-однозначное отображение X в Y (инъекция);
г — взаимно-однозначное отображение X на Y (биекция).

Если для любых двух разных элементов x_1 и x_2 из X их образы $y_1=\Gamma(x_1)$ и $y_2=\Gamma(x_2)$ также разные, то отображение называется *инъекцией* (рис. 16,*в*). Отображение, которое является одновременно сюръективным и инъективным (рис. 16, *г*), называется *биекцией (наложением)*. В этом случае говорят, что $\Gamma: X \to Y$ *есть взаимно-однозначное* отображение, а между элементами X и Y есть взаимно-однозначное соответствие. При этом, обратное отношение Γ^{-1} также взаимно-однозначное отображение, $x=\Gamma^{-1}(y)$ равносильно $y=\Gamma(x)$ и $(\Gamma^{-1})^{-1}$ совпадает с Γ.

Любое отображение Γ из X в Y есть элемент множества $U(X\times Y)$, которое обозначается также через Y^X (напомним, что $U(X\times Y)$ — это множество всех подмножеств прямого произведения $X\times Y$, а элементами последнего являются упорядоченные пары (x, y), где $x \in X$ и $y \in Y$). Если Γ — взаимно-однозначное отображение, а множества X и Y совпадают ($X=Y$), то $\Gamma: X \to X$ называют *отображением множества X на себя.* Элементы $(x, x) \in X\times X$ образовывают *тождественное отображение e,* причем

$$\Gamma\Gamma^{-1} = \Gamma^{-1}\Gamma = e.$$

Отображения, заданные на одном множестве

Важным частным случаем отображения является случай, когда множества X и Y совпадают. При этом отображение Γ: $X \to X$ будет представлять собой отображение множества X самого в себя и будет определяться парой

$$(X, \Gamma), \tag{53}$$

где $\Gamma \subseteq X^2$.

Подробным изучением таких отображений занимается теория графов, которая будет рассматриваться в других разделах дискретной математики. Затронем здесь лишь некоторые операции над подобными отображениями.

Пусть Γ и Δ — отображение множества X в X. Композицией этих отображений назовем отображение $\Gamma\Delta$, которое в соответствии с правилом, приведенным в п. 1.4 определяется следующим образом:

$$(\Gamma\Delta)x = \Gamma(\Delta x). \tag{54}$$

В частном случае, если $\Delta=\Gamma$, получаем отображение

$$\Gamma^2 x = \Gamma(\Gamma x); \tag{55}$$

$$\Gamma^3 x = \Gamma(\Gamma^2 x) \text{ и т.д.} \tag{56}$$

Таким образом, в общем случае для любого $s \geq 2$

$$\Gamma^s x = \Gamma(\Gamma^{s-1} x). \tag{1.57}$$

Специальным определением введем соотношение

$$\Gamma^0 x = x. \tag{58}$$

Это дает возможность распространить соотношение (57) и на отрицательные s. Действительно, согласно (57)

$$\Gamma^0 x = \Gamma(\Gamma^{-1} x) = \Gamma\Gamma^{-1} x = x. \tag{59}$$

Это означает, что $\Gamma^{-1} x$ представляет собой обратное отображение. Тогда

$$\Gamma^{-2} x = \Gamma^{-1}(\Gamma^{-1} x) \tag{60}$$

и т.д.

Пример 19. Пусть X — множество людей. Для каждого человека $x \in X$ обозначим через Γx множество его детей. Тогда $\Gamma^2 x$— множество внуков x; $\Gamma^3 x$— множество правнуков x; $\Gamma^{-1} x$— множество родителей x и т. д.

Изображая людей точками и рисуя стрелки, которые идут из x в Γx, получаем родословное или гениалогическое дерево (рис. 17).

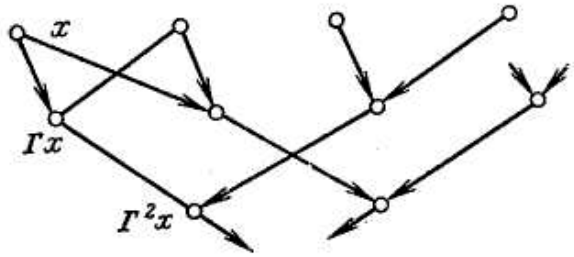


Рис. 17. Гениалогическое дерево

Пример 20. Рассмотрим шахматную игру. Обозначим через x некоторое положение (расположение фигур на доске), которое может создаваться в процессе игры, а через X множество всевозможных положений. Тогда Γx для любого $x \in X$ будет означать множество положений, которые можно получить из x, делая один ход при соблюдении правил игры. При этом $\Gamma x=\varnothing$, если x матовое или патовое положение; $\Gamma^3 x$ — множество положений, которые можно получить из x тремя ходами; $\Gamma^{-1} x$ — множество положений, из которых данное положение может быть получено за один ход.

Для отображений, заданных на одном множестве часто используют некоторые другие названия, которые у нас встретятся в дальнейшем.

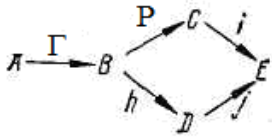
Так, если элементы $x \in X$ представляют собой состояние динамической системы, то отображение Γx может рассматриваться как множество состояний, в которые система может перейти из данного состояния. В этом случае удобно использовать термин *преобразования состояния динамической системы*. Для обозначения некоторых специальных видов отображений, заданных на одном том же множестве, используется также термин *отношение.*

Композиция отображений.

Если $\Gamma: X \to Y$ и $P: Y \to Z$, то их композиция $(P \circ \Gamma): X \to Z$, причем $(P \circ \Gamma)(x) = P(\Gamma(x))$. Пусть, например, $\Gamma=\sin$, $P=\ln$; тогда

$$(P \circ \Gamma)(x) = (\ln \circ \sin)x = \ln \sin x.$$

Для наглядности представления соотношений, где встречается несколько отображений, пользуются диаграммами, например:



Такая диаграмма называется *коммутативной,* если в любом случае, когда можно пройти от одного множества к другому по различным последовательностям стрелок, соответствующие композиции совпадают (в приведенном выше примере условие коммутативности $i \circ P = j \circ h$).

Подстановки как отображение.

Взаимно-однозначное отображение множества $N=\{1, 2, ..., n\}$ на себя называется *подстановкой n* чисел (или подстановкой *n*-й степени). Обычно принято записывать подстановку двумя строками, заключенными в скобки. Первая строка содержит аргументы (первые координаты) подстановки, а вторая - соответствующие им образы (вторые координаты). Например, взаимно-однозначное соответствие четырех чисел, заданное множеством упорядоченных пар
{(1, 2), (2, 4), (3, 3), (4, 1)} запишется как подстановка *a* четвертой степени

$$a = \begin{pmatrix} 1\,2\,3\,4 \\ 2\,4\,3\,1 \end{pmatrix}$$

в которой 1 переходит в 2, 2 — в 4, 3 — в 3 и 4 — в 1.

Так как безразлично, в каком порядке идут упорядоченные пары отображения, то одна и та же подстановка допускает различные представления:

$$\begin{pmatrix} 1\,2\,3\,4 \\ 2\,4\,3\,1 \end{pmatrix} = \begin{pmatrix} 4\,2\,3\,1 \\ 1\,4\,3\,2 \end{pmatrix} = \begin{pmatrix} 2\,1\,4\,3 \\ 4\,2\,1\,3 \end{pmatrix} \text{ и т.д.}$$

Каждая строка в записи подстановки *n*-й степени содержит *n* различных чисел, расположенных в определенном порядке, т.е. представляет собой некоторую *перестановку n* чисел 1, 2, ..., *n*. Если обозначить *i*-е элементы перестановок через α_i и β_i ($i = 1, 2,..., n$), причем $\alpha_i, \beta_i \in N$, то подстановку *n*-й степени можно представить как

$$a = \begin{pmatrix} \alpha_1, \alpha_2, ..., \alpha_n \\ \beta_1, \beta_2, ..., \beta_n \end{pmatrix}.$$

Поскольку число всех перестановок из n чисел равно $n!$, то число всех различных подстановок n-й степени, как и число всевозможных способов записи каждой из таких подстановок, также равно $n!$

Тождественная подстановка n-й степени e_n переводит каждое число в себя. Очевидно, одной из записей e_n является следующая:

$$e_n = \begin{pmatrix} 1\,2\,\ldots\,n \\ 1\,2\,\ldots\,n \end{pmatrix}.$$

Если в подстановке a поменяем местами ее перестановки, то получим подстановку a^{-1}, *симметричную* a. Например

$$a = \begin{pmatrix} 1\,2\,3\,4 \\ 2\,4\,3\,1 \end{pmatrix}; \qquad a^{-1} = \begin{pmatrix} 2\,4\,3\,1 \\ 1\,2\,3\,4 \end{pmatrix}$$

Композицией подстановок n-й степени a и b называется подстановка n-й степени $c = ab$, являющаяся результатом последовательного выполнение сначала a, потом b. Например:

$$c = ab = \begin{pmatrix} 1\,2\,3\,4 \\ 2\,4\,3\,1 \end{pmatrix} \begin{pmatrix} 1\,2\,3\,4 \\ 1\,4\,3\,2 \end{pmatrix} = \begin{pmatrix} 1\,2\,3\,4 \\ 4\,2\,3\,1 \end{pmatrix}$$

так как 1 переходит в 2 и 2 — в 4, т.е. в результате 1 переходит в 4 и т.д.

Очевидно, если a — подстановка n-й степени, то

$$ae_n = e_n a = a, \quad aa^{-1} = a^{-1}a = e_n.$$

Подстановка называется *четной*, если общее число инверсий в ее строках (перестановках) четно, и *нечетной* — в противном случае. Как известно, *инверсию* образуют два числа в перестановке, когда меньшее из них расположено правее от большего. Каждой перестановке можно сопоставить число инверсий в ней, которое подсчитывается следующим образом: для каждого из чисел определяется количество стоящих правее его меньших чисел, и полученные результаты складываются. Например, подстановка

$$\begin{pmatrix} 4\,2\,5\,1\,3\,6 \\ 5\,3\,1\,4\,2\,6 \end{pmatrix}$$

нечетная, так как количество инверсий в верхней перестановке

$$3+1+2+0+0+0=6$$

и в нижней перестановке

$$4+2++0+1+0+0=7,$$

т.е. общее число инверсий $6+7=13$.

Разложение подстановки в циклы.

Всякую подстановку можно разложить в *произведение циклов,* множество элементов которых попарно не пересекаются. *Цикл* — это такая подстановка

$$\begin{pmatrix} \alpha_1, \alpha_2,..., \alpha_{k-1}, & \alpha_k, \alpha_{k+1},..., \alpha_n \\ \alpha_2, \alpha_3,..., \alpha_k, & \alpha_1, \alpha_{k+1}, ..., \alpha_n \end{pmatrix} = (\alpha_1, \alpha_2,..., \alpha_k)$$

которая переводит α_1 в α_2, α_2 в α_3, ..., α_{k-1} в α_k и α_k в α_1, а другие элементы α_{k+}, ..., α_n переходят в самих себя.

Сокращенная запись цикла $(\alpha_1, \alpha_2, ..., \alpha_k)$ сводится к перечислению множества элементов, которые циклически переходят друг в друга, а количество этих элементов k определяет *длину (порядок) цикла.* Так,

$$\begin{pmatrix} 4\,2\,5\,1\,3\,6 \\ 5\,3\,1\,4\,2\,6 \end{pmatrix} = (1, 4, 5)(2, 3)(6).$$

Цикл длины 1 представляет собой тождественную подстановку и часто не записывается. Подстановка, все n элементов которой образуют цикл, называется *круговой* или *циклической.* Цикл длины 2 называют *транспозицией* (это подстановка, которая переставляет только два элемента). Всякая подстановка представляется произведением транспозиций, например:

$$\begin{pmatrix} 1\,2\,3\,4\,5 \\ 2\,5\,4\,1\,3 \end{pmatrix} = \begin{pmatrix} 1\,2\,3\,4\,5 \\ 2\,1\,3\,4\,5 \end{pmatrix} \cdot \begin{pmatrix} 2\,1\,3\,4\,5 \\ 2\,5\,3\,4\,1 \end{pmatrix} \cdot \begin{pmatrix} 2\,5\,3\,4\,1 \\ 2\,5\,4\,3\,1 \end{pmatrix} \cdot \begin{pmatrix} 2\,5\,4\,3\,1 \\ 2\,5\,4\,1\,3 \end{pmatrix} =$$

$=(1, 2)(1, 5)(3,4)(1,3).$

Заметим, что подобное разложение может содержать циклы с общими элементами и при этом оно не является единственным. В то же время разложеник подстановки на *независимые циклы* (без общих элементов) всегда можно осуществить только единственным способом.

Разность между числом всех элементов подстановки n и количеством ее циклов m (с учетом циклов длины 1) называется *декрементом* подстановки $d= n-m$. Четность подстановки совпадает с парностью ее декремента.

2.2. Функция

Рассмотрим некоторое отображение

$$f: X \to Y \qquad (61)$$

Это отображение называется *функцией,* если оно является однозначным, т.е. если для любых пар $(x_1, y_1) \in f$ и $(x_2, y_2) \in f$ из $x_2 = x_1$ следует $y_2 = y_1$.

Из определения отображения и из приведенных ранее примеров следует, что элементами множества X и Y могут быть объекты любой природы. Однако в задачах вычислительных систем особый интерес представляют отображения, которые являются однозначными и множество значений которых представляет собой множество вещественных чисел R. Однозначное отображение f, которое определяется (61) называется функцией с вещественными значениями, если $Y \subseteq R$.

Понятие функция является чрезвычайно широким и изучению отдельных классов функций посвящены многие математические дисциплины (алгебра, тригонометрия и т.п.). Мы рассмотрим только некоторые общие наиболее фундаментальные свойства функции, не касаясь свойств конкретных классов функций.

Пример 21. Из данного города в другой можно проехать по железной дороге, автобусом или самолетом. Стоимость билета будет соответственно 70, 90 и 120 грн. Стоимость билета в этом примере можно представить как функцию от вида транспорта. Для этого рассмотрим множества

$$X = \{\text{ж.д., авт., сам.}\}; \quad Y = \{70, 90, 120\}.$$

Функция $f : X \to Y$, получаемая из условий примера, может быть записана в виде множества

$$f = \{(\text{ж.д.}, 70), \ (\text{авт.}, 90), \ (\text{сам.}, \ 120)\}.$$

Значение y в любой из пар $(x, y) \in f$ называется функцией от данного x, которая записывается в виде $y = f(x)$.

Такая запись позволяет ввести следующее формальное определение функции:

$$f = \{(x, y) \in X \times Y \mid y = f(x)\}. \qquad (62)$$

Таким образом, символ f используется при определении функции в двух смыслах:

1) f является множеством, элементами которого являются пары (x, y), которые принимают участие в соответствии;

2) $f(x)$ является обозначением для $y \in Y$, соответсвующего данному $x \in X$.

Формальное определение функции в виде соотношения (62) позволяет установить способы задания функции.

1. Перечисление всех пар (x, y), составляющих множество f, как это было сделано в примере 21. Такой способ задания функции

применим, если X является конечным множеством. Для большей наглядности пары (x, y) удобно располагать в виде таблицы.

2. Во многих случаях как X, так и Y представляют собой множества вещественных или комплексных чисел. В таких случаях очень часто под $f(x)$ понимают формулу, т.е. выражение, которое содержит перечень математических операций (сложение, вычитание, деление, логарифмирование и т.п.), которые нужно выполнить над $x \in X$, чтобы получить y.

Пример 22. Пусть

$$X = Y = R \quad и \quad f = \{(x, y) \in R^2 \mid y = x^2\}.$$

Тогда

$$f(x) = x^2$$

Иногда для разных подмножеств множества X функции приходится пользоваться различными формулами. Пусть $A_1, \ldots, A_n$ — попарно непересекающиеся подмножества X. Обозначим через $f_i(x)$, $i=1, 2, \ldots, n$ формулу, которая определяет y при $x \in A_i$. Тогда функция $f(x)$ будет определяться выражением

$$f(x) = \begin{cases} f_1(x) \text{ при } x \in A_1; \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ f_n(x) \text{ при } x \in A_n. \end{cases} \qquad (.63)$$

Так, функцию $y=f(x) = |x|$ можно задать в виде

$$y = \begin{cases} x \text{ при } x \geq 0; \\ -x \text{ при } x < 0. \end{cases}$$

3. Если X и Y— множества вещественных чисел, то элементы $(x, y) \in f$ можно изобразить в виде точек на плоскости R^2. Полная совокупность таких точек будет представлять собой *график* функции $f(x)$.

Если в выражении (61) $X=U\times V$, то приходим к функции от двух переменных u и v, обозначаемой через $f(u, v)$, где $u \in U$ и $v \in V$. Формальное определение функции двух вещественных переменных будет следующим :

$$f = \{(u, v, y) \in U \times V \times Y \mid y = f(u, v)\}. \qquad (64)$$

Рассмотрим определение функции с применением декартова произведения множеств.

Напомним, что декартовым произведением $M_a \times M_b$ множеств M_a и M_b называется множество M вида

$$M = \{(m_i, m_j) / m_i \in M_a, m_j \in M_b\}.$$

Подмножество $F \in M_x \times M_y$, называется *функцией*, если для каждого элемента x, $x \in M_x$, имеется не более одного элемента y, $y \in M_y$ вида $(x, y) \in F$; при этом если для каждого элемента x существует один элемент y вида $(x, y) \in F$, то функция называется *всюду (полностью) определенной,* в противном случае - *частично определенной (недоопределенной).* Множество M_x образует *область определения функции F,* множество M_y *— область значений функции F.* Часто вместо записи $(x, y) \in F$ используют запись $y=F(x)$; при этом элемент x называют *аргументом* или *независимой переменной,* а y *— значением функции F, или зависимой переменной.*

Сопоставим с декартовым произведением двух множеств прямоугольную решетку, узлы которой взаимно однозначно соответствуют элементам декартова произведения. Подмножество декартова произведения на рисунках будем отмечать штриховкой соответствующих элементов.

Пример 23. На рис. 18,*а* изображено подмножество декартова произведения множеств $M_x=\{x_1\ x_2,\ x_3,\ x_4\}$ и $M_y=\{y_1,\ y_2,\ y_3\}$, не являющееся функцией; на рис. 18,*б*, - являющееся полностью определенной функцией; на рис. 18, *в* — частично определенной функцией.

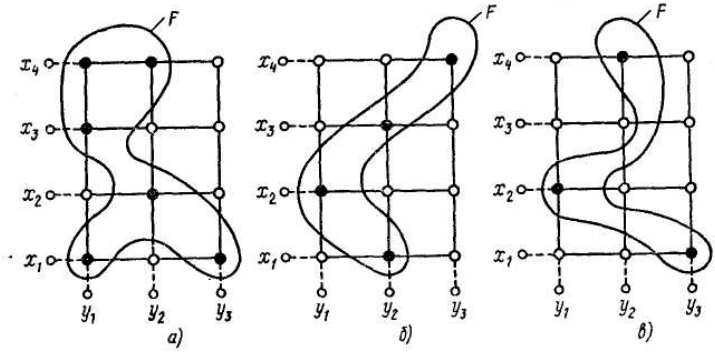


Рис. 18.

Количество аргументов определяет *местность функции.* Выше были рассмотрены одноместные функции.

Аналогично понятию декартова произведения двух множеств определим декартово произведение n множеств.

Декартовым произведением

$$M_1 \times M_2 \times ... \times M_n = \prod_{i=2}^{n} M_i$$

множеств M_1, M_2,…,M_n называется множество

$$M = \{(m_{i_1}, m_{i_2}, ..., m_{i_n}) / m_{i_1} \in M_1, m_{i_2} \in M_2, ..., m_n \in M_{i_n}\}$$

Элементами декартова произведения $M_1 \times M_2 \times ... \times M_n$ являются всевозможные последовательности, каждая из которых состоит из n элементов, причем первый элемент принадлежит множеству M_1, второй - множеству M_2,..., n-й элемент — множеству M_n.

Если множество M_x в определении функции $y=F(x)$ является декартовым произведением множеств

$$M_{x_1} \times M_{x_2} \times ... \times M_{x_n},$$

то получаем определение n-местной *функции*

$$y = F(x_1, x_2, ..., x_n).$$

Частным случаем n-местной функции $y=F(x_1,x_2,...,x_n)$ является n-местная операция. Под n-местной *операцией* O_n в множестве M понимается n-местная функция $y=F(x_1,x_2,...,x_n)$, в которой область определения аргументов и область значений функции совпадают:

$$M_{x_1} = M_{x_2} = ... = M_{x_n} = M_y.$$

Таким образом, n-местная операция по n элементам множества M определяет $(n+1)$-й элемент этого же множества. **Сужение и продолжение функции.** Пусть функция $f: X \to Y$ определена на множестве X, а f_1 — на множестве $Q \subset X$, причем для каждого $x \in Q$ значение функций f и f_1 совпадают. Тогда f_1 называют *ограничением (сужением) функции* f на Q, а f — *продолжением функции* f_1 *на* X.

Например, функция $f(x)=x^3$ (другая запись $x \to x^3$), определенная на множестве действительных чисел R, отображает это множество на

себя. Если ограничить область определения этой функции множеством целых чисел Z, то получим сужение $f_1(x)$ функции $f(x)$ на Z, причем $f_1(x)$ отображает множество Z в Z (а не на Z), так как не всякое число является кубом целого числа. Операцию сужения функции часто используют для табличной задачи функций с бесконечной областью определения X. В качестве множества A берут обычно выборку равнозначных значений x множества X. Получаемое при этом сужение f_A функцию f уже легко представить в виде таблицы. По этому принципу построены таблицы логарифмов, тригонометрических функций и другие. Функции f и g равны, если их область определения – то же самое множество A и для любого $a \in A$ $f(a)= g(a)$.

Пример 24. 1) Функция $f(x)=2^x$ является отображением N в N и N на M_{2^n}

2) Всякая нумерация счетного множества есть его отображением на N.

3) Функция $f(x)=\sqrt{x}$ не полностью определена, если ее тип $N \to N$, и полностью определена, если ее тип $N \to R$ или $R_+ \to R$ (R_+ положительное подмножество R).

4) Пусть зафиксирован список $\{a_1, ..., a_n\}$ всех элементов конечного множества A. Тогда любой вектор $v_i = (a_{i_1},...,a_{i_n})$ из A^n можно рассматривать как описание функции f_i: $A \to A$ (т.е. преобразование A), определяемой следующим образом: $f_i(a_j)=a_{ij}$, т.е. значение f_i для a_j равно j-й компоненте v_i. Число всех преобразований A равно, следовательно, $|A^n| = n^n$. Аналогично всякую функцию типа $N \to N$ можно представить бесконечной последовательностью элементов N, т.е. натуральных чисел; отсюда нетрудно показать, что множество всех преобразований счетного множества континуально.

5) Каждое натуральное число n единственным образом разлагается на произведение простых чисел (простых делителей этого числа). Поэтому, если договориться располагать простые делители n в определенном порядке (например, в порядке неубывания), то получим функцию $q(n)$ типа

$$N \to \bigsqcup_{i=1}^{\infty} N^i ,$$

которая отображает N в множество векторов произвольной длины. Например,

$$q(42)=(2, 3, 7),\ q(23)=23,\ q(100)=(2, 2, 5, 5).$$

Это отображение не является сюръективным, так как в область значений q не входят векторы, для компонентов которых не выполнено условие неубывания.

6) Каждому человеку соответствует множество его знакомых. Если зафиксировать момент времени (например, 10 января 2010 г., 5 ч. 00 мин), то это соответствие будет однозначным и является отображением множества M людей, которые живут в этот момент, в множество подмножеств M.

Пример 25. Функция $\sin x$ имеет тип $R \to R$. Отрезок $[-\pi/2, \pi/2]$ она взаимно-однозначно отображает на отрезок $[-1, 1]$. Поэтому на отрезке $[-1, 1]$ для нее есть обратная функция $\arcsin x$.

Пример 26. 1) Функции $\sin x$ и $\sqrt{x}$ имеют тип $R \to R$, т.е. отображают одно и то же множество в себя. Поэтому их композиция возможна в произвольном порядке и дает функции $\sin\sqrt{x}$ и $\sqrt{\sin x}$. Заметим, что области определения их различны: первая функция определена на положительной полуоси; вторая функция определена на множестве отрезков $[2k\pi, (2k+1)\pi]$, где $k = 0, \pm1, \pm2$ ... Таким образом, область определения композиции может быть уже области определения обеих исходных функций и даже быть пустой.

2) Множество $K = \{k_1 ..., k_m\}$ команд ЭВМ отображается в машинные коды этой ЭВМ, т.е. в натуральные числа. Кодирующая функция φ имеет тип $K \to N$. С помощью суперпозиции этой функции и арифметических функций оказываются возможными арифметические действия над командами (которые сами по себе числами не является), т.е. функции вида $\varphi(k_1)+\varphi(k_2)$, $\varphi(k_1)+4$ и т.д.

3) В функции $f_1(x_1, x_2, x_3) = x_1+2x_2+7x_3$ переименование x_3 в x_2, приводит к функции $f_1(x_1, x_2, x_2)=x_1+2x_2+7x_2$, что равно функции двух аргументов $f_2(x_1, x_2)=x_1+9x_2$. Переименование x_1 и x_3 в x_2 приводит к одноместной функции $f_3(x_2)=10x_2$.

4) Элементарной функцией в математическом анализе называется всякая функция f, которая является суперпозицией фиксированного (т.е. не зависимого от значений аргументов f) числа арифметических функций, а также функций e^x, $\log x$, $\sin x$, $\arcsin x$. Например, функция $\log^2(x_1+x_2)+3\sin\sin x_1+x_3$ элементарна, так как является результатом нескольких последовательных суперпозиций x_1+x_2, x^2, $\log x$, $3x$, $\sin x$.

5) Всякая непрерывная функция n переменных представима в виде суперпозиции непрерывных функций двух переменных.

Числовые функции. Проиллюстрируем введенные понятия на функциях, определенных на числовых множествах, элементами которых являются действительные числа. Такая функция каждому числу x из области определения ставит в соответствие число $y=f(x)$ из области ее значений. Иначе говоря, числовая функция f определяется множеством упорядоченных пар чисел (x, y).

Говоря геометрическим языком, множеству действительных чисел отвечает множество *точек прямой (числовой оси).* Пары чисел (x, y) представляются в декартовой системе координат *точками плоскости* с координатами $x \in X$ и $y \in Y$, причем первая координата x — *абсцисса,* а вторая y — *ордината* точки. Числовые оси, которые отвечают множествам X и Y, есть *осями координат,* а декарто произведение $X \times Y$, представляет собой множество точек плоскости. Таким образом, между элементами множества $X \times Y$ и точками плоскости устанавливается взаимно-однозначное соответствие.

Различные подмножества действительных чисел, на которых определяется функция, отвечают подмножествам точек прямой. В качестве таких подмножеств часто используют следующие:

отрезок (замкнутый интервал) $[a, b] = \{x \mid a \le x \le b\}$;
полуинтервал, открытый слева $(a, b] = \{x \mid a < x \le b\}$;
полуинтервал, открытый справа $[a, b) = \{x \mid a \le x < b\}$;
открытый интервал (или просто *интервал)* $(a, b) = \{x \mid a < x < b\}$.

Область определения функции может быть задана и отдельными точками числовой прямой. Множество точек плоскости, которая отвечает множеству упорядоченных пар $(x,y) \in f$, называется графиком функции f. На рис. 19 изображен график функции $y=f(x)$, определенной на множестве G с областью значений F.

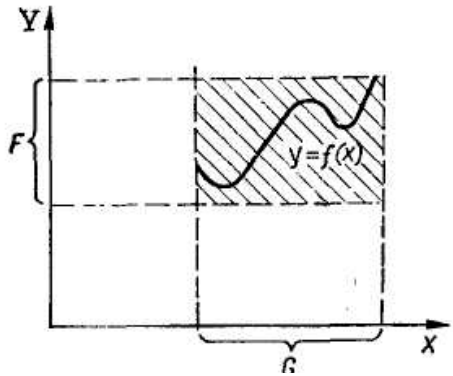


Рис. 19. График числовой функции $y=f(x)$ (G — область определения; F — область значений).

В заключение отметим, что при более строгом рассмотрении между отображением и функцией все же имеется некоторое различие, характеризуемое способом определения этих отношений на множестве X, причем отображение следует рассматривать как частный случай функции. Функциональное отношение $A \subset X \times Y$ называют отношением множества X в Y, если это отношение всюду определено на X, т. е. его область определения $D_0(A)$ совпадает с множеством X.

Отношение $A \subset X \times Y$ называют функциональным, если все его элементы (упорядоченные пары) имеют различные первые координаты, т. е. каждому элементу $x \in X$, такому, что $(x, y) \in A$, соответствует один и только один элемент $y \in Y$. При этом первая координата x упорядоченной пары $(x, y) \in A$ является аргументом (переменной), а вторая y —образом (значением) функции.

Пример. Во множестве $N = \{1, 2, 3, 4, 5, 6\}$ заданы отношения:

$$\{(1, 3),\ (2,\ 4),\ (2, 6), (3, 5), (3,\ 2)\}, \qquad \text{(а)}$$

$$\{(1, 6),\ (2, 2),\ (3, 5), (4, 5), (5, 6)\}. \qquad \text{(б)}$$

Какие из этих отношений являются функциями и какие отображениями?

Решение. В выражениях (а) и (б) первое отношение является отображением, второе — функцией, так как для второго отношения все первые координаты отличны друг от друга, а для первого это условие не выполняется.

Рассмотрим пример конструирования печатной платы. Пусть x — некоторое исходное расположение конструктивных элементов на плате; X — множество различных расположений таких элементов на плате. Тогда Γx для любого $x \subset X$ —множество положений, которые можно получить из x, например с помощью парных перестановок конструктивных элементов, делая один шаг перестановок в направлении улучшения некоторого показателя качества размещения. При этом $\Gamma^4 x$—множество перестановок конструктивных элементов, которые можно выполнить из состояния x четырьмя шагами; $\Gamma^{-1} x$ — множество положений (состояний) конструктивных элементов, из которых данное положение может быть получено за один шаг. Если из положения x перестановками с другими элементами не удается улучшить показатель качества размещения (достичь локальный оптимум показателя качества), то $\Gamma x = \varnothing$.

Обратная функция

Понятие обратной функции может быть применено для такого отображения $f:X\to Y$, которое, во-первых, является однозначным, т.е. для любых $(x_1, y_1)\in f$ и $(x_2, y_2)\in f$ из $x_2=x_1$ следует $y_2=y_1$ и, во-вторых, является взамно-однозначным, т.е. из $x_2\neq x_1$ следует $y_2\neq y_1$. При выполнении этих условий отображение $f:X\to Y$ является однозначным, т.е. определяет функцию $y=f(x)$. Обратное отображение $f^{-1}:Y\to X$ также является однозначным и определяет функцию $y=f^{-1}(x)$, которую называют обратной по отношению к функции $y=f(x)$. При аналитическом задании функции f принято аргумент как прямой, так и обратной функции обозначать одной и той же буквой, например, x. Поэтому для нахождения обратной функции нужно уравнение $y=f(x)$ решить относительно x и поменять обозначения, заменив x на y и y на x. При этом обратная функция запишется в виде $y=f^{-1}(x)$.

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 таким образом: оно действует из A в B с помощью σ, а потом из B в C с помощью 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 $G_{\rho\circ\sigma}=\sigma^{-1}G_\rho$. Чтобы проиллюстрировать ситуацию, рассмотрим рис. 20.

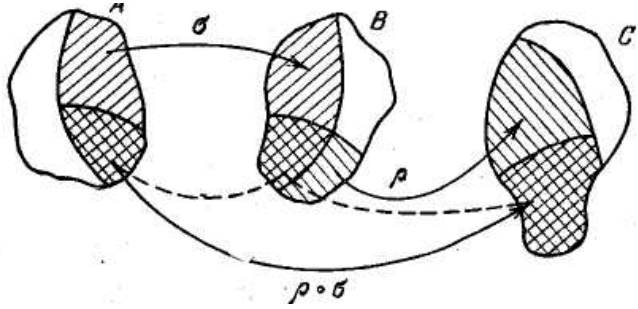


Рис. 20.

Области определения и значений σ и ρ заштрихованы в разных направлениях. Следовательно, сегменты с двойной штриховкой на A, B и C представляют собой $G_{\rho\circ\sigma}$ G_ρ $\square F_\rho$, $F_{\rho\circ\sigma}$ 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 принят в той книге.

Пример 27. 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

$$G_\rho=\{x^2: x\in N\},\quad G_\sigma=\{x: x, x+1\in N=N,$$

$G_{\rho\circ\sigma}=\sigma^{-1}G_\rho=\{x: x\in N$ и $x+1=y^2$, где $y\in N\}=\{3, 8, 15, 24, ...\}$ (рис. 21).

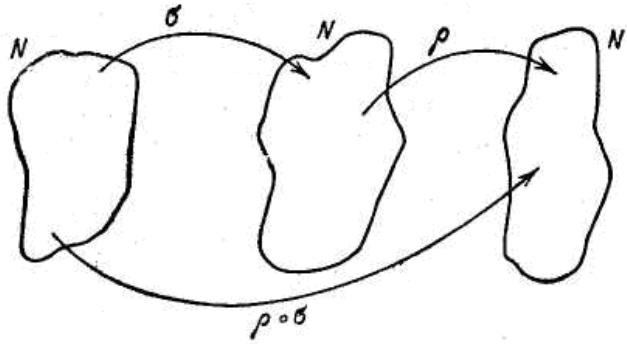


Рис. 21

Используя результаты, полученные выше, выполним исследование сложных функций. Пусть дана функция f: $A\to B$; в этом случае f^{-1} является функцией тогда и только тогда, когда f инъективна, а отображением тогда и только тогда, когда f биективна. В большинстве рассматриваемых нами случаев f — биекция; тогда f^{-1} — также биекция, а функции $f^{-1}\circ f$ и $f\circ f^{-1}$ являются тождественными отображениями.

Рассмотрим функции f: $A\to B$ и g: $B\to C$. Тогда:

а) если f и g инъективны, то существует $g\circ f$;

б) если f и g сюрективны, то также существует $g\circ f$.

Обратным отношением к $g\circ f$ есть $f^{-1}\circ g^{-1}$. Порядок должен быть обратным, как указано на рис. 22.

Заметим, что если g — отображение, т.е. $G_g=B$, то $F_f \subseteq G_g$ и, следовательно, $G_{g\circ f}=F_g$. Аналогично, если $F_f \supseteq G_g$, то $F_{g\circ f}=F_g$. Если f и g инъективны, то существует $g\circ f$; следовательно, $f^{-1}\circ g^{-1}$ — функция.

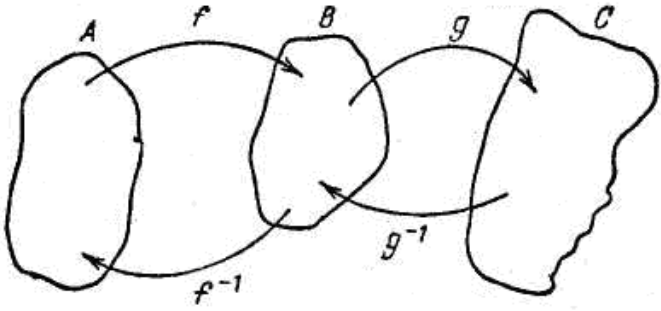


Рис. 22.

Подытоживая вышесказанное, имеем: из $F_f=G_g$ следует, что $g\circ f: G_f \to F_g$ — отображение; если f и g также инъективны, то $f^{-1}\circ g^{-1}: F_f \to G_g$ — биекция. Очевидно, что эти критерии выполняются, если f и g — биекции.

2.3. Формализация мощности множеств и счетность

Мы почти подошли к тому моменту, когда появляется возможность использовать понятие биекции для формализации понятия мощности и процесса вычислений. Вычисление важно не только само по себе, но также и потому, что **функция является *вычислимой* тогда и только тогда, когда связанное с ней множество счетно**. Вначале дадим определение множества N. **Чтобы прояснить наши намерения, заметим, что любое число (1, 2 и т.д.) может быть использовано двумя различными способами: как существительное или как прилагательное, которое дает номер другого существительного. Мы будем рассматривать числа как существительные.**

В качестве предварительного определения N положим

$$N=\{1\} \cup \{n+1: n \in N\}.$$

Из этого рекурсивного определения следует, что $1 \in N$, и если к произвольному элементу из N прибавить 1, то полученный результат также принадлежит N. Следовательно, N содержит 1, 1+1(=2), 2+1(=3), 3+1(=4) и т.д.

К сожалению, это определение неприемлемо по причинам, которые будут рассмотрены ниже. Тем не менее здесь есть несколько важных моментов. Например, так как N (по крайней мере интуитивно) бесконечно, то мы должны иметь механизм, с помощью которого можно конструировать последующие элементы из конечного множества,— другими словами, никогда не сможем написать точное представление N. Мы также должны придумать имя числу, которое называем «один», и аналогично для «два» (сокращение 1+l), «три» и т.д. Конечно, мы могли бы выбрать любые имена или символы для этой цели, однако было бы неправильно использовать неудобные обозначения. Перейдем теперь к недостаткам данного определения. Проверка его обнаруживает, что оно содержит два новых символа: «1» и «+»; другие символы известны из построения множеств. Символ «1» можно объяснить вышеуказанным способом. Однако символ «+» означает операцию *на* N и, следовательно, не может быть использован для определения N. (Операции будут определены ниже.)

Чтобы выйти из этого затруднения, возвратимся к основам теории множеств. Напомним, что нам нужно было построить число, которое на 1 больше максимального из всех предыдущих чисел. Легко получить аналогичный процесс для множеств (который называется построением надмножеств с количеством элементов на 1 больше, чем в данном множестве).

Пример 28. Пусть $A=\{x, y, z\}$ и $B=\{x, y, z, A\}$. Тогда $A \subseteq B$ и $A \in B$, поэтому $B \backslash A = \{A\}$ и имеет только один элемент.

Эта конструкция может быть перенесена на произвольное множество. Начиная с множества X, мы можем определить следующее множество (обозначается $X^{\oplus}$): $X^{\oplus} = X \square \{X\}$. Чтобы использовать этот процесс для построения N, требуется некоторое начальное множество. Выберем в качестве такого множество {Ø}. Оно имеет один элемент (многие авторы начинают с Ø. Это порождает множество $\{0\} \cup N$. Мы не считаем 0 натуральным числом, и в этом причина такого выбора начального элемента. Не существует универсального условия относительно 0 и N. Всегда следует проверять условные обозначения, принятые в других книгах, при обращении к ним.) Из {Ø} создадим последовательности $\{Ø\}:\{Ø\}^{\oplus} = \{Ø,\{Ø\}\}$, $\{Ø\}^{\oplus\oplus} = \{Ø,\{Ø\},\{Ø,\{Ø\}\}\}$ и т.д.

Это приводит к прогрессии, которая является более привлекательной, чем

$$1, 1 + 1, 1 + 1 + 1, ...,$$

по крайней мере ее конструкция является строго определенной. Чтобы навести порядок в вышесказанном, выберем временно имена для этих множеств.

Переименуем {Ø} как 1, $1^{\oplus}$ как 2, $2^{\oplus}$ как 3 и т.д. Тогда

$$1=\{\,\emptyset\,\},$$
$$2=\{\,\emptyset, 1\},$$
$$3=\{\,\emptyset, 1,2\},,.,$$

Поэтому, например, множество 3 имеет три элемента. Во избежание неточности, давайте снова изменим обозначение и определим множества

$$N_m=m\backslash\{\emptyset\}=\{1, 2, ...,m\},$$
$$N = N_1 \square (N_{m_{\oplus}} : m \in N).$$

Тогда из определения следует, что $|N_m|=m$ (число m) и что если $a, b \in N$, то $a \leq b$ тогда и только тогда, когда $N_a \subseteq N_b$. Поэтому наша вера в упорядоченность N формально обоснована. Итак, множества N и N_m (для каждого $m \in N$) определены и могут быть использованы в дальнейшем. Введем некоторые понятия.

Определение. Два множества *биективны* (обозначается $X \sim Y$), если между ними существует биекция. Непустое множество *конечно,* если оно биективно некоторому N_m ($m \in N$). Если $X \sim N_m$, то мощность множества (обозначается $|X|$) равна m. (Числа в данном случае используются как прилагательные. Например, если P — множество всех людей и $X \subseteq P$ таково, что $X \sim N_m$, то X есть множеством из m людей.) Напомним, что пустое множество Ø биективно только по отношению к себе, является конечным и имеет мощность 0, т.е. $|\emptyset| =0$.

Говорят, что множество *счетно,* если оно биективно N. Символ $\aleph_0$ (алеф-нуль) часто используют для обозначения мощности N. Множество называется *счетным,* если оно конечно или счетно, и может быть сосчитано с использованием биекции $f: N \to X$, если X счетно, или биекции $f: N_m \to X$, если $|X|=m$, или $f: \emptyset \to X$; i-й элемент X является образом i отображения f.

Перед тем как установить несколько полезных результатов, отметим одно важное свойство множеств и биекций: отношение σ, которое определено на множестве S посредством

$$\sigma = \{(X, Y): X \sim Y\},$$

является отношением эквивалентности, а подмножества S, которые входят в классы эквивалентности, состоят из множеств, которые

имеют одинаковую мощность. Следовательно, чтобы продемонстрировать тот факт, что два множества имеют один и тот же размер, трkбуется построить биективное отображение между ними.

Пример 29. Покажем, что $|N|=|Z|$.

Отображение

$$\psi : n \to \begin{cases} \dfrac{1-n}{2}, & \textit{если } n \textit{ нечетное,} \\ n/2 & \textit{в противном случае} \end{cases}$$

является биекций между N и Z.

В примере 29 греческая буква ψ использовалась для обозначения биекции. Использование греческих букв φ, ψ, χ..... для обозначения произвольных биекций является общепринятым в текстах по логике и будет здесь использоваться в этом контексте (среди других). Однако, чтобы не было путаницы с пустым множеством Ø, мы будем избегать использования этих букв в этом разделе.

Пример 30. Покажем, что $|N|=|Q|$. Это требует несколько более сложных соображений. Вначале рассмотрим счетное количество копий N, каждая из них соответствует своему номеру $n \in N$. Мы можем записать это множество как $N \times N$ и упорядочивать его элементы, как показано на рис. 23. Такое упорядочение является биекцией $N \times N \to N$, которая задается отношением

$$(x, y) \to \frac{(x+y-1)(x+y-2)}{2} + y.$$

Каждый положительный элемент Q может быть связан с дробью (p, q), где p и q взаимно простые, и связан с элементом (p, q) множества $N \times N$ естественным образом. Поэтому, записывая

$$T=\{x: x \in Q, x>0\},$$

получаем $$|T| \le |N \times N| = |N|$$

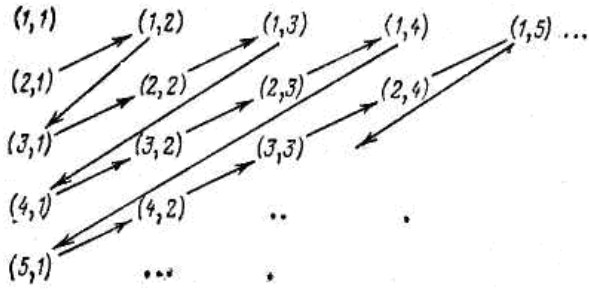


Рис. 23.

(Использование отношения «≤» между этими бесконечными числами не обосновано. Соотношение $|A|\leq|B|$ следует читать как «A биективно подмножеству B». Доказательство неочевидного факта, что это отношение является отношением порядка, лежит за пределами этой книги). С другой стороны, каждое $p\in N$ может быть представлено как $p/1$ и, следовательно, связанно с парой $(p,\ 1)$. Поэтому $|N|\leq|T|\leq|N|$, откуда следует, что $|T|=|N|$. Для достижения нашей цели необходимо сказать следующее. Возьмем диагонально упорядоченное множество $N\times N$ (см. рис. 23) и выбросим из него пары, которые имеют нетривиальный общий множитель. Это дает метод нумерации элементов T, однако трудно дать формулу, которая связывала бы элементы T с элементами N. Теперь мы должны повторить наше рассуждение применительно ко всем рациональным числам. Это можно сделать несколькими способами. Выберем для наглядности следующий. Расширяя уже полученное соответствие между T и $N\times N$ до оператора между Q и $Z\times N$, получаем

$$|N|=|Z|\leq|Q|\leq|Z\times N|=|N\times N|=|N|,$$

что дает требуемый результат (построение биекции между $Z\times N$ и $N\times N$ оставляем как упражнение).

Предыдущий пример несколько длинен, однако при его рассмотрении возникло несколько важных моментов, которые мы сейчас отметим.

1. Если S конечно и $\chi: S\to S$ — инъективное отображение, тогда χ биективно.

Доказательство. Пусть выполнены условия утверждения. Если $S=\varnothing$, то требуемый результат тривиален. Если $S\neq\varnothing$, тогда существует биекция $\psi: N_m\to S$ для некоторого $m\in N$ и отображение $\psi^{-1}\circ\chi\circ\psi$ инъективно: $N_m\to N_m$ и, следовательно, является биекцией. (Доказательство этого факта оставляем как упражнение.).

Основная идея заключается в переупорядочивании m объектов и известная по названию «принцип раскладывания по гнездам». Дано m гнезд, каждое в своем ящике. Любая схема переселения, при которой в одном ящике может быть не более одного гнезда, должна использовать все m ящиков, т.е. $\psi^{-1}\circ\chi\circ\psi$ является перестановкой N_m (рис. 24). Однако ψ — биекция; следовательно, $\chi\circ\psi$ и χ также биекции. (Обратно, если χ: $S\to S$ не сюръективно, то S должно быть бесконечным.)

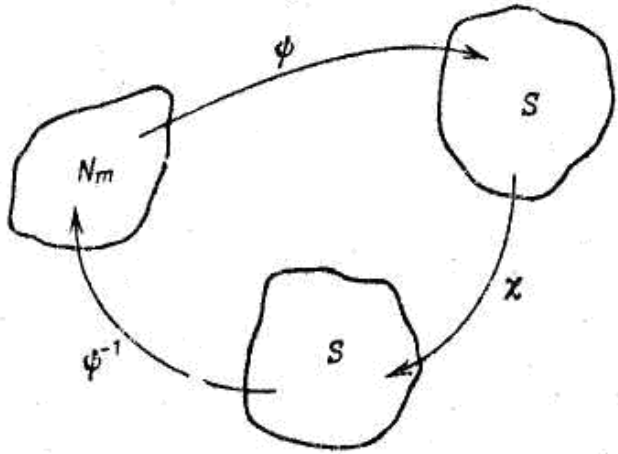


Рис. 24.

2. Множество N бесконечно, так как отображение на N, определенное как $n \square n+1$, инъективно, но не биективно (нет элемента, который отображается в 1). Следовательно, обращая предыдущий результат, получаем, что N не может быть конечным.

Подмножество A из R *ограничено сверху (снизу)* если существует *верхняя (нижняя) граница. А ограничено,* если оно ограничено сверху и снизу.

3. Ограниченное подмножество из N конечно.

Доказательство. Каждое подмножество из N ограничено снизу нулем. Пусть $A\subseteq N$ ограничено сверху некоторым $m\in N$. Определим отображение χ: $A\to N$ так, что если $A=\{a_1, a_2,..., a_t, ...\}$ и $a_1<a_2< a_3<....<m$ (такой порядок возможнен, так как $A\subseteq N$), то $\chi(a_i)=u$.

Следствием этого является соотношение $\chi(a_i)\le a_i$ и χ, очевидно, инъективно. Оно также должно быть биекцией, т.е. χ: $A\to N$ для некоторого $n\le m$. Если это не так, тогда существует $a_p\in A$ такое, что

$\chi(a_p)>m$ и, таким образом, $a_p \geq \chi(a_p)>m$. Однако A ограничено m; поэтому мы пришли к противоречию. Следовательно, χ — биекция на N_n и A конечно.

4. Каждое подмножество конечного множества конечно.

Доказательство. Пусть $A \subseteq B$ и B конечно. Если $B=\varnothing$, то $A=\varnothing$, и утверждение доказано. В противном случае $B \sim N_m$ для некоторого $m \in N$. Тогда существует биекция $\chi: B \rightarrow N_m$. Применение χ к A дает подмножество из N_m и потому $\chi(A)$ ограничено. Из случая 3 следует конечность $\chi(A)$. Поэтому, так как $\chi(A)$ биективно с A , то A конечно.

5. Прямым следствием случая 4 является тот факт, что любое множество, которое имеет бесконечное подмножество, само бесконечно.

Доказывая некоторые неочевидные факты о размерности множеств Z, Q, N, $N \times N$, ..., разумно задаться вопросом: существуют ли множества, мощность которых больше мощности N? Ответ на этот вопрос утвердительный. Действительно, из данного произвольного множества мы можем (используя диагональную процедуру Кантора; см. ниже) создать множество, мощность которого строго больше мощности N.

Мы не будем рассматривать общий случай, и ограничимся рассмотрением примера, который носит фундаментальный характер и отвечает целям нашего изложения.

Пример 31. Покажем, что

$$|[0, 1[|>|N|, \quad [0, 1[=\{x: x \in R, 0 \leq x<1\}.$$

Доказательство (А. Кантор). Каждое число между 0 и 1 может быть записано в виде бесконечной десятичной дроби $0,d_{n1}d_{n2}d_{n3}...$ Предположим, что эти числа могут быть перенумерованы и что n-е число имеет значение, которое дано выше. Будем конструировать следующее число: возьмем n-ю цифру десятичной формы, равную n-му числу. Это дает нам число $0,d_{11}d_{22}d_{33}...$. Построим новое число: $0,\delta_{11}\delta_{22}\delta_{33...}$, где каждая цифра δ_{ii} отличается от соответствующей цифры d_{ii}. Тогда это построенное число будет отличаться от каждого числа из первоначального перечня, а именно от n-го числа оно будет отличаться n-й цифрой. Следовательно, мощность $[0, 1[$ строго больше, чем мощность N, и счетной биекции не существует.

Принцип построения числа $0,\delta_{11}\delta_{22}\delta_{33...}$ в предыдущем доказательстве есть наиболее важным моментом, хотя следует заметить, что существуют задачи, где десятичное представление чисел не единственно. Это встречается тогда, когда представление заканчивается бесконечной последовательностью 0 или 9. (Например,

0,3999... и 0,4000...) Чтобы построенное число $0,\delta_{11}\delta_{22}\delta_{33}...$ отличалось от уже выписанного списка чисел, мы можем оговорить, что δ_{ii} должны отличаться от 0, d_{ii} и 9. Соответствующие проблемы возникают и в других подобных конструкциях, однако, чтобы не отвлекать внимания от основных идей, в большинстве случаев мы будем их игнорировать.

На самом деле можно показать, что $[0, 1[\sim R$. Очевидно, что R — важное множество. Поэтому его мощность обозначают специальным символом, о чем мы уже говорили, $\aleph_1$ (алеф-один). Рассмотрим теперь способы, с помощью которых можно объединять множества и отношения между мощностями отдельных множеств и мощностью результирующего множества. Первый из них довольно очевиден.

Теорема. *Если $A \sim B$ и $C \sim D$, то $(A \times C) \sim (B \times D)$.*

Доказательство. Пусть $\chi: A \to B_m$ и: $\psi: C \to D$ — биекции. Тогда $(a, c) \square (\chi(a), \chi(c))$ является биекцией между $A \times C$ и $B \times D$.

Теорема. *Если Z — конечное множество и $\{X, Y\}$ — разбиение Z, тогда $|Z| = |X| + |Y|$.*

Доказательство. Так как Z конечно, то $Z \sim N_m$ для некоторого $m \in N$, и существует биекция $\chi: Z \to N_m$ (рис. 25).

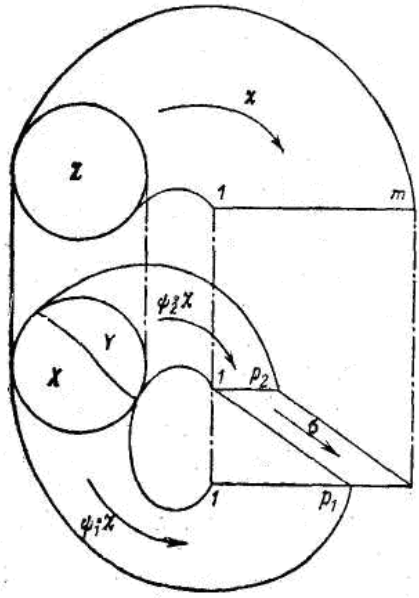


Рис. 25.

Боее того, так как $X \subseteq Z$ и $Y \subseteq Z$, то $\chi(X) \subseteq N_m$ и $\chi(Y) \subseteq N_m$. Пусть ψ_1 является биекцией из $\chi(X)$ в N_{p_1} для некоторого $p_1 \leq m$, где $(\psi_1 \circ \chi)(X) = N_{p_1}$. Аналогично пусть ψ_2 — биекция из $\chi(Y)$ в N_{p_2} для некоторого $p_2 \leq m$, где $(\psi_2 \circ \chi)(Y) = N_{p_2}$. Тогда если σ: $N \to N$ определяется как σ: $x \square x + p_1$, то σ является биекцией между N_{p_2} и $N_{p_1+p_2} \setminus N_{p_1}$.

Следовательно, отображение

$$z \text{ a } \begin{cases} \psi_1 \circ \chi(z), & \text{если} \quad z \in X, \\ \sigma \circ \psi_2 \circ \chi(z), & \text{если} \quad z \in Y. \end{cases}$$

является инъекцией и биективно (так как Z конечно) между Z и $N_{p_1+p_2}$. Таким образом, $Z \sim N_{p_1+p_2}$. Следовательно, $m = p_1 + p_2$ и $|Z| = |X| + |Y|$.

Пример 32. Чтобы проиллюстрировать применение предыдущей теоремы, рассмотрим случай, когда

$$Z = (a, b, c, d, e, f),$$
$$X = (b, e),\ Y = (a, c, d, f).$$

Тогда биекция между Z и N_6 задается следующим образом:

$$\chi = \{(a,5), (b,1), (c,6), (d,3), (e,4), (f,2)\},$$

и поэтому

$$\chi(X) = \{1, 4\},\ \chi(Y) = \{2, 3, 5, 6\}.$$

Подходящими отображениями ψ_3 являются

$$\psi_1 = \{(1, 2), (4, 1)\},\ \psi_2 = \{(3, 1), (5, 2), (6, 3), (2, 4)\}.$$

Поэтому если σ определяется как $x \text{ a } x + 2$, то

$$\psi_1 \circ \chi(X) = \{1, 2\},$$
$$\psi_2 \circ \chi(Y) = \{1, 2, 3, 4\},$$
$$\sigma \circ \psi_2 \circ \chi(Y) = \{3, 4, 5, 6\}.$$

Комбинирование $\psi_1 \circ \chi$ и $\sigma \circ \psi_2 \circ \chi$ дает нужный результат.

Предыдущий пример не является характерным в том смысле, что он требует очень тщательных манипуляций с биекциями. Обычно в этом нет необходимости, так как почти всегда можно сослаться на хорошо известные результаты.

Закончим обсуждение следующим основным результатом. Если A и B — конечные множества, то $A \times B$ обычно и $|A \times B| = |A| \bullet |B|$. Этот результат дает способ для введения *доказательства по индукции.*

Доказательство по индукции использует два основных понятия, которые содержатся в определении N:

(I) Существует некоторый начальный элемент (в N это 1).

(II) Для заданного утверждения, соответствующего некоторому элементу, существует метод, позволяющий перейти к следующему (в случае N это — создание следующего числа за наибольшим до сих пор числом, включенным в N).

Более конкретно, если для некоторого $n_0 \in N$ (обычно $n_0=1$, но не обязательно) мы можем доказать, что утверждение $P(n_0)$ справедливо и для любого $n \in N$ ($n \geq n_0$) справедливость $P(n)$ влечет справедливость $P(n+1)$ (здесь $n+1$ — следующий за n элемент), то заключаем, что $P(n)$ справедливо для всех $m \geq n_0$. Шаг (I) является *основанием индукции,* шаг (II)— *шагом индукции.*

Весь процесс, по существу, является прямым доказательством

$$P(n_0) \Rightarrow \ldots \Rightarrow P(m)$$

и, следовательно, осуществляет непосредственную проверку промежуточных результатов.

Рассмотрим следующий пример.

Пример 33. Если A и B конечны, то

$$|A \times B| = |A| * |B|.$$

Доказательство. Поскольку A и B конечны, то $A \sim N_m$ с биекцией τ: $A \to N_m$ и $B \to N_n$ для некоторых m, $n \in N$. Будем использовать индукцию по n — размерности B. Заметим, что от размерности A требуется конечность, так как мы допускаем знакомство только с умножением конечных величин.

Основание индукции. Если $B=\varnothing$, то $A \times B = \varnothing$, и поэтому имеем тривиальное равенство

$$|A \times B| = 0 = |A| \cdot 0 = |A| \cdot |B|.$$

Если $B=\{b\}$, то отображение $A \to A \times B$ такое, что $a \square (a, b)$, видимо, биективно, и поэтому

$$|A \times B| = |A| = |A| \cdot 1 = |A| \cdot |B|.$$

Шаг индукции. Предположим, что

$$|A \times B_k| = |A| \cdot |B_k|,$$

где $B_k \subseteq B$ и $|B_k| = k \in \mathrm{N}$. Тогда

$$|A| \cdot |B_k| = m \cdot k \in \mathrm{N}$$

и существует биекция ψ: $A \times B_k \to N_{m+k}$. Если $k<n$, то можно взять подмножество B, которое имеет $j=k+1$ элементов. Пусть $B_j = B_k \cup \{x\}$, где B_k — множество из k элементов, и пусть отображение χ, $\chi: A \times B_j \to N$ определяется следующим образом:

$$\chi:(a,x) \square\ \tau(a)+m\bullet k,$$

$\chi: A\times\{x\}\to\{m\bullet k+1,\ldots,m\bullet k+m\}\chi:(a,b)$ а $\psi(a,b)$,если $b\in B_k$

Очевидно, что χ является биекцией на N_{m*k+m} и

$$m\bullet k+m= m\bullet (k+1)= m\bullet j$$

Поэтому $|A\times B_j|=m\bullet j=|A|\bullet |B_j|$. Следовательно, тождественность справедлива для всех подмножеств B, которые содержатся в B, и поэтому $|A\times B|=|A|\bullet |B|$.

2.4. Некоторые специальные классы функций

В этом разделе мы немного отойдем от основной темы обсуждения для того, чтобы коротко рассмотреть следующих три важных класса функций: *подстановки, последовательности, функционалы*.

Эти функции часто используются; особенно отметим их приложение к теории графов, к трассированию вычислений, к определению языков программирования и перевода, к машинной графике.

Начнем из подстановок и перстановок. Частично мы их уже рассматривали выше.

Понятие подстановок и последовательности

Определение. *Подстановкой* множества A называется биекция на A.

Подстановки конечных множеств представляют особый интерес в вычислениях. Когда A конечно, мы можем вычислить число разных подстановок A.

Пусть $|A|=n\in N$. Обозначим через ${}_nP_n$ число таких подстановок. Значение ${}_nP_n$ легко вычислить. Можно рассматривать задачу построения биекции на A как задачу заполнения ящиков, пронумерованных от 1 до n (рис. 26), объектами $a_1, \ldots, a_n$.

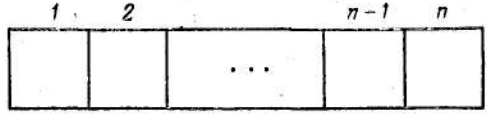


Рис. 26

Порядок, в котором заполняются ящика - несущественен (любой другой порядок можно получить перемешиванием ящиков). Поэтому будем заполнять их слева направо. Первый ящик может быть заполнен n способами, так как мы имеем свободный выбор из всего множества A. Убирая выбранный элемент из A, получим множество из n — 1

элементов. Следовательно, второй ящик может быть заполнен n — 1 способами, третий ящик — n — 2 способами и т.д. Продолжая этот процесс, получим, что (n — 1)-й ящик может быть заполнен двумя способами, а ящик с номером n — единственным оставшимся элементом из A. Следовательно, число разных подстановок из A равно

$$n \cdot (n - 1) \cdot (n - 2) \cdot \ldots \cdot 3 \quad 2 \quad 1.$$

Это произведение называется *факториалом n* (обозначается $n!$). Следовательно, ${}_nP_n = n!$

Так как $A \sim N_n$, то можно свести наше рассмотрение к N_n. Любая подстановка на N_n должна определять образ каждого элемента в N_n (который, безусловно, должен быть единственным и отличным от других). Пусть ψ — подстановка на N_n. Тогда ψ можно определить как множество из n пар следующим образом:

$$\psi = \{(1, x_1), (2, x_2), \ldots, (n, x_n)\},$$

где

$$\{x_1 \ldots, x_n\} = N_n.$$

Не обязательно, конечно, должно быть $x_1 = 1$ и т.д. Можно также представить ψ следующим образом;

$$\psi = \begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ x_1 & x_2 & x_3 & \ldots & x_n \end{pmatrix}$$

Пример 34. Пусть σ — подстановка на N_6 :

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 3 & 1 & 4 & 2 \end{pmatrix}$$

Тогда $\sigma(1) = 5$, $\sigma(3) = 3$ и т.д.

Достоинством этого обозначения - является простота, с которой могут быть вычислены сложные подстановки. Предположим, что ψ — подстановка на N_n, которая определена выше, а χ - другая подстановка на том же самом множестве. Тогда подстановка χ может быть записана как совокупность пар в порядке, определяемом $x_1, x_2, \ldots, x_n$. Если две последовательности записать одну над другой (первая применяемая подстановка должна быть записана первой), то верхняя и нижняя строки дадут результирующую подстановку.

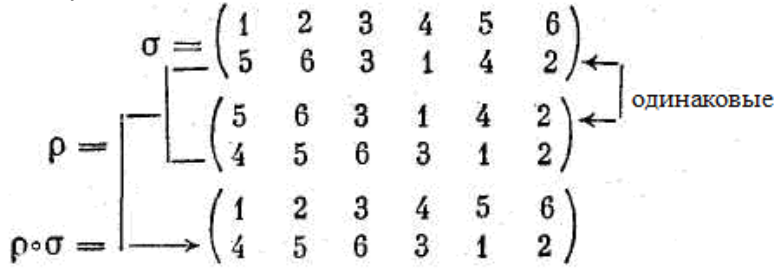
Пример 35. Пусть σ - подстановка из примера 34 и

$$\rho = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 6 & 1 & 4 & 5 \end{pmatrix}$$

Можно переписать ρ в виде

$$\rho = \begin{pmatrix} 5 & 6 & 3 & 1 & 4 & 2 \\ 4 & 5 & 6 & 3 & 1 & 2 \end{pmatrix}$$

Поэтому $\rho\circ\sigma$ может быть вычислено следующим образом:



Следовательно, например,

$$\rho\circ\sigma\ (2)(= \rho(\sigma(2))= \rho(6))=5 \text{ и т.д.}$$

Отсюда следует, что представление обратной (конечной) подстановки выходит перестановкой строк, которые представляют исходную подстановку. Хотя такое представление полезно в вычислениях, оно требует много лишнего места, особенно в тех случаях, когда много элементов не меняются в процессе подстановки. Существует более простое определение, которое может употребляться непосредственно для некоторых простых подстановок и косвенно для всех конечных.

Определение. Пусть $A=\{a_1, \ldots, a_n\}$. Подстановку ρ называют *циклом (циклической подстановкой),* если

$$\rho = \begin{pmatrix} a_1 & a_2 & \ldots & a_{n-1} & a_n \\ a_2 & a_3 & \ldots & a_n & a_1 \end{pmatrix}.$$

Предположим, что $A \subseteq B$ и B конечно. Распространяя ρ на все B, можно определить подстановку σ так, что

$$\sigma : x \text{ a } \begin{cases} \rho(x), & \textit{если} \quad x \in A, \\ x, & \textit{если} \quad x \in B \setminus A. \end{cases}$$

В этом случае σ ведет себя подобно ρ во всех случаях, когда элементы B не остаются на месте. Применение σ к A передвигает элементы по кругу циклическим образом, и, если известна область A, мы можем обозначить подстановку как $(a_1,a_2,\ldots,a_n)$. Эта подстановка называется *циклом длины* n.

Пример 36. Рассмотрим опять подстановку

$$\rho = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 6 & 1 & 4 & 5 \end{pmatrix}$$

Подстановка является циклом длины 5 и может быть записана как

(1, 3, 6, 5, 4).

Не все подстановки являются циклами. Например, подстановка σ в примере 34 не является циклом. Напомним, что σ имела вид

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 3 & 1 & 4 & 2 \end{pmatrix}$$

Поэтому σ(1)=5, σ(5)=4, σ(4)=1, откуда следует, что σ содержит цикл (1, 5, 4). Начиная с 2, получаем другой цикл — (2, 6). Таким образом, имеем σ =(1, 5, 4)∘(2,6) и σ =(2, 6)∘ (1, 5, 4).

В действительности каждая конечная подстановка может быть представлена как произведение циклов, при этом циклы могут располагаться в любом порядке. Из построения следует, что один элемент не может встретиться более чем в одном цикле, т.е. циклы *не пересекаются.*

Теорема. *Каждая подстановка* ρ *на конечном множестве A выражается в виде произведения непересекающихся циклов.*

Доказательство. Поскольку $|A|=n \in N$, то $A \sim N_n$. Поэтому без потери общности мы можем ограничиться рассмотрением подстановки ρ на N_n.

В теореме утверждается, что $\rho = \sigma_1 \circ \sigma_{2\ldots} \circ \sigma_r$, где каждое σ_i является циклом и циклы не пересекаются. Для доказательства теоремы построим необходимые циклы. Сначала найдем наименьший элемент $x_1 \in N_n$ такой, что $\rho(x_1) \neq x_1$ и $\rho(x)=x$ для всех x, $1 \leq x < x_1$. Если такого x_1 не существует, то ρ=1 (т.е. ρ есть тривиальным пустым произведением циклов). В противном случае вычислим x_1, $\rho(x_1)$, $\rho^2(x_1)$, $\rho^3(x_1)$ и т.д. Все эти элементы находятся в N_n. Поэтому элементы в этой последовательности должны содержать повторение. Предположим, что $\rho^k(x_1)$ - первый такой элемент (который уже повторялся в последовательности). Покажем, что $\rho^k(x_1)=x_1$. Предположим, что это соотношение не выполняется. Тогда $\rho^l(x_1) = \rho^k(x_1)$ для некоторого l, $0<l<k$. Следовательно,

$$\rho^{l-1}(x_1) = \rho^{-1}(x_1) \circ \rho^l(x_1) = \rho^{-1}(x_1) \circ \rho^k(x_1) = \rho^{k-1}(x_1) \text{ и т.д.}$$

Поэтому $\rho^{l-l}(x_1)=\rho^{k-l}(x_1)$, т.е. $\rho^{k-l}(x_1)=\rho^0(x_1)=x_1$, что противоречит минимальности k (так как k — $l<k$). Таким образом, $\rho^k(x_1)=x_1$, и подстановка

$$\sigma_1=(x_1, \rho(x_1), \rho^2(x_1), \rho^3(x_1),\ldots,\rho^{k-1}(x_1))$$

задает цикл внутри ρ.

Если все элементы $x \in N_n$ такие, что $\rho(x) \neq x$ (будем называть такие элементы *нестационарными),* содержатся в σ_1, то $\rho=\sigma_1$ — единственный цикл (который, естественно, не пересекается). В противном случае найдем следующий наименьший элемент $x_2 \in N_n$ такой, что $\rho(x_2) \neq x_2$ и x_2 не встречается в σ_1. Из x_2 строим множество различных степеней ρ:

$$\sigma_2=(x_2, \rho(x_2), \rho^2(x_2), \rho^3(x_2),\ldots,\rho^m(x_2))...$$

Это цикл длины не менее 2, и он не пересекается с σ_1. Если все нестационарные элементы исчерпаны, то $\rho=\sigma_1\circ\sigma_2=\sigma_2\circ\sigma_1$. Очевидно, что множество нестационарных элементов, которые не входят в эти циклы, можно уменьшить, и в конце концов придем к Ø. Следовательно, $\rho=\sigma_1\circ\sigma_2\circ\sigma_{3...}\circ\ \sigma_r$, для некоторого $r \in N$.

Рассмотрим теперь несколько другую ситуацию. Возьмем множества $A: |A|=n$ и $B \subseteq A$, $|B|=r \leq n$. Возникает вопрос: сколько биективных функций существует из A в B? Или, что эквивалентно, сколько существует инъективных отображений из B в A? Число перестановок (без повторений) из n элементов по r обозначается ${}_nP_r$, и вычисляется так же, как и ${}_nP_n$, за исключением того факта, когда процесс прекращается после заполнения r ящиков. Таким образом,

$${}_nP_r = n \cdot (n-1) \cdot \ldots \cdot (n-r+1).$$

Легко видеть, что, продолжая процесс заполнения ящиков, оставшиеся n-r элементов можно разместить по последним n-r ящикам ${}_{n-r}P_{n-r}$ способами. Поэтому и

$${}_nP_r = \frac{{}_nP_n}{{}_{n-r}P_{n-r}} = \frac{n!}{(n-r)!}.$$

При вычислении ${}_nP_r$ мы находим число биективных функций из A в B. Подсчитаем число таких функций.

Определение. Пусть A — конечное множество и $B \subseteq A$, $|A|=n \geq r=|B|$. Множество B называется *сочетанием* (без повторений) из n элементов по r. Число таких сочетаниий обозначается через C^r_n.

Вычисление C^r_n производится следующим образом. Положим $|A|=n$. Возьмем произвольное подмножество $B \subseteq A$ такое, что $|B|=r$. Тогда B является образом подстановки из n элементов по r. Число инъективных функций на A, которые имеют B своим образом, является ${}_nP_r$. Если f является такой функцией и g — другая такая функция, которая имеет ту же самую область значений, то g связана с f

соотношением $g=\varphi\circ f$, где φ — подстановка на B. Функции g и f определяют одну и ту же комбинацию, и в действительности число функций, которые определяют эту комбинацию, равно числу подстановок φ на B. Следовательно,

$$ {}_nP_r = C^r{}_n \bullet {}_rP_r $$

откуда

$$ C^r{}_n = \frac{{}_nP_r}{{}_rP_r} = \frac{n!}{r!(n-r)!}. $$

Поскольку относительные дополнения единственны и $|A\backslash B|=n\text{-}r$, то отсюда следует, что $C^r{}_n = C^{n-r}{}_n$.

Вернемся теперь к математическим объектам, которые упоминались нами раньше, но которые не рассматривались как функции.

Определение. *Последовательностью* на множестве S называют отображение $N \to S$.

Если $\sigma: N \to S$ заданная последовательность и $\sigma(n)=s_n$, то, обычно, обозначают последовательность не σ, а (s_n) или $(s_1, s_2, .., s_n , ...)$. В этом случае s_n называют n-м членом последовательности.

Часто при изучении свойств последовательностей возникает понятие «расстояние» между соседними элементами последовательности (скажем, s_n и s_{n+1}) и между элементами s_n при $n \geq n_0$ (где n_0 — некоторый фиксированный элемент N) и фиксированным элементом из S_1.

Мы возвратимся к этим вопросам чуть позже, поскольку в данный момент у нас в общем случае нет понятия расстояния.

Понятие функционала

Говоря об отображении $f: X \to Y$ как о функции с вещественными значениями, мы не накладывали на характер элементов множества X каких-либо особых ограничений. В простейших задачах множество X, как и множество Y представляет собой множество вещественных чисел. В этом случае каждая пара $(x, y) \in f$ ставит в соответствие одному вещественному числу x другое вещественное число y. Однако важным для практики есть случай, когда множество X представляет собой множество функций, а множество Y — множество вещественных чисел. Этот случай приводит к понятию функционала, подробное рассмотрение которого удобно провести на примере.

Представим себе некоторую линию $y=f(x)$, которая соединяет фиксированные точки A и B, как показано на рис. 27, по которой скатывается свободно движущийся шарик. Обозначим через t время, которое шарик затратит на перемещение из точка A в точку B. Это время зависит от характера линии AB, т.е. от вида функции $f(x)$. Если обозначить через $F(x)$ множество различных функций, которые изображают линию AB, а через T множество вещественных чисел t, определяющих время движения шарика, то зависимость времени движения от вида функции может быть записана как отображение

$$J:F(x)>T.$$

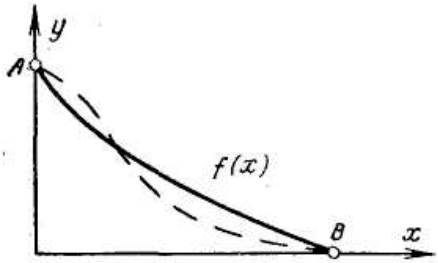


Рис. 27. Линия наискорейшего спуска.

Элементами множества J будут пары $(f(x), t))$, в которых $f(x) \in F(x)$, а $t \in T$. В этом случае говорят, что вещественное число $t \in T$ представляет собой функционала J от функции $f(x) \in F(x)$, и записывают это в виде

$$t=J\,[f(x)]$$

В задачах управления функционалы используются как критерии качества управления. Так, в рассмотренном примере время перемещения шарика из точки A в точку B можно трактовать как критерий «качества» избранной функции $f(x)$. При этом говорят об *оптимальном управлении* как о таком, при котором соответствующий критерий качества оборачивается в минимум. С этой точки зрения определения «оптимального» вида функции $f(x)$ сводится к выполнению условия

$$\min_{f \in F} J\,[f(x)],$$

при котором время t будет минимальным. В математике подобная линия наискорейшего спуска получила название ***брахистохроны***.

Особый интерес представляют функционалы в задачах, связанных с трансляцией языков программирования. Это связано с тем, что рассматриваемые функционалы используются как объекты особого

рода, в ряде случаев отличные от элементов, которые были в области определения и области значения функций. Конечно, множества функций могут рассматриваться так же, как и любые другие множества.

В развитых языках программирования имена целых переменных не отличаются от имен переменных функций и могут изучаться аналогичными способами. Хотя эти функции являются довольно сложными, языки программирования редко дают примеры важности функционалов.

Пример 37. Пусть P — множество программ, т.е. текстов программ (строк символов), что должны быть обработаны компилятором. Аналогично пусть I и O — множества соответственно входных и выходных значений, которые доступны программе для ввода и вывода. Тогда компилятор (с соответствующего языка) является функционалом типа $P \to [\, I \to O]$; для данной $p \in P$ он должен создать машинный код, который при выполнении будет читать $i \in I$ и выдавать $o \in O$.

Пример 38. Пусть все данные принадлежат R. Тогда, если

f: $a \square$ [x a $a+x$], то

$$f(2): x \text{ a } 2+x \text{ и } f(2)(3) = 5,$$

в то время, как $f(3)$: x a $3 + x$ и $f(3)(3) = 6$ и т.д.

Обращение с функционалами не вызывает трудностей при условии, что ссылка делается на основной функционал (т. е. $A \to B$ или $A \to [B \to C]$). Следовательно, в дальнейшем мы будем рассматривать их просто как функции, имеющие нетривиальные области значений, и будем обращаться с ними соответствующим образом.

В заключение определим функции, которые сохраняют некоторые структуры. Из дальнейшего будет видно, что в некоторых ситуациях желательно сохранить многие из алгебраических свойств, которыми множества могут обладать. Ограничимся вначале рассмотрением простейшего случая.

Определение. Пусть X — множество, на котором задано отношение эквивалентности ρ. Тогда X *разбивается отношением* ρ *на* ρ *-эквивалентные классы;* множество классов обозначается как X/ρ.

Определение. Пусть X и Y— множества, ρ_X и ρ_Y — отношения эквивалентности на них, и пусть f: $X \to Y$—отображение. Обозначим через $\hat{f}$ *отношение*

$$\hat{f}: X/\rho_X \to Y/\rho_Y$$

такое, что

$$\hat{f} = \{([x],\ [f(x)]):\ x \in X\},$$

где $[x]$ — класс эквивалентности x. Если f — функция, то

$$x_1 \rho_x x_2 \Rightarrow \hat{f}([x_1]) = \hat{f}([x_2]),$$

и f является *отображением, сохраняющим эквивалентность*. В этом случае говорят, что f: $X \to Y$ *индуцирует* отображение

$$\hat{f}:\ X/\rho_x \to Y/\rho_y. \ //$$

Наглядный способ представления такого отображения дан на рис. 28.

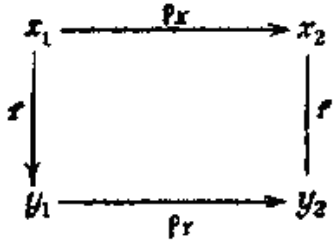


Рис. 28

Если рассмотреть отображение f, согласованное с отношением эквивалентности, то можно переходить от x_1 к y_1 или через x_2, используя соотношения $y_2=f(x_2)$ и $x_2 \rho_X x_1$, или через y_1, используя соотношения $y_1=f(x_1)$ и $y_2 \rho_Y y_1$.

Пример. Пусть $X = \{1, 2, 3\}$, $Y = \{1, 4, 9\}$, и пусть ρ_X и ρ_Y таковы, что

$$X/\rho_x = \{\{1\}, \{2, 3\}\}, \quad Y/\rho_y = \{\{1\}, \{4, 9\}\},$$

и f: $X \to Y$ такое, что x a x^2. Тогда

$$\hat{f}([1]) = [f(1)] = [1] = \{1\},$$
$$\hat{f}([2]) = [4] = \{4, 9\},$$
$$\hat{f}([3]) = [9] = \{4, 9\}.$$

В этом случае $\{2, 3\} \in X/\rho_x \Rightarrow 2 \rho_x 3 \Rightarrow [2] = [3]$ и $\hat{f}([2]) = \hat{f}([3])$. Поэтому f является функцией и f сохраняет отношения эквивалентности.

Пример. Пусть X, Y и f те же, что и раньше, и отношение эквивалентности σ_X и σ_Y индуцируют разбиения $\{\{1\}, \{2, 3\}\}$ и $\{\{1, 4\}, \{9\}\}$ соответственно. В этом случае индуцированные отношения дают

$$\widehat{f}([2]) = [f(2)] = [4] = \{1, 4\},$$
$$\widehat{f}([3]) = [f(3)] = [9] = \{9\}.$$

Так как $2\sigma_X 3$, то [2] = [3] в X/σ_X, но $(4, 9) \notin \sigma_Y$, поскольку $[4] \neq [9]$ в Y/σ_Y. По сравнению с рис. 1.28 этот пример дает отношения, показанные на рис. 29.

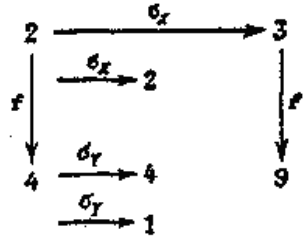


Рис. 29

Так как нельзя соединить стороны прямоугольника во всех случаях, то отношения эквивалентности не сохраняются. Эти диаграммы могут быть использованы для определения операций таким образом, чтобы соединить углы прямоугольника. После этого можно будет объединять диаграммы подобно строительным блокам.

Функция времени

В основе понятия функции времени лежит множество $T \subseteq R$ с элементами t, которое называют множеством моментов времени. ***Время обладает той характерной особенностью, что имеет направление***. Это означает, что если $t_1, t_2 \in T$ и $t_1 < t_2$, то момент t_1 предшествует моменту t_2. Другими словами T — **упорядоченное множество.**

Функция времени определяет отображение f множества моментов времени T на множество вещественных чисел R :

$$f: T \rightarrow R. \qquad (1.65)$$

Элементами f будут пары (t, x), которые обозначаются также через $x(t)$, где $t \in T$, $x \in R$. Каждая такая пара определяет значение функции в момент t и называется ***событием*** или ***мгновенным значением функции***. Полная совокупность пар (t, x), т.е. значений $x(t)$ для всех $t \in T$, и представляет собой функцию времени. Дальнейшее уточнение функции время связано с уточнением ее области определения, т.е. вида множества T.

Если $T=R$, т.е. t может принимать любое вещественное значение от $-\infty$ до $+\infty$, то функция $x(t)$ называется функцией с *непрерывным*

временем. Примером может служить синусоидальная функция времени $x(t)=A\sin(\omega t+\varphi)$, описывающая напряжение в сети переменного тока.

Однако нас обычно не интересуют весьма удаленные моменты времени как в прошлому, так и в будущем. Поэтому производят сужение $x(t)$ на ограниченный интервал $t_1<t\le t_2$, который обычно считают полузакрытым интервалом и обозначают $(t_1,t_2]$. Полузакрытые интервалы времени удобны тем, что допускают последовательное сочленение друг с другом. Так, если интервал $(t_1,t_2]$ разбить моментом t' на два интервала $(t_1, t']$ и $(t', t_2]$, то не будет сомнений, к которому интервалу отнести t'.

Сужение функции $x(t)$, заданной на интервале $-\infty<t<+\infty$, на интервал $(t_1, t_2]$ называется *отрезком функции* $x(t)$ и обозначается $x_{(t_1,t_1]}$. Итак, по определению

$$x_{(t_1,t_1]}=\{(t)|t\in(t_1,t_2]\} \qquad (66)$$

Для осуществления операции сужения часто используют специальную функцию времени, которую называют *единичной функцией* или единичным скачком:

$$1(t-\lambda)=\begin{cases}0 & при \quad t\le\lambda;\\ 1 & при \quad t>\lambda,\end{cases} \qquad (67)$$

приведенную на рис. 30,*а*.

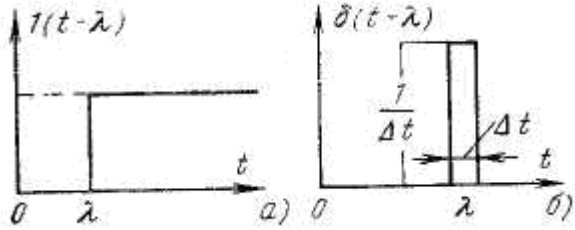


Рис. 1.30. Единичный скачок и импульсная функция .

Так, напряжение, которое подается на вход прибора, который подключается к сети в момент $t=\lambda$, будет равно:

$$u(t)=\mathbf{1}(t-\lambda)x(t)=\mathbf{1}(t-\lambda)A\sin(\omega t+\varphi).$$

Другой широко используемой функцией времени является импульсная функция $\delta(t-\lambda)$, определяемая соотношениями:

$$\delta(t-\lambda)=\begin{cases}0 & при \quad t\neq\lambda;\\ \infty & при \quad t=\lambda,\end{cases} \tag{68}$$

$$\int_{\lambda-\varepsilon}^{\lambda+\varepsilon}\delta(t-\lambda)\,dt=1, \quad \varepsilon>0. \tag{69}$$

Функцию $\delta(t—\lambda)$ можно рассматривать как предельный случай приведенного на рис. 30,*б* прямоугольного импульса шириной Δt и высотой $1/\Delta t$, появляющегося в момент $t=\lambda$ при $\Delta t\to 0$.

Импульсная функция позволяет выделять мгновенные значения функции $x(t)$ для фиксированных моментов времени. Так, если $t_1<\lambda<t_2$, то

$$\int_{t_1}^{t_2}x(t)\delta(t-\lambda)\,dt=x(\lambda)\int_{\lambda-\varepsilon}^{\lambda+\varepsilon}\delta(t-\lambda)\,dt=x(\lambda). \tag{70}$$

Если множество T представляет собой множество натуральных чисел

$$\ldots\ , -2, -1, 0,\ 1, 2,\ ,.,\ , n,\ \ldots\ ,$$

то говорят о функциях *с дискретным временем.* В этом случае элементы множества T обозначают через n, так что пара $(n,\ x)$, которая обозначается также $x[n]$ или x_n, определяет значение функции в момент n. На рис. 31 приведен пример функции с дискретным временем.

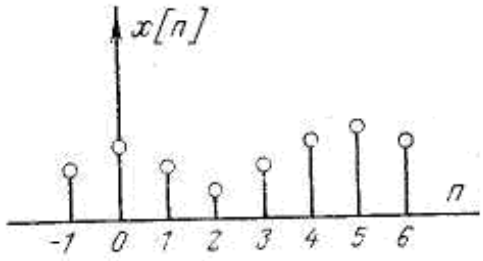


Рис. 31. Функция с дискретным временем.

Понятие оператора

Оператором L называется отображение

$$L:X\to Y, \tag{71}$$

в котором множества X и Y являются множествами функций с элементами $x(t)$ и $y(t)$, так что элементами множества L будут пары $(x(t)$ и $y(t))$. В этом случае говорят, что оператор L преобразует функцию $x(t)$ в функцию $y(t)$, и пишут:

$$y(t)= L\,[x(t)].$$

Рис. 32. Представление управляющей системы в виде оператора.

Примером оператора служит оператор дифференцирования p, который ставит в соответствие функции $f(x)$ другую функцию $f'(x)=df(x)/dx$, что может быть записано в виде

$$f'(x)=p\,[f(x)].$$

В задачах управления роль оператора часто выполняет сама управляющая система, которая преобразует по некоторому закону L входной сигнал $x(t)$ в выходной сигнал $y(t)$, как это показано на рис. 32.

Последовательности

Здесь содержится материал, использующий теорию множеств. Цель, которая при этом преследуется, состоит не в развитии техники вычислений, а в создании строгих утверждений типа:

«Предел $f(x)$ при x, стремящемся к 0, есть y»,

«Наклон графика f в точке a равен b»,

«f имеет гладкий график» и т. п.

(Два последних понятия относятся к графике.) Мы приведем основные определения, которые используются при получении некоторых результатов. Этого достаточно для того, чтобы проиллюстрировать доказательства большинства теорем.

Определение. *Вещественной последовательностью* называется отображение N на R.

Последовательность записывают в виде (a_n). Если при возрастании n члены a_n становятся «близкими» к некоторому фиксированному значению $a \in \mathrm{R}$, то говорят, что последовательность (a_n) имеет предел a или что a_n стремится к a при стремлении n к бесконечности. Дадим строгое определение сказанному.

Определение. Если (a_n) - вещественная последовательность и для любого $\varepsilon > 0$ существует $N_\varepsilon \in N$ такое, что $N > N_\varepsilon \Rightarrow |a_N - a| < \varepsilon$, то говорят, что (a_n) имеет *предел a,* и записывают это как

$$\lim_{n\to\infty} a_n = a$$

или $a_n \to a$ при $n \to \infty$. (Здесь $|x|$ обозначает модуль числа $x \in R$).

Если (a_n) имеет предел, то говорят, что последовательность *сходится.* Если последовательность не имеет предела, то говорят, что она *расходится.*

Пример.

1. Последовательность (a_n), где $a_n = 1/n$, имеет предел 0; для $\varepsilon > 0$ можно выбрать N_ε — любое натуральное число, большее $1/\varepsilon$. Тогда

$$N > N_\varepsilon \Rightarrow |a_N - 0| = 1/N < 1/N_\varepsilon < \varepsilon;$$

следовательно,

$$\lim_{n\to\infty} \frac{1}{n} = 0.$$

2. Последовательность (a_n), где $a_n = (-1)^n$, расходящаяся.

Предложение. *Если* (s_n) *и* (t_n) *—последовательности и* $\lambda \in R$, *тогда* $(s_n + t_n)$, $(s_n t_n)$ *и* (λs_n) *также являются последовательностями, и если* $\lim_{n\to\infty} s_n = s$ *и* $\lim_{n\to\infty} t_n = t$,

то:

а) $\lim_{n\to\infty} (s_n + t_n) = s + t$;

б) $\lim_{n\to\infty} (s_n t_n) = st$;

в) $\lim_{n\to\infty} (\lambda s_n) = \lambda s$;

г) *если* $\varepsilon \neq 0$, *то* $\varepsilon_n/t_n \to s/t$ *при* $n \to \infty$.

Доказательство. Пусть $\varepsilon > 0$. Тогда существует $N_\varepsilon \in N$ такое, что

$$|s_N - s| < \varepsilon/2 \text{ и } |t_N - t| < \varepsilon/2$$

при $N > N_\varepsilon$. Так как при $N > N_\varepsilon$,

$$|s_N + t_N - (s + t)| = |s_N - s + t_N - t| \leqslant$$
$$\leqslant |s_N - s| + |t_N - t| < \varepsilon,$$

то $\lim_{n\to\infty} (s_n + t_n) = s + t$.

Аналогично для случая б)

$$|s_N t_N - st| = |s_N t_N - s_N t + s_N t - st| \leqslant$$
$$\leqslant |s_N t_N - s_N t| + |s_N t - st| \leqslant |s_N| \cdot |t_N - t| + |s_N - s| \cdot |t|.$$

Пусть задано $\varepsilon > 0$. Тогда существует $N_\varepsilon \in \mathrm{N}$ такое, что для $N > N_\varepsilon$ справедливы неравенства

$$|s_N - s| < \frac{1}{2}\frac{\varepsilon}{|t|+1},$$

$$|t_N - t| < \frac{1}{2}\frac{\varepsilon}{|s|+1}, \quad |s_N| < |s| + 1.$$

Следовательно,

$$|s_N||t_N - t| + |s_N - s||t| \leqslant (|s| + 1)|t_N - t| +$$

$$+ |s_N - s||t| < \frac{1}{2}\varepsilon + \frac{1}{2}\frac{\varepsilon}{|t|+1}|t| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon,$$

откуда получаем $(s_n t_n) \to st$. Доказательство случаев в), г) предложения оставляем в качестве упражйения.

Определение. Пусть (a_n) — последовательность в R. Последовательность

$$s_n = \sum_{i=1}^{n} a_i$$

определяет *ряд* $\sum a_n$. При этом s_n называют *n-й частичной суммой ряда.* Если последовательность (s_n) сходится, то говорят, что ряд *сходящийся,* и число

$$\lim_{n\to\infty} s_n$$

называют *суммой ряда.* Оно обозначается

$$\sum_{n=1}^{\infty} a_n.$$

Операции

Определение. *Операцией над множеством* S называется функция $f: S^n \to S, n \in \mathrm{N}$.

В этом определении есть два важных момента, которые заслуживают особого вспоминания. Во-первых, раз операция является функцией, то результат применения операции *однозначно определен.* Поэтому данный упорядоченный набор из n элементов S функция f переводит только в один элемент S. Во-вторых, поскольку область значений операции лежит в S, на которое операция действует, будем говорить, что операция *замкнута* на S,

Говорят, что операция $S^n \to S$ *имеет порядок* n. Ограничимся рассмотрением ситуаций, когда порядок равен 1 или 2. В этом случае

операции называют *монадическими* (или *унарными)* и *диадическими* (или *бинарными)* соответственно. Элементы набора из n элементов в области определения называют операндами. Операции обычно обозначают символами, которые называют операторами. В случае унарных операций обычно символ оператора ставят перед операндом.

Наиболее простым примером является операция изменения знака на R. В предположении, что операция сложения уже определена, $-x$ определяет операцию $x \square\ y$: $x+y=0$ (x отображается в y: $x+y=0$).

Определение. Бинарные операции обозначают одним из трех способов. В первом случае оператор ставится между операндами *(infix),* во второму — перед операндами *(prefix)* и в третьему — после операндов *(postfix).*

Пример 39.

$a+b$ *infix,*
$+ab$ *prefix,*
$ab+$ *postfix.*

Переход от одной формы к другой нетруден и лучше всего описывается в терминах ориентированных графов.

В соответствии с большинством математических традиций, кроме некоторых работ по алгебре и формальной логике, мы будем использовать обозначение *infix.* Другие обозначения имеют то преимущество, что не требуют скобок при определении порядка вычислений сложных выражений, и это делает их особенно удобными для автоматической обработки. Можно проверить соответствие между следующими парами выражений, записанными в формах *infix* и *postfix* соответственно:

а) $a+b\bullet\ c+(d+e\bullet\ (f+g),$
$abc\bullet\ +\ defg+\bullet\ ++;$
б) $(a+b)\bullet\ c+d+e\bullet\ f+g,$
$ab+c\bullet\ d+ef\bullet\ +g+;$
в) $a+(b\bullet\ (c+d)+e)\bullet\ f+g,$
$abed+\ e+f\ +g+.$

Пример 40. Рассмотрим алгебраическое выражение

$$a + b \bullet c + (d + e \bullet (f + g))$$

и его представление на рис. 33, которое называют деревом.

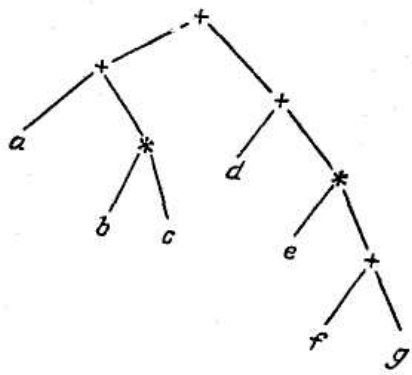


Рис. 33.

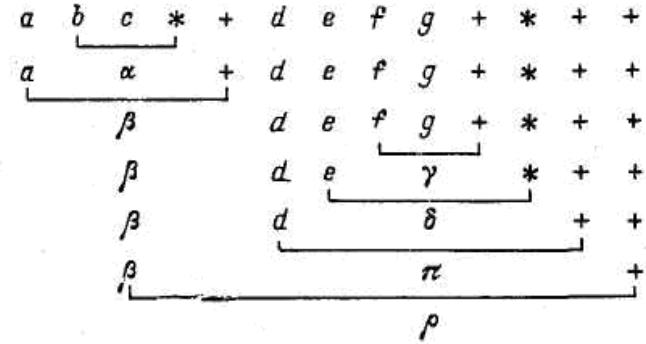
Из свойств арифметических операций мы знаем, что значение этого выражения можно вычислить многими способами. Однако если двигаться слева направо и снизу вверх, то получаем

$$\alpha \leftarrow b \bullet c,\ \beta \leftarrow a + \alpha,\ \gamma \leftarrow f + g,$$
$$\delta \leftarrow e \bullet \gamma,\ \pi \leftarrow d + \delta,\ \rho \leftarrow \beta + \pi.$$

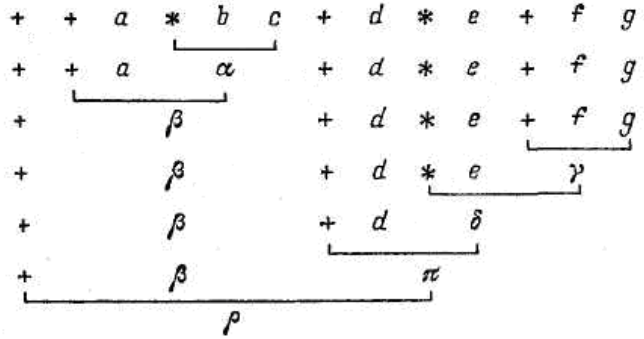
Здесь греческими буквами обозначаются промежуточные результаты, за исключением ρ - искомого результата.

Вычисление значения этого выражения с помощью дерева выполняется очень просто, однако если работать непосредственно с исходным выражением, то это можно сделать по-иному. Действительно, обычно *(infix)* выражения, как это показано в примере, нерегулярно потому, что некоторые подвыражения заключены в скобки, а некоторые нет. Особенно такая ситуация будет наблюдаться в том случае, если проинтегрировать информацию о разных символах на дереве (поскольку на самом деле его нет). Очевидно, что формы записи *prefix* и *postfix* этого выражения несут больше информации.

Вычисление значения выражения в форме *postfix* осуществляется следующим образом:



Аналогично в форме *prefix* вычисления осуществляются следующим образом:



«Переходы» по дереву показаны на рис. 34, *а* (форма *prefix)* на рис. 34,*b* (форма *postfix)* и на рис. 34, *с* (форма *infix)* со скобками:

$$((a + (b\bullet c)) + (d + (e\bullet (g + g)))).$$

К этим вопросам мы возвратим позже.

Мы уже знакомы с многими бинарными операциями, например с арифметическими операциями +, •, —,/ и операциями над множествами — объединением (☐) и пересечением (I).

Операции, которые определены на конечных множествах, часто удобнее задавать с помощью таблиц.

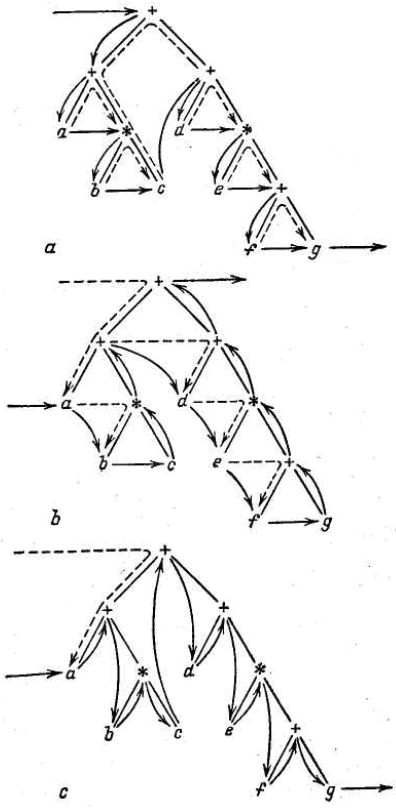


Рис. 34.

Пример 41. Пусть операция $\otimes$ определена на множестве $\{a, b, c\}$ с помощью таблицы

$\otimes$	a	b	c
a	a	a	b
b	b	a	c
c	a	b	b

Следовательно,

$$a \otimes b = a,$$
$$b \otimes b = a,$$
$$c \otimes b = b, \ldots$$

Такие символы, как $\otimes$ и $\oplus$, будут использоваться для обозначения разных операций, которые будут вводиться в процессе изложения.

Очевидно, что использование таблиц имеет важное значение, так как некоторые операции, с которыми приходится иметь дело в дискретной математике, непригодны для словесного задания.

Обратим теперь внимание на ***свойства операций***. **Операции вместе со своими следствиями обеспечивают основу всех алгебраических вопросов математики, так как они определяют порядок работы с объектами.**

Определение. Говорят, что бинарная операция $\otimes$ на множестве A *коммутативна,* если

$$a \otimes b = b \otimes a \quad \text{для всех} \quad a, b \in A.$$

Следовательно, обычная операция сложения на Z коммутативна, а вычитания — нет.

Определение. Говорят, что операция $\otimes$ на множестве A *ассоциативна,* если

$$(a \otimes b) \otimes c = a \otimes (b \otimes c) \quad \text{для всех} \quad a,b,c \in A.$$

Заметим, что в определении ассоциативности порядок операндов a, b и c сохранен (операция может быть некоммутативной!) и использованы круглые скобки, чтобы определить порядок вычислений.

Таким образом, выражение $(a \otimes b) \otimes c$ требует, чтобы сначала вычислялось $a \otimes b$ и результат этого (скажем, x) принимал участие в операции с c, т.е. давал $x \otimes c$. Если операция ассоциативна, то порядок вычислений несуществен и, следовательно, скобки не требуются.

Пример 42. Над Z имеем

$$(1+2)+3 = 1+2+3 = 1+(2+3),$$

но

$$(1-2)-3 = -4 \quad \text{и} \quad 1-(2-3) = 2.$$

Таким образом, операция вычитания не ассоциативна.

Коммутативность и ассоциативность являются двумя важными свойствами, которые могут быть определены для простых операций. Перед тем как описывать свойства, которые связывают две операции, определим некоторые термины, относящиеся к специальным элементам множеств, к которым эти операции применяются.

Определение. Пусть $\otimes$ — бинарная операция на множестве A и $l \in A$ такая, что

$$l \otimes a = a \quad \text{для всех} \quad a \in A.$$

Тогда l называется *левой единицей* относительно $\otimes$ на A. Аналогично, если существует $r \in A$ такое, что

$$r \otimes a = a \text{ для всех } a \in A,$$

то r является *правой единицей* относительно $\otimes$. Далее, если существует элемент e, который является и левой, и правой единицей, т.е.

$$e \otimes a = a \otimes e = a \text{ для всех } a \in A,$$

то e называется *(двусторонней) единицей* по отношению к $\otimes$.

Пример 43. Над R 0 является правой единицей по отношению к вычитанию и единицей по отношению к сложению, так как

$$a — 0 = a,$$

но

$$0 — a \neq a, \text{ если } a \neq 0;$$
$$a + 0 = a \text{ и } 0 + a = a \text{ для всех } a.$$

Определение. Пусть $\otimes$ — операция на A с единицей e и $x \otimes y = c$. Тогда говорят, что x — *левый обратный* элемент к y, а y — *правый обратный* элемент к x. Дальше, если x и y такие, что

$$x \otimes y = e = y \otimes x,$$

это y называется *обратным элементом* к x по отношению к $\otimes$, и наоборот.

Замечание. В некоторых книгах левые (правые) обратные элементы относят к левой (правой) единицы, однако, как мы в скором времени увидим, в большинстве случаев единицы являются двусторонними и, следовательно, не требуется делать никаких различий. Для решения уравнений необходимо существование и единственность единиц и обратных элементов. Менее общим свойством операций является идемпотентнисть, хотя оно используется в алгебре логики.

Определение. Пусть операция $\otimes$ на множестве A и произвольный элемент $x \in A$ таковы, что $x \otimes x = x$. Тогда говорят, что x *идемпотентен* по отношению к $\otimes$.

Очевидно, что любое подмножество идемпотентно по отношению к операциям пересечения и объединения.

Определение. Пусть дано множество A, на котором определены две операции $\otimes$ и $\oplus$. Тогда, если

$$a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c) \text{ для всех } a,b,c \in A,$$

то говорят, что $\otimes$ *дистрибутивна* по отношению к $\oplus$.

Если сказанное выше не совсем понятно, следует провести соответствие между этим тождеством и обычной арифметикой на R, например,

$$3*(1 + 2) = (3*1)+(3*2).$$

Наиболее общеизвестная алгебра может быть построена из относительно небольшого набора основных правил. Сейчас мы продемонстрируем, как из элементарных предположений можно извлечь некоторые простые следствия; большинство примеров дано в виде упражнений.

Пример 44. Пусть $\otimes$ — операция на множестве A и существует единица по отношению к $\otimes$. Тогда единичный элемент единствен.

Доказательство. Предположим, что x и y — единицы по отношению к $\otimes$, т.е.

$$x \otimes a = a \otimes x = a,$$
$$y \otimes a = a \otimes y = a \quad \text{для всех } a \in A.$$

Тогда $x = x \otimes y$, так как y — единица, и $x \otimes y = y$, поскольку x — единица. Итак, $x = y$.

Пример 45. Пусть $\otimes$ — ассоциативная операция на множестве A и e — единица по отношению к $\otimes$. Тогда если $a \in A$ и x имеет обратный, то обратный элемент единствен по отношению к $\otimes$.

Доказательство. Допустим, что x' и x'' - обратные элементы к x, так что

$$x \otimes x' = x' \otimes x = e \text{ и } x \otimes x'' = x'' \otimes x = e.$$

Тогда

$$x' = x' \otimes e = x' \otimes (x \otimes x'') = (x' \otimes x) \otimes x'' = e \otimes x'' = x''.$$

2.5. Числа и последовательности Фибоначчи

Древняя история богата выдающимися математиками. Многие достижения древней математической науки до сих пор вызывают восхищение остротой ума их авторов, а имена Евклида, Архимеда, Герона известны каждому образованному человеку.

Иначе обстоит дело с математикой средневековья. Кроме Виеты, жившего, впрочем, уже в шестнадцатом столетии, и математиков более близких нам времен школьный курс математики не называет ни одного имени, относящегося к средним векам. Это, конечно, не случайно. Математика в эту этоху развивалась чрезвычайно медленно, и крупных математиков тогда было очень мало.

Тем больший интерес представляет для нас сочинение «Liber abaccі» («Книга об абаке»), написанная знаменитым итальянским математиком Леонардо из Пизы, который известен больше по своему прозвищу Фибоначчи (Fibonacci — сокращенное filius Bonacci, т. е. сын Боначчи). Эта книга, написанная в 1202 г., дошла до нас во втором своем варианте, который относится к 1228 г.

«Liber abaccі» представляет собой объемистый труд, содержащий почти все арифметические и алгебраические сведения того времени и сыгравший заметную роль в развитии математики в Западной Европе в течение нескольких следующих столетий. В частности, именно по этой книге европейцы познакомились с индусскими («арабскими») цифрами.

Сообщаемый в «Liber abaccі» материал поясняется на большом числе задач, составляющих значительную часть этого трактата.

Рассмотрим одну такую задачу, помещенную на стр. 123—124 рукописи 1228 г.

«Сколько пар кроликов в один год от одной пары рождается?»

«Некто поместил пару кроликов в некоем месте, огороженном со всех сторон стеной, чтобы узнать, сколько пар кроликов родится при этом в течение года, если природа кроликов такова, что через месяц пара кроликов производит на свет другую пару, а рождают кролики со второго месяца после своего рождения. Так как первая пара в первом месяце дает потомство, удвой, и в этом месяце окажутся 2 пары; из них одна пара, а именно первая, рождает и в следующем месяце, так что во втором месяце оказывается 3 пары; из них в следующем месяце 2 пары будут давать потомство, так что в третьем месяце родятся еще 2 пары кроликов, и число пар кроликов в этом месяце достигнет 5; из них в этом же месяце будут давать потомство 3 пары, и число пар кроликов в четвергом месяце достигнет 8; из них 5 пар произведут другие 5 пар, которые, сложенные с 8 парами, дадут в пятом месяце 13 пар; из них 5 пар, рожденных в этом месяце, не дают в том же месяце потомства, а остальные 8 пар рождают, так что в шестом месяце оказывается 21 пара; сложенные с 13 парами, которые родятся в седьмом месяце, они дают 34 пары; сложенные с 21 парой, рожденной в восьмом месяце, они дают в этом месяце 55 пар; сложенные с 34 парами, рожденными в девятом месяце, они дают 89 пар; сложенные вновь с 55 парами, которые рождаются в десятом месяце, они дают в этом месяце 144 пары; снова сложенные с 89 парами, которые рождаются в одиннадцатом месяце, они дают в этом месяце 233 пары; сложен-ные вновь с 144 парами, рожденными в последнем месяце, они дают 377 пар; столько пар произвела первая пара в данном месте к концу одного года. Действительно, на этих полях ты можешь увидеть, как мы это

делаем; именно, мы складываем первое число со вторым, т. е. 1 и 2; и второе с третьим; и третье с четвертым; и четвертое с пятым; и так одно за другим, пока не сложим десятое с одиннадцатым, т. е. 144 с 233; и мы получим общее число упомянутых кроликов, т. е. 377; и так можно делать по порядку до бесконечного числа месяцев».

Пара	Первый	Второй	Третий	Четвертый	Пятый	Шестой	Седьмой	Восьмой	Девятый	Десятый	Одиннадцатый	Двенадцатый
1	2	3	8	8	13	21	31	55	89	144	233	377

Перейдем теперь от кроликов к числам и рассмотрим следующую числовую последовательность:

$$u_1, u_2, ..., u_n, \qquad (75)$$

в которой каждый член равен сумме двух предыдущих членов, т. е. при всяком $n > 2$

$$u_n = u_{n-1}+u_{n-2}. \qquad (76)$$

Такие последовательности, в которых каждый член определяется как некоторая функция предыдущих, часто встречаются в математике и называются *рекуррентными* или, по-русски, *возвратными* последовательностями. **Сам процесс последовательного определения элементов таких последовательностей называется *рекуррентным процессом,* а равенство (76) — *возвратным (рекуррентным) уравнением.***

Заметим прежде всего, что по одному только условию (76) члены последовательности (75) вычислять нельзя. Можно составить сколько угодно различных числовых последовательностей, удовлетворяющих этому условию; например,

2, 5, 7, 12, 19, 31, 50, ...,

1, 3, 4, 7, 11, 18, 29.....

— 1, -5, -6, —11, -17, ... и т. д.

Значит, для однозначного построения последовательности (75) условия (76) явно недостаточно, и нам следует указать некоторые дополнительные условия. Например, мы можем задать несколько первых членов последовательности (75). Сколько же первых членов последовательности (75) мы должны задать, чтобы можно было вычислять все следующие ее члены, пользуясь при этом только условием (76)?

Начнем с того, что не всякий член последовательности (75) может быть получен при помощи (76) уже хотя бы потому, что не у каждого члена (75) имеется два предшествующих; например, перед первым чле-

ном последовательности вообще не стоит ни одного члена, а перед вторым ее членом стоит только один. Значит, вместе с условием (76) для определения последовательности (75) нам нужно знать два ее первых члена.

Этого, очевидно, уже достаточно для того, чтобы иметь возможность вычислить любой член последовательности (75). В самом деле, u_3 можно вычислить как сумму заданных нам u_1 и u_2; u_4 — как сумму u_2 и уже вычисленного ранее u_3; u_5— как сумму уже вычисленных u_3 и u_4 и т. д. «по порядку до бесконечного числа членов». Переходя таким образом от двух соседних членов последовательности к непосредственно следующему за ними члену, мы можем дойти до члена с любым наперед заданным номером и вычислить его.

Обратимся теперь к важному частному случаю последовательности (75), когда $u_1 = 1$ и $u_2 = 1$. Условие (76), как было только что отмечено, дает нам возможность вычислять последовательно один за другим все члены этого ряда. Нетрудно проверить, что в этом случае первыми четырнадцатью его членами будут числа

1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377,

которые уже встречались нам в задаче о кроликах.

В честь автора этой задачи вся последовательность (75) при $u_1 = u_2 = 1$ называется *рядом Фибоначчи,* а члены ее — *числами Фибоначчи.*

Числа Фибоначчи обладают целым рядом интересных и важных свойств, простейшие из которых будут рассмотрены ниже

1. Вычислим сначала сумму первых n чисел Фибоначчи. Именно, докажем, что

$$u_1 + u_2 + \ldots + u_n = u_{n+2} - 1. \qquad (77)$$

В самом деле, мы имеем:

$$\begin{aligned} u_1 &= u_3 - u_2, \\ u_2 &= u_4 - u_3, \\ u_3 &= u_5 - u_4, \\ &\cdots\cdots \\ u_{n-1} &= u_{n+1} - u_n, \\ u_n &= u_{n+2} - u_{n+1}. \end{aligned}$$

Сложив все эти рааенства почленно, мы получим

$$u_1 + u_2 + \ldots + u_n = u_{n+2} - u_2,$$

и нам остается вспомнить, что $u_2 = 1$.

2. Сумма чисел Фибоначчи с нечетными номерами:

$$u_1 + u_3 + u_5 + \ldots + u_{2n-1} = u_{2n}. \qquad (78)$$

Для доказательства этого равенства напишем

$$u_1 = u_2,$$
$$u_3 = u_4 - u_2,$$
$$u_5 = u_6 - u_4,$$
$$\ldots\ldots$$
$$u_{2n-1} = u_{2n} - u_{2n-2}.$$

Сложив эти равенства почленно, мы и получим требуемое.

3. Сумма чисел Фибоначчи с четными номерами:

$$u_2 + u_4 + \ldots + u_{2n} = u_{2n+1} - 1. \qquad (79)$$

На основании п. 1 мы имеем

$$u_1 + u_2 + u_3 + \ldots + u_{2n} = u_{2n+2} - 1;$$

вычтя почленно из этого равенства равенство (1.178), мы получим

$$u_2 + u_4 + \ldots + u_{2n} = u_{2n+2} - 1 - u_{2n} = u_{2n+1} - 1,$$

а это нам и требовалось.

Вычитая, далее, почленно (79) из (78), получаем

$$u_1 - u_2 + u_3 - u_4 + \ldots + u_{2n-1} - u_{2n} = -u_{2n-1} + 1. \qquad (80)$$

Прибавим теперь к обеим частям (80) по u_{2n+1}:

$$u_1 - u_2 + u_3 - u_4 + \ldots - u_{2n} + u_{2n+1} = u_{2n} + 1. \qquad (81)$$

Объединяя (80) и (81), получаем выражение для знакопеременной суммы чисел Фибоначчи:

$$u_1 - u_2 + u_3 - u_4 + \ldots + (-1)^{n+1} u_n = (-1)^{n+1} u_{n-1} + 1. \qquad (82)$$

4. Формулы (77) и (78) были выведены при помощи почленного сложения целой серии очевидных равенств. Еще одним примером применения этого приема может служить вывод формулы для суммы квадратов первых n чисел Фибоначчи:

$$u_1^2 + u_2^2 + \ldots + u_n^2 = u_n u_{n+1}. \qquad (83)$$

Заметим для этого, что

$$u_k u_{k+1} - u_{k-1} u_k = u_k (u_{k+1} - u_{k-1}) = u_k^2.$$

Сложив равенства

$$u_1^2 = u_1 u_2,$$
$$u_2^2 = u_2 u_3 - u_1 u_2,$$
$$u_3^2 = u_3 u_4 - u_2 u_3,$$
$$\cdots\cdots\cdots\cdots$$
$$u_n^2 = u_n u_{n+1} - u_{n-1} u_n$$

почленно, мы получаем формулу (83).

5. Многие соотношения между числами Фибоначчи удобно доказывать при помощи метода *полной индукции.*

Сущность метода полной индукции (называемого часто также методом математической индукции) состоит в следующем: для доказательства, что некоторое утверждение справедливо для всякого натурального числа, достаточно установить, что:

а) оно имеет место для числа 1;

б) из справедливости доказываемого утверждения для какого-либо произвольно выбранного натурального числа n следует его справедливость для числа $u + 1$.

Всякое индуктивное доказательство утверждения, справедливого для любого натурального числа, состоит поэтому из двух частей.

В первой части (обычно сравнительно простой) устанавливается справедливость доказываемого утверждения для единицы. Справедливость доказываемого утверждения для единицы называют иногда *основанием индукции.* Во второй части доказательства (как правило, более сложной) делается предположение о справедливости доказываемого утверждения для некоторого произвольного (но фиксированного) числа n, и из этого предположения, которое часто называют *индуктивным предположением,* выводится, что и для числа n+1 доказываемое утверждение имеет место. Вторая часть доказательства называется *индуктивным переходом.*

Иногда применяется индуктивное рассуждение, которое можно назвать переходом «от всех чисел, меньших n, к n». При этом необходимость в специальном доказательстве основания индукции отпадает, так как, говоря формально, доказательство для случая $n = 1$ и есть переход от «всех» целых положительных чисел, меньших единицы (которых просто нет), к единице.

Именно таким является доказательство возможности разложения любого натурального числа на простые множители.

Предположим, что каждое из чисел, меньших некоторого n, разложимо в произведение простых множителей. Если число n оказывается простым, то оно само и является своим разложением. Если же число n составное, то его, по определению, можно представить в

виде произведения хотя бы двух сомножителей: $n = n_1 n_2$, где $n_1 \neq 1$ и $n_2 \neq 1$. Но тогда $n_1 < n$ и $n_2 < n$, а по индуктивному предположению как n_1, так и n_2 разлагаются на простые множители. Тем самым и n разложимо на простые множители.

6. Простейшей реализацией идеи индукции в применении к числам Фибоначчи является само определение чисел Фибоначчи. Оно, как разъяснялось выше, состоит в указании двух первых чисел Фибоначчи: $u_1 = 1$ и $u_2 = 1$ и в индуктивном переходе от u_n и u_{n+1} к u_{n+2}, даваемым рекуррентным соотношением

$$u_n + u_{n+1} = u_{n+2}.$$

В частности, отсюда автоматически следует, что если некоторая последовательность чисел начинается с двух единиц, а каждое из следующих получается сложением двух предыдущих, то эта последовательность является последовательностью чисел Фибоначчи.

В качестве примера рассмотрим так называемую «задачу о прыгуне». Она состоит в следующем.

Прыгун может прыгать в одном направлении вдоль разделенной на клетки полосы, перемещаясь при каждом прыжке либо в соседнюю клетку, либо через клетку. Сколькими способами может он сдвинуться на n— 1 клетку и, в частности, переместиться из первой клетки в n-ю? (Способы прыгания считаются одинаковыми, если в ходе каждого из них прыгун побывает в одних и тех же клетках.)

Обозначим искомое число через x_n. Очевидно, $x_1 = 1$ (ибо переход из первой клетки в первую же осуществляется только одним способом — отсутствием прыжков) и $x_2 = 1$ (переход из первой клетки во вторую также единствен: он состоит в одном непосредственном прыжке на соседнюю клетку). Пусть целью прыгуна является достижение n + 2-й клетки. Общее число способов осуществления этой цели в наших обозначениях равно x_{n+2}. Но с самого начала эти способы разбиваются на два класса: начинающиеся с прыжка во вторую клетку и начинающиеся с прыжка в третью клетку. Из второй клетки прыгун может переместиться в n+2-ю x_{n+1} способами, а из третьей x_n способами. Таким образом, последовательность чисел $x_1, x_2, \ldots, x_n, \ldots$ удовлетворяет рекуррентному соотношению

$$u_n + u_{n+1} = u_{n+2}$$

и поэтому совпадает с последовательностью чисел Фибоначчи: $x_n = u_n$.

7. Докажем по индукции следующую важную формулу:

$$u_{n+m} = u_{n-1}u_m + u_n u_{m+1}. \qquad (84)$$

Доказательство этой формулы будем вести индукцией по m. При $m = 1$ эта формула принимает вид $u_{n+1} = u_{n-1}u_n + u_n u_2$, что очевидно. При $m = 2$ формула (84) также верна, потому что

$$u_{n+2} = u_{n-1}u_2 + u_nu_3 = u_{n-1} + 2u_n =$$
$$= u_{n-1} + u_n + u_n = u_{n+1} + u_n.$$

Основание индукции, таким образом, доказано. Индуктивный переход докажем в следующей форме: предполагая, что формула (84) справедлива при $m = k$ и при $m = k+1$, докажем, что она имеет место и при $m = k + 2$.

Итак, пусть

$$u_{n+2} = u_{n+1} + u_n = u_{n-1} + u_n + u_n =$$
$$= u_{n-1} + 2u_n = u_{n-1}u_2 + u_nu_3.$$

Сложив последние два равенства почленно, мы получим

$$u_{n+k+2} = u_{n-1}u_{k+2} + u_nu_{k+3},$$

а это и требовалось.

Формулу (84) легко интерпретировать (и даже доказать) в терминах задачи о прыгуне.

Именно, общее число способов перемещения прыгуна из первой клетки в $n + m$-ю равно u_{n+m}. Среди этих способов будут как те, при которых прыгун перепрыгнет через n-ю клетку, так и те, при которых он побывает в ней.

При способах первого класса прыгун обязан достичь n—1-й клетки (он может сделать это u_{n-1} способами), затем совершить прыжок на n+1-ю клетку и, наконец, сместиться на оставшиеся $(n+m)—(n+1)=m—1$ клеток (это осуществимо u_m способами). Следовательно, первый класс насчитывает $u_{n-1}u_m$ способов. Аналогично, при способах второго класса прыгун достигает n-й клетки (это возможно u_n способами), после чего переходит в $n+m$-ю клетку (одним из u_{m+1} способов). Поэтому во втором классе имеется u_nu_{m+1} способов, и формула (84) доказана.

8. Полагая в формуле (84) $m = n$, мы получаем

$$u_{2n} = u_{n-1}u_n + u_nu_{n+1},$$

или

$$u_{2n} = u_n(u_{n-1} + u_{n+1}). \qquad (85)$$

Из написанного равенства видно, что u_{2n} делится на u_n.

Так как

$$u_n = u_{n+1} - u_{n-1},$$

формулу (85) можно переписать так:

$$u_{2n} = (u_{n+1} - u_{n-1})(u_{n+1} + u_{n-1}),$$

или

$$u_{2n} = u_{n+1}^2 - u_{n-1}^2,$$

т. е. разность квадратов двух чисел Фибоначчи, номера которых отличаются на два, есть снова число Фибоначчи.

Аналогично (полагая $m = 2n$) можно показать, что

$$u_{3n} = u_{n+1}^3 + u_n^3 - u_{n-1}^3.$$

9. В дальнейшем нам пригодится следующая формула:

$$u_n^2 = u_{n-1}u_{n+1} + (-1)^{n+1}. \qquad (86)$$

Докажем ее индукцией по n. Для $n = 2$ (1.186) принимает вид

$$u_2^2 = u_1u_3 - 1,$$

что очевидно.

Предположим тонерь формулу (86) доказанной для некоторого n. Прибавим к обеим частям ее по u_nu_{n+1}. Мы получим

$$u_n^2 + u_nu_{n+1} = u_{n-1}u_{n+1} + u_nu_{n+1} + (-1)^{n+1},$$

или

$$u_n(u_n + u_{n+1}) = u_{n+1}(u_{n-1} + u_n) + (-1)^{n+1},$$

или

$$u_nu_{n+2} = u_{n+1}^2 + (-1)^{n+1},$$

или

$$u_{n+1}^2 = u_nu_{n+2} + (-1)^{n+2}.$$

Этим индуктивный переход обоснован, и формула (86) доказана для любого n.

10. Аналогично только что доказанным свойствам чисел Фибоначчи можно установить еще и такие свойства этих чисел:

$$u_1u_2 + u_2u_3 + u_3u_4 + \ldots + u_{2n-1}u_{2n} = u_{2n}^2,$$

$$u_1u_2 + u_2u_3 + u_3u_4 + \ldots + u_{2n}u_{2n+1} = u_{2n+1}^2 - 1,$$

$$nu_1 + (n-1)u_2 + (n-2)u_3 + \ldots + 2u_{n-1} + u_n = u_{n+4} - (n+3),$$

$$u_1 + 2u_2 + 3u_3 + \ldots + nu_n = nu_{n+2} - u_{n+3} + 2.$$

Доказательство предоставляется провести читателю.

11. Не менее замечательными, чем числа Фибоначчи, являются другие числа, называемые *биномиальными коэффициентами.*

Биномиальными коэффициентами называются коэффициенты при степенях x в разложениях степеней $(1+x)^n$:

$$(1+x)^n = C_n^0 + C_n^1x + C_n^2x^2 + \ldots + C_n^nx^n. \qquad (87)$$

Очевидно, числа C^2_n однозначно определены при всех целых неотрицательных n и всех целых неотрицательных k, не превосходящих n.

Использование биномиальных коэффициентов оказывается весьма удобным во многих математических рассуждениях. Пригодятся они нам и при изучении свойств чисел Фибоначчи. Кроме того, биномиальные коэффициенты связаны с числами Фибоначчи и непосредственно, и мы выявим некоторые закономерности, связывающие эти два класса чисел.

Предварительно установим некоторые свойства биномиальных коэффициентов.

Положив в (87) $n = 1$, мы видим, что

$$C_1^0 = C_1^1 = 1;$$

кроме того, имеет место следующая лемма.

Лемма. $C_n^k + C_n^{k+1} = C_{n+1}^{k+1}$.

Доказательство. Мы имеем

$$(1 + x)^{n+1} = (1 + x)^n (1 + x),$$

или, пользуясь определением биномиальных коэффициентов,

$$C_{n+1}^0 + C_{n+1}^1 x + \ldots + C_{n+1}^{k+1} x^{k+1} + \ldots + C_{n+1}^{n+1} x^{n+1} =$$
$$= (C_n^0 + C_n^1 x + \ldots + C_n^k x^k + C_n^{k+1} x^{k+1} + \ldots$$
$$\ldots + C_n^n x^n)(1 + x) = C_n^0 + (C_n^0 + C_n^1) x + \ldots$$
$$\ldots + (C_n^k + C_n^{k+1}) x^{k+1} + \ldots + (C_n^{n-1} + C_n^n) x^n + C_n^n x^{n+1}.$$

Но слева и справа в этом равенстве стоит *один и тот же* полином. Поэтому и коэффициенты при одинаковвых степенях x слева и справа должны быть равны, В частности, должно быть

$$C_{n+1}^{k+1} = C_n^k + C_n^{k+1},$$

а это и требовалось.

Из доказанной леммы следует, что биномиальные коэффициенты можно вычислять при помощи некоторого рекуррентного процесса, подобного процессу получения чисел Фибоначчи, только значительно более сложной природы. Это же обстоятельство дает нам возможность доказывать по индукции разного рода утверждения о биномиальных коэффициентах.

12. Расположим биномиальные коэффициенты в виде следующей таблицы, называемой *треугольником Паскаля:*

$$
\begin{array}{l}
C_0^0 \\
C_1^0\ C_1^1 \\
C_2^0\ C_2^1\ C_2^2 \\
\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\
C_n^0\ C_n^1\ C_n^2\ \ldots\ C_n^n \\
\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot
\end{array}
$$

т. е.

$$
\begin{array}{l}
1 \\
1\ \ 1 \\
1\ \ 2\ \ 1 \\
1\ \ 3\ \ 3\ \ 1 \\
1\ \ 4\ \ 6\ \ 4\ \ 1 \\
1\ \ 5\ 10\ 10\ \ 5\ \ 1 \\
1\ \ 6\ 15\ 20\ 15\ \ 6\ \ 1 \\
\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot
\end{array}
$$

Строки треугольника Паскаля принято нумеровать сверху вниз, причем верхняя строка, состоящая из единственной единицы, считается нулевой.

Из предыдущего вытекает, что крайние члены в каждой из строк треугольника Паскаля равны единице, а каждый из остальных членов таблицы получаются путем сложения двух других, стоящих непосредственно над ним.

13. Формула (87) позволяет сразу вывести два важных соотношения, связывающих биномиальные коэффициенты, составляющие одну строку треугольника Паскаля.

Полагая в (87) $x = 1$, получаем

$$2^n = C_n^0 + C_n^1 + C_n^2 + \ \ldots\ + C_n^n.$$

Если же принять $x = -1$, то получим

$$0 = C_n^0 - C_n^1 + C_n^2 + \ \ldots\ + (-1)^n C_n^n.$$

14. Докажем индукцией по n, что

$$C_n^k = \frac{n(n-1)\ldots(n-k+1)}{1 \cdot 2 \cdot \ldots \cdot k}. \qquad (88)$$

Эта формула часто принимается за определение биномиальных коэффициентов. Она характеризует биномиальный коэффициент C_n^k как число сочетаний из n элементов по k. Мы пошли здесь по иному, более формальному пути, который в данном случае предпочтительнее.

Если согласиться считать, что произведение нулевого числа сомножителей всегда равно единице, то при k=0 из (88) получаем уже известное нам C_n^0=1. Имея это в виду, мы можем ограничиться случаем $k \geq 1$.

При n = 1 мы имеем

$$C_1^1 = \frac{1}{1} = 1.$$

Пусть теперь при некотором данном n формула (88) справедлива при любом значении k = 0, 1, ,...., n.

Рассмотрим число C_{n+1}^k. Так как $k \geq 1$, мы можем написать

$$C_{n+1}^k = C_n^{k-1} + C_n^k,$$

или, воспользовавшись индуктивным предположением (88),

$$C_n^{k-1} + C_n^k =$$

$$= \frac{n(n-1)\ldots(n-k+2)}{1\cdot 2\cdot \ldots \cdot (k-1)} + \frac{n(n-1)\ldots(n-k+2)(n-k+1)}{1\cdot 2\cdot \ldots \cdot (k-1)\,k} =$$

$$= \frac{n(n-1)\ldots(n-k+2)}{1\cdot 2\cdot \ldots \cdot (k-1)} \left(1 + \frac{n-k+1}{k}\right) =$$

$$= \frac{n(n-1)\ldots(n-k+2)}{1\cdot 2\cdot \ldots \cdot (k-1)} \cdot \frac{k+n-k+1}{k} =$$

$$= \frac{(n+1)\,n(n-1)\ldots(n-k+2)}{1\cdot 2\cdot \ldots \cdot (k-1)\,k} = C_{n+1}^k.$$

Последнее равенство является формулой (88) для биномиальных коэффициентов из следующей, n+1-й строки треугольника Паскаля.

15. Проведем через числа треугольника Паскаля линии, идущие под углом 45 градусов к его строкам, и назовем их *восходящими диагоналями* треугольника Паскаля. Восходящими диагоналями будут, например, прямые, проходящие через числи 1, 4, 3 или 1, 5, 6, 1.

Покажем, что сумма чисел, лежащих на некоторой восходящей диагонали, есть число Фибоначчи.

В самом деле, первая, самая верхняя восходящая диагональ треугольника Паскаля состоит только из единицы. Только из единицы состоит и вторая его диагональ. Для доказательство интересующего нас предложения достаточно показать, что сумма всех чисел, составляющих n-ю и n+1-ю диагонали треугольника Паскаля, равна сумме чисел, составляющих его n + 2-ю диагональ.

Но на n-й диагонали расположены числа

$$C_{n-1}^0,\ C_{n-2}^1,\ C_{n-3}^2, \ldots,$$

а на n + 1-й — числа

$$C_n^0,\ C_{n-1}^1,\ C_{n-2}^2, \ldots$$

Сумму всех этих чисел запишем так:

$$C_n^0 + \left(C_{n-1}^0 + C_{n-1}^1\right) + \left(C_{n-2}^1 + C_{n-2}^2\right) + \ldots,$$

или, принимая во внимание лемму в п. 11,

$$C_{n+1}^0 + C_n^1 + C_{n-1}^2 + \ldots$$

Последнее выражение есть сумма чисел, лежащих на n + 2-й восходящей диагонали треугольника.

Из только что доказанного на основании формулы (77) мы получаем: сумма всех биномиальных коэффициентов, лежащих выше n-й восходящей диагонали треугольника Паскаля (включая саму эту диагональ), равна u_{n+2}— 1.

Используя формулы (78), (79), (80) и подобные им, читатель без труда может получить дальнейшие тождества, связывающие числа Фибоначчи с биномиальными коэффициентами.

16. До сих пор мы определяли число Фибоначчи рскуррентно, т. е. индуктивно, по их номеру. Оказывастся, однако, что любое число Фибоначчи можно определить и непосредственно, как некоторую функцию его номера.

Исследуем для этого различные последовательности $u_1, u_2, \ldots, u_n, \ldots$, удовлетворяющие соотношению

$$u_n = u_{n-2} + u_{n-1}. \tag{89}$$

Все такие последовательности мы будем называть *решениями уравнения* (89).

Будем обозначать буквами V, V и V'' соответственно последовательности

$$v_1, \quad v_2, \quad v_3, \quad \ldots$$
$$v_1', \quad v_2', \quad v_3', \quad \ldots$$
$$v_1'', \quad v_2'', \quad v_3'', \quad \ldots$$

Сначала докажем две леммы.

Лемма 1. *Если V есть решение уравнения* (89), *а c — произвольное число, то последовательность cV (т. е. последовательность $cv_1, cv_2, cv_3, \ldots$) есть также решение уравнения* (89).

Доказательство. Умножив соотношение

$$v_n = v_{n-2} + v_{n-1}$$

почленно на c, мы получаем

$$cv_n = cv_{n-2} + cv_{n-1},$$

а это и требовалось.

Лемма 2. *Если последовательности V и V'' являются решениями уравнения* (89), *то и их сумма $V + V''$ (т. е. последовательность*

$(v'_1 + v''_1,\ v'_2 + v''_2,\ v'_3 + v''_3,\ \ldots)$ *также является решением уравнения* (89).

Доказательство. Из условия леммы мы имеем

$$v'_n = v'_{n-2} + v'_{n-1}$$

и

$$v''_n = v''_{n-2} + v''_{n-1}.$$

Сложив эти два равенства почленно, мы получим

$$v'_n + v''_n = (v'_{n-2} + v''_{n-2}) + (v'_{n-1} + v''_{n-1}).$$

Этим лемма доказана.

Пусть теперь V и V'' — два непропорциональных решения уравнения (89) (т. е. два таких решения уравнения (89), что при любом постоянном c найдется такой номер n, для которого $\frac{v'_n}{v''_n} \neq c$). Покажем, что всякую последовательность V, являющуюся решением уравнения (89), можно представить в виде

$$c_1 V' + c_2 V'', \tag{90}$$

где c_1 и c_2 — некоторые постоянные. Поэтому принято говорить, что (90) является *общим решением* уравнения (89).

Предварительно докажем, что если решения V и V'' уравнения (89) непропорциональны, то

$$\frac{v'_1}{v''_1} \neq \frac{v'_2}{v''_2} \tag{91}$$

(т. е. что эта непропорциональность обнаруживается уже в первых двух членах последовательностей V и V'').

Доказательство (91) ведется от противного. Пусть для непропорциональных решений V и V'' уравнения (89)

$$\frac{v'_1}{v''_1} = \frac{v'_2}{v''_2}. \tag{92}$$

Написав производную пропорцию, мы получаем

$$\frac{v'_1 + v'_2}{v''_1 + v''_2} = \frac{v'_2}{v''_2}$$

или, принимая, во внимание, что V и V'' являются решениями уравнения (89),

$$\frac{v'_3}{v''_3} = \frac{v'_2}{v''_2}.$$

Аналогично убеждаемся (индукция!) в том, что

$$\frac{v_1'}{v_1''} = \frac{v_2'}{v_2''} = \ldots = \frac{v_n'}{v_n''} = \ldots$$

Таким образом, из (92) следует, что последовательности V и V'' пропорциональны, а это противоречит предположению. Значит, справедливо (91).

Возьмем теперь некоторую последовательность V, являющуюся решением уравнения (89). Эта последовательность, как уже было выяснено ранее, вполне определена, если заданы два ее первых члена, v_1 и v_2.

Найдем такие c_1 и c_2, чтобы имело место

$$\begin{aligned} c_1 v_1' + c_2 v_1'' &= v_1, \\ c_1 v_2' + c_2 v_2'' &= v_2. \end{aligned} \tag{93}$$

Тогда на основании лемм 1 и 2 $c_1V'+c_2V''$ даст нам последовательность V.

Ввиду условия (91) система уравнений (93) разрешима относительно c_1 и c_2, каковы бы ни были числа v_1 и v_2:

$$c_1 = \frac{v_1 v_2'' - v_2 v_1''}{v_1' v_2'' - v_1'' v_2'}, \qquad c_2 = \frac{v_1' v_2 - v_2' v_1}{v_1' v_2'' - v_1'' v_2'}.$$

(Условие (91) означает, что общий знаменатель этих дробей отличен от нуля.) Подставив вычисленные значения c_1 и c_2 в (90), мы и получим требуемое представление последовательности V.

Значит, для описания *всех* решений уравнения (89) нам достаточно найти *какие-нибудь два* его непропорциональных решения.

Будем искать эти решения среди геометрических прогрессий. В соответствии с леммой 1 достаточно ограничиться рассмотрением только таких прогрессий, у которых первый член равен единице. Итак, возьмем прогрессию

$$1, q, q^2, \ldots$$

Чтобы она была решением уравнения (89), необходимо, чтобы при всяком n выполнялось

$$q^{n-2} + q^{n-1} = q^n,$$

или, сокращая на q^{n-2},

$$1+q = q^2. \tag{94}$$

Корни этого квадратного уравнения, т. е. $\frac{1+\sqrt{5}}{2}$ и $\frac{1-\sqrt{5}}{2}$, и будут искомыми знаменателями прогрессий. Мы будем их обозначать соответственно через α и β. Подчеркнем, что

для чисел α и β, как для корней уравнения (94), должно иметь место

$$1 + \alpha = \alpha^2,\ 1 + \beta = \beta^2 \text{ и } \alpha\beta = -1.$$

Мы получили, таким образом, две геометрические прогрессии, являющиеся решениями уравнения (89). Поэтому все последовательности вида

$$c_1 + c_2,\quad c_1\alpha + c_2\beta,\quad c_1\alpha^2 + c_2\beta^2,\ \ldots \tag{95}$$

являются решениями уравнения (1.189). Так как найденные прогрессии имеют разные знаменатели и потому непропорциональны, формула (95) при различных c_1 и c_2 дает все решения уравнения (89).

В частности, при некоторых c_1 и c_2 формула (95) должна дать и ряд Фибоначчи. Для этого, как указывалось выше, нужно определить c_1 и c_2 из уравнений

$$c_1 + c_2 = u_1$$

и

$$c_1\alpha + c_2\beta = u_2,$$

т. е. из системы

$$c_1 + c_2 = 1,$$
$$c_1\frac{1+\sqrt{5}}{2} + c_2\frac{1-\sqrt{5}}{2} = 1.$$

Решив эту систему, мы получаем

$$c_1 = \frac{1+\sqrt{5}}{2\sqrt{5}},\qquad c_2 = -\frac{1-\sqrt{5}}{2\sqrt{5}},$$

откуда

$$u_n = c_1\alpha^{n-1} + c_2\beta^{n-1} =$$
$$= \frac{1+\sqrt{5}}{2\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n-1} - \frac{1-\sqrt{5}}{2\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{n-1},$$

т. о.

$$u_n = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n}{\sqrt{5}}. \tag{96}$$

Формула (96) называется *формулой Бине* (по имени математика, который ее вывел). Очевидно, подобные формулы можно указать и для других решений (89).

17. Мы видели, что $\alpha^2 = \alpha + 1$. Ясно поэтому, что любую целую положительную степень числа α можно представить в виде $a\alpha + b$ с целыми коэффициентами a и b. Так,

$$\alpha^3 = \alpha\alpha^2 = \alpha(\alpha+1) = \alpha^2 + \alpha = \alpha + 1 + \alpha = 2\alpha + 1,$$
$$\alpha^4 = \alpha\alpha^3 = \alpha(2\alpha+1) = 2\alpha^2 + \alpha = 2\alpha + 2 + \alpha = 3\alpha + 2$$

и т. д.

Покажем (по индукции), что

$$\alpha^n = u_n\alpha + u_{n-1}.$$

Действительно, для n=2, 3 это справедливо. Предположим, что

$$\alpha^k = u_k\alpha + u_{k-1},$$
$$\alpha^{k+1} = u_{k+1}\alpha + u_k.$$

Сложив эти равенства, получим

$$\alpha^k + \alpha^{k+1} = (u_k + u_{k+1})\alpha + (u_{k-1} + u_k),$$

или

$$\alpha^{k+2} = u_{k+2}\alpha + u_{k+1},$$

и индуктивный переход обоснован. Аналогично из $\beta^2 = \beta + 1$ следует

$$\beta^n = u_n\beta + u_{n-1}.$$

18. При помощи формулы Бине удобно суммировать многие ряды, связанные с числами Фибоначчи.

Найдем, например, чему равна сумма

$$u_3 + u_6 + u_9 + \ldots + u_{3n}.$$

Мы имеем

$$u_3 + u_6 + \ldots + u_{3n} = \frac{\alpha^3 - \beta^3}{\sqrt{5}} + \frac{\alpha^6 - \beta^6}{\sqrt{5}} + \ldots + \frac{\alpha^{3n} - \beta^{3n}}{\sqrt{5}} =$$
$$= \frac{1}{\sqrt{5}}(\alpha^3 + \alpha^6 + \ldots + \alpha^{3n} - \beta^3 - \beta^6 - \ldots - \beta^{3n})$$

или, суммируя встретившиеся нам геометрические прогрессии,

$$u_3 + u_6 + \ldots + u_{3n} = \frac{1}{\sqrt{5}}\left(\frac{\alpha^{3n+3} - \alpha^3}{\alpha^3 - 1} - \frac{\beta^{3n+3} - \beta^3}{\beta^3 - 1}\right).$$

Но

$$\alpha^3 - 1 = \alpha + \alpha^2 - 1 = \alpha + \alpha + 1 - 1 = 2\alpha,$$

и аналогично $\beta^2 - 1 = 2\beta$. Поэтому

$$u_3 + u_6 + \ldots + u_{3n} = \frac{1}{\sqrt{5}}\left(\frac{\alpha^{3n+3} - \alpha^3}{2\alpha} - \frac{\beta^{3n+3} - \beta^3}{2\beta}\right),$$

или, произведя сокращения,

$$u_3 + u_6 + \ldots + u_{3n} = \frac{1}{\sqrt{5}}\left(\frac{\alpha^{3n+2} - \alpha^2 - \beta^{3n+2} + \beta^2}{2}\right) =$$
$$= \frac{1}{2}\left(\frac{\alpha^{3n+2} - \beta^{3n+2}}{\sqrt{5}} - \frac{\alpha^2 - \beta^2}{\sqrt{5}}\right) = \frac{1}{2}(u_{3n+2} - u_2) = \frac{u_{3n+2} - 1}{2}$$

19. В качестве следующего примера применения формулы Бине вычислим сумму кубов первых n чисел Фибоначчи.

Заметим предварительно, что

$$u_k^3 = \left(\frac{\alpha^k - \beta^k}{\sqrt{5}}\right)^3 = \frac{1}{5}\,\frac{\alpha^{3k} - 3\alpha^{2k}\beta^k + 3\alpha^k\beta^{2k} - \beta^{3k}}{\sqrt{5}} =$$

$$= \frac{1}{5}\left(\frac{\alpha^{3k} - \beta^{3k}}{\sqrt{5}} - 3\alpha^k\beta^k\,\frac{\alpha^k - \beta^k}{\sqrt{5}}\right) =$$

$$= \frac{1}{5}\left(u_{3k} - (-1)^k\, 3u_k\right) = \frac{1}{5}\left(u_{3k} + (-1)^{k+1}\, 3u_k\right).$$

Поэтому

$$u_1^3 + u_2^3 + \ldots + u_n^3 =$$

$$= \frac{1}{5}\left((u_3 + u_6 + \ldots + u_{3n}) + 3\left(u_1 - u_2 + u_3 - \ldots + (-1)^{n+1} u_n\right)\right),$$

или, пользуясь формулой (1.182) и результатами предыдущего пункта,

$$u_1^3 + u_2^3 + \ldots + u_n^3 = \frac{1}{5}\left(\frac{u_{3n+2} - 1}{2} + (-1)^{n+1}\, 3u_{n-1} + 3\right) =$$

$$= \frac{u_{3n+2} + (-1)^{n+1}\, 6u_{n-1} + 5}{10}.$$

20. Поставим вопрос о том, как быстро растут числа Фибоначчи при увеличении их номеров. Формула Бине дает достаточно исчерпывающий овет и на этот вопрос.

Докажем следующую теорему.

Теорема. *Число Фибоначчи u_n есть ближайшее целое число к $\frac{\alpha^n}{\sqrt{5}}$, т. е. к n-му члену a_n геометрической прогрессии, первый член которой есть*

$$\frac{\alpha}{\sqrt{5}},$$

а знаменатель равен а.

Доказательство. Очевидно, достаточно установить, что абсолютная величина разности между u_n и a_n всегда меньше $\frac{1}{2}$. Но

$$|u_n - a_n| = \left|\frac{\alpha^n - \beta^n}{\sqrt{5}} - \frac{\alpha^n}{\sqrt{5}}\right| = \left|\frac{\alpha^n - \alpha^n - \beta^n}{\sqrt{5}}\right| = \frac{|\beta|^n}{\sqrt{5}}.$$

Так как $\beta = -0{,}618...$, то $|\beta| < 1$, а значит, $|\beta|^n < 1$ при любом n и тем более (так как $\sqrt{5} > 2$) должно быть $\frac{|\beta|^n}{\sqrt{5}} < \frac{1}{2}$. Теорема доказана.

Используя теорию пределов, легко сможет, несколько видоизменив доказательство этой теоремы, показать, что

$$\lim_{n\to\infty} |u_n - a_n| = 0.$$

Пользуясь доказанной теоремой, можно вычислять числа Фибоначчи при помощи таблиц логарифмов.

Вычислим, например, u_{14} (u_{14}, как легко сообразить, должно являться ответом задачи Фибоначчи о кроликах):

$$\sqrt{5} = 2{,}2361, \qquad \lg\sqrt{5} = 0{,}34949;$$

$$\alpha = \frac{1+\sqrt{5}}{2} = 1{,}6180, \qquad \lg\alpha = 0{,}20898;$$

$$\lg\frac{\alpha^{14}}{\sqrt{5}} = 14 \cdot 0{,}20898 - 0{,}34949 = 2{,}5762,$$

$$\frac{\alpha^{14}}{\sqrt{5}} = 376{,}9.$$

Ближайшим целым числом к 376,9 является 377; это и есть u_{14}.

При вычислении чисел Фибоначчи с большими номерами мы уже не сможем по таблицам логарифмов определить все цифры числа, а сможем указать только несколько первых цифр его, так что вычисление оказывается приближенным.

В виде упражнения читатель может доказать, что в десятичной системе счисления u_n при $n \geq 17$ имеет не более $\frac{1}{4}$ и не менее $\frac{n}{5}$ цифр. А из скольких цифр состоит u_{100}?

21. Результат предыдущего пункта можно уточнить. Следующая теорема пригодится нам в дальнейшем.

Теорема.

$$\frac{\alpha^{n-\frac{1}{n}}}{\sqrt{5}} \leqq u_n \leqq \frac{\alpha^{n+\frac{1}{n}}}{\sqrt{5}}.$$

Доказательство. Мы ограничимся доказательством левой стороны неравенства: правая доказывается аналогично. Поскольку согласно формуле Бине

$$u_n = \frac{1}{\sqrt{5}}(\alpha^n - \beta^n),$$

а $\alpha\beta = -1$, для наших целей будет достаточно показать, что

$$\alpha^{n-\frac{1}{n}} \leqq \alpha^n - \frac{1}{\alpha^n},$$

или

$$\alpha^{2n-\frac{1}{n}} \leqq \alpha^{2n} - 1,$$

или, возводя в степень n,

$$\alpha^{2n^2-1} \leqq (\alpha^{2n} - 1)^n. \qquad (97)$$

Будем доказывать это неравенство по индукции. При $n = 1$ оно превращается в

$$\alpha \leqq \alpha^2 - 1,$$

что действительно имеет место (именно, со знаком равенства). При $n = 2$ (97) означает

$$\alpha^7 \leqq (\alpha^4 - 1)^2. \qquad (98)$$

Это неравенство можно проверить и прямым вычислением. Однако его можно и доказать, воспользовавшись соотношением, выведенным в п. 17. В данном случае мы имеем

$$\alpha^4 = 3\alpha + 2,$$

$$(\alpha^4 - 1)^2 = (3\alpha + 1)^2 = 9\alpha^2 + 6\alpha + 1 = 15\alpha + 10,$$

и (1.198) переписывается как

$$\alpha^7 = 13\alpha + 8 \leqq 15\alpha + 10,$$

что очевидно. Наконец, при $n = 3$ (97) переписывается как

$$\alpha^{17} \leqq (\alpha^6 - 1)^3,$$

что проверяется аналогично предыдущему.

Предположим теперь, что $n>2$ и (97) имеет место, и докажем, что

$$\alpha^{2(n+1)^2-1} \leqq (\alpha^{2n+2} - 1)^{n+1}.$$

Для этого достаточно показать, что при увеличении n на единицу правая часть (97) растет быстрее левой части. Но левая часть, очевидно, возрастает в α^{4n+2} раз. Оценим увеличение правой части.

Мы имеем

$$\frac{(\alpha^{2(n+1)} - 1)^{n+1}}{(\alpha^{2n} - 1)^n} = (\alpha^{2(n+1)} - 1)\left(\frac{\alpha^{2(n+1)} - 1}{\alpha^{2n} - 1}\right)^n.$$

Последняя дробь больше, чем α^2, и притом на

$$\frac{\alpha^{2(n+1)}-1}{\alpha^{2n}-1}-\alpha^2=\frac{\alpha^{2n+2}-1-\alpha^{2n+2}+\alpha^2}{\alpha^{2n}-1}=\frac{\alpha^2-1}{\alpha^{2n}-1}=$$

$$=\frac{1}{\alpha^{2n-2}+\alpha^{2n-4}+\ \dots\ +\alpha^2+1}>\frac{1}{\alpha^{2n-1}}.$$

Следовательно, пользуясь формулой бинома,

$$\left(\frac{\alpha^{2(n+1)}-1}{\alpha^{2n}-1}\right)^n>\left(\alpha^2+\frac{1}{\alpha^{2n-1}}\right)^n=\alpha^{2n}+n\,\frac{\alpha^{2n-2}}{\alpha^{2n-1}}+\dots,$$

где точки стоят вместо положительных слагаемых.

Ввиду того, что $n>2$, написанное выражение больше, чем $\alpha^{2n}+1$. Значит,

$$\frac{(\alpha^{2(n+1)}-1)^{n+1}}{(\alpha^{2n}-1)^n}>(\alpha^{2(n+1)}-1)(\alpha^{2n}+1)=$$

$$=\alpha^{4n+2}+\alpha^{2n+2}-\alpha^{2n}-1=\alpha^{4n+2}+\alpha^{2n}(\alpha^2-1)-1=$$

$$=\alpha^{4n+2}+\alpha^{2n+1}-1>\alpha^{4n+2},$$

и теорема доказана.

22. Рассмотрим еще один класс последовательностей, основанных на числах Фибоначчи. Пусть x — произвольное число. Вычислим сумму

$$s_n(x)=u_1x+u_2x^2+\ \dots\ +u_nx^n.$$

Для этого воспользуемся прежде всего формулой Бине:

$$s_n(x)=\frac{\alpha-\beta}{\sqrt{5}}x+\frac{\alpha^2-\beta^2}{\sqrt{5}}x^2+\ \dots\ +\frac{\alpha^n-\beta^n}{\sqrt{5}}x^n=$$

$$=\frac{1}{\sqrt{5}}(\alpha x+\alpha^2x^2+\ \dots\ +\alpha^nx^n)-$$

$$-\frac{1}{\sqrt{5}}(\beta x+\beta^2x^2+\ \dots\ +\beta^nx^n).$$

Здесь в скобках написаны суммы двух геометрических прогрессий со знаменателями αx и βx. Известная формула, выражающая сумму геометрической прогрессии, справедлива в том случае, когда знаменатель прогрессии отличен от единицы. Если же он ранен единице, то все члены прогрессии равны друг другу, и их сумма вычисляется совсем просто.

В соответствии со сказанным рассмотрим сначала случай, когда $\alpha x\neq 1$ и $\beta x\neq 1$, т. е. когда

$$x\neq\frac{1}{\alpha}=-\beta\ \text{и}\ x\neq\frac{1}{\beta}=-\alpha.$$

В этих случаях, суммируя в (99) геометрические прогрессии, мы получаем

$$s_n(x) = \frac{1}{\sqrt{5}} \frac{\alpha^{n+1}x^{n+1} - \alpha x}{\alpha x - 1} - \frac{1}{\sqrt{5}} \frac{\beta^{n+1}x^{n+1} - \beta x}{\beta x - 1},$$

или, выполняя естественные преобразования,

$$s_n(x) = \frac{1}{\sqrt{5}} \frac{(\alpha^{n+1}x^{n+1} - \alpha x)(\beta x - 1) - (\beta^{n+1}x^{n+1} - \beta x)(\alpha x - 1)}{(\alpha x - 1)(\beta x - 1)}$$

и далее —

$$s_n(x) = \frac{1}{\sqrt{5}} \left(\frac{\alpha^{n+1}\beta x^{n+2} - \alpha^{n+1}x^{n+1} + \alpha x}{\alpha\beta x^2 - (\alpha + \beta)x + 1} - \right.$$
$$\left. - \frac{\alpha\beta^{n+1}x^{n+2} - \beta^{n+1}x^{n+1} + \beta x}{\alpha\beta x^2 - (\alpha + \beta)x + 1} \right).$$

Вспоминая, что

$$\alpha\beta = -1, \quad \alpha + \beta = 1, \quad \text{а} \quad \alpha - \beta = \sqrt{5},$$

имеем

$$s_n(x) = \frac{1}{\sqrt{5}} \frac{x\sqrt{5} - (\alpha^n - \beta^n)x^{n+2} - (\alpha^{n+1} - \beta^{n+1})x^{n+1}}{1 - x - x^2}$$

и окончательно

$$s_n(x) = \frac{x - u_n x^{n+2} - u_{n+1}x^{n+1}}{1 - x - x^2}. \qquad (100)$$

В частности, полагая в этой формуле x=1, получим

$$s_n(1) = u_1 + u_2 + \ldots + u_n = \frac{1 - u_n - u_{n+1}}{-1} = u_{n+2} - 1,$$

что соответствует сказанному в п. 1.

При x= — 1 имеем

$$s_n(-1) = u_1 - u_2 + \ldots + (-1)^{n-1}u_n =$$
$$= \frac{-1 - u_n(-1)^{n+2} - u_{n+1}(-1)^{n+1}}{-1} = (-1)^{n+1}u_{n-1} + 1$$

(ср. формулу (82)).

Рассмотрим теперь оставшиеся «особые» случаи.

Пусть $x = \frac{1}{\alpha} = -\beta$. Тогда в (99) каждый член первой прогрессии равен единице, и сумма этой прогрессии равна n. Во второй же прогрессии знаменатель оказывается равным $-\beta^2$,

Таким образом,

$$s_n\left(\frac{1}{\alpha}\right)=\frac{1}{\sqrt{5}}\left(n-\left(\beta^2-\beta^4+\ \dots\ +(-1)^{n-1}\beta^{2n}\right)\right)=$$

$$=\frac{1}{\sqrt{5}}\left(n-\frac{\beta^2-(-1)^n\beta^{2n+2}}{1+\beta^2}\right)=$$

$$=\frac{1}{\sqrt{5}}\left(n-\frac{\beta^2}{1+\beta^2}+(-1)^n\beta^{2n}\frac{\beta^2}{1+\beta^2}\right).$$

Замечая, что

$$1+\beta^2=2+\beta=2+\frac{1-\sqrt{5}}{2}=\frac{5-\sqrt{5}}{2},$$

а

$$\frac{\beta^2}{1+\beta^2}=\frac{1+\beta}{2+\beta}=\frac{3-\sqrt{5}}{5-\sqrt{5}}=\frac{(3-\sqrt{5})(5+\sqrt{5})}{(5-\sqrt{5})(5+\sqrt{5})}=$$

$$=\frac{10-2\sqrt{5}}{20},$$

мы получаем окончательно

$$s_n\left(\frac{1}{\alpha}\right)=\frac{n}{\sqrt{5}}-\frac{\sqrt{5}-1}{10}+(-1)^n\beta^{2n}\frac{\sqrt{5}-1}{10}. \qquad (101)$$

Наконец, пусть $x=\frac{1}{\beta}$. В этом случае в (99) единице равен знаменатель, второй прогрессии, а знаменатель первой прогрессии равен — α^2. Мы имеем

$$s_n\left(\frac{1}{\beta}\right)=\frac{1}{\sqrt{5}}\left(\left(\alpha^2-\alpha^4+\ \dots\ +(-1)^{n-1}\alpha^{2n}\right)-n\right).$$

Аналогично предыдущему получаем

$$s_n\left(\frac{1}{\beta}\right)=\frac{1}{\sqrt{5}}\left(\frac{\alpha^2-(-1)^n\alpha^{2n+2}}{1+\alpha^2}-n\right)=$$

$$=\frac{1}{\sqrt{5}}\left((-1)^{n+1}\alpha^{2n}\frac{\alpha^2}{1+\alpha^2}+\frac{\alpha^2}{1+\alpha^2}-n\right)$$

и в итоге

$$s_n\left(\frac{1}{\beta}\right)=(-1)^{n+1}\frac{1+\sqrt{5}}{10}\alpha^{2n}+\frac{1+\sqrt{5}}{10}-\frac{n}{\sqrt{5}}. \qquad (102)$$

23. Посмотрим, как ведет себя сумма $s_n(x)$ при фиксированном x и неограниченно возрастающем n.,

Переходя в равенстве (99) к пределу по n, получаем

$$\lim_{n\to\infty} s_n(x) = \lim_{n\to\infty} \frac{1}{\sqrt{5}}((\alpha x + \alpha^2 x^2 + \ldots + \alpha^n x^n) - (\beta x + \beta^2 x^2 + \ldots + \beta^n x^n)) =$$

$$= \frac{1}{\sqrt{5}} \lim_{n\to\infty} (\alpha x + \alpha^2 x^2 + \ldots + \alpha^n x^n) - \frac{1}{\sqrt{5}} \lim_{n\to\infty} (\beta x + \beta^2 x^2 + \ldots + \beta^n x^n).$$

Здесь под знаками двух последних пределов стоят суммы геометрических прогрессий. Поэтому сами пределы являются суммами соответствующих бесконечных геометрических прогрессий. Но, как известно, для того, чтобы можно было говорить о сумме бесконечной геометрической прогрессии, необходимо и достаточно, чтобы ее знаменатель по абсолютной величине был меньше единицы. В имеющихся у нас прогрессиях знаменатели равны αx и βx. Здесь $|\alpha|>|\beta|$ Поэтому из $|\alpha x|<1$ следует $|\beta x|<1$. Таким образом, выполнение неравенства $|\alpha x|<1$ будет обеспечивать существование всех интересующих нас в данный момент пределов. Итак, предел

$$\lim_{n\to\infty} s_n(x) \qquad (103)$$

существует, если $|x| < \frac{1}{\alpha}$. Обозначим этот предел через *s(x)*. Для его вычисления мы можем воспользоваться формулой (100).

Заметим для этого, что на основании сказанного в п. 20

$$u_n \leqq \frac{\alpha^n}{\sqrt{5}} + 1.$$

Поэтому

$$\lim_{n\to\infty} u_n x^{n+2} \leqq \lim_{n\to\infty} \left(\frac{\alpha^n}{\sqrt{5}} + 1\right) x^{n+2} = \frac{x^2}{\sqrt{5}} \lim_{n\to\infty} (\alpha x)^n + \lim_{n\to\infty} x^{n+2}.$$

Ввиду $|\alpha x|<1$ должно быть и $|x|<1$, так что оба написанных предела равны нулю. По тем же причинам и

$$\lim_{n\to\infty} u_{n+1} x^{n+1} = 0.$$

Следовательно, переходя в формуле (100) к пределу по *n* при неограниченном возрастании *n,* получаем

$$s(x) = \lim_{n\to\infty} s_n(x) = \lim_{n\to\infty} \frac{x - u_n x^{n+2} - u_{n+1}x^{n+1}}{1 - x - x^2} =$$

$$= \frac{1}{1 - x - x^2}\left(x - \lim_{n\to\infty} u_n x^{n+2} - \lim_{n\to\infty} u_{n+1}x^{n+1}\right) = \frac{x}{1 - x - x^2}.$$

Найденный результат можно переписать в развернутом виде как

$$u_1 x + u_2 x^2 + \ldots + u_n x^n + \ldots = \frac{x}{1 - x - x^2}. \qquad (104)$$

Придавая переменной x те или иные значения, будем получать различные конкретные формулы. Например, полагая $x = \frac{1}{2}$, обнаружим, что

$$\frac{u_1}{2} + \frac{u_2}{2^2} + \ldots + \frac{u_n}{2^n} + \ldots = 2.$$

24. Формулу (104) можно получить также при помощи несколько иных рассуждений. Напишем

$$u_1 x + u_2 x^2 + \ldots + u_n x^n + \ldots = s(x) \qquad (105)$$

(помня при этом, что выражение $s(x)$ имеет смысл лишь при $|x| < \frac{1}{a}$) и умножим это равенство почленно на x и на x^2:

$$u_1 x^2 + u_2 x^3 + \ldots + u_n x^{n+1} + \ldots = x s(x), \qquad (106)$$

$$u_1 x^3 + u_2 x^4 + \ldots + u_n x^{n+2} + \ldots = x^2 s(x). \qquad (107)$$

Вычитая из равенства (105) оба равенства (106) и (107), мы после приведения подобных членов получим

$$u_1 x + (u_2 - u_1) x^2 + (u_3 - u_2 - u_1) x^3 +$$
$$+ (u_4 - u_3 - u_2) x^4 + \ldots + (u_n - u_{n-1} - u_{n-2}) x^n + \ldots =$$
$$= (1 - x - x^2) s(x).$$

Все заключенные в скобках выражения в левой части равенства, кроме первого равны нулю, и это равенство превращается в

$$x = (1 - x - x^2) s(x),$$

откуда и следует (104).

25. Говоря о числе Фибоначчи u_n, мы пока все время предполагали, что его номер n является целым положительным числом. Однако основное рекуррентное соотношение, определяющее числа Фибоначчи, может быть записано и как

$$u_{n-2} = u_n - u_{n-1}. \qquad (108)$$

При этом оно будет служить для выражения чисел Фибоначчи с меньшими номерами через числа с большими.

Полагая последовательно в (108) $n = 2, 1, 0, -1, \ldots$, мы можем вычислить

$$u_0 = 0, \qquad u_{-1} = 1, \qquad u_{-2} = -1, \qquad u_{-3} = 2, \ldots$$

и вообще, как легко убедиться (пусть читатель убедится сам!),

$$u_{-n} = (-1)^{n+1} u_n. \tag{109}$$

Это простое выражение числа Фибоначчи с произвольным целым номером позволяет сводить все задачи о таких числах Фибоначчи к задачам об обычных числах Фибоначчи с натуральными номерами.

Например, для вычисления суммы n «первых назад» чисел Фибоначчи

$$u_{-1} + u_{-2} + \ldots + u_{-n}$$

достаточно переписать ее в соответствии с (109):

$$u_1 - u_2 + \ldots + (-1)^{n-1} u_n,$$

и вспомнить формулу (1.182):

$$u_{-1} + u_{-2} + \ldots + u_{-n} = (-1)^{n+1} u_{n-1} + 1 = -u_{-n+1} + 1.$$

Опирающееся на основное рекуррентное соотношение индуктивное рассуждение о числах Фибоначчи типа переходов «от n и n+1 к n+2» можно в связи с соотношением (108) проводить по схеме «от n и n — 1 к n — 2». В частности, таким образом без труда доказывается для любых целых n и m формула (84)

$$u_{n+m} = u_{n-1}u_m + u_n u_{m+1}.$$

26. Основные уравнения для чисел α и β:

$$\alpha^{n+2} = \alpha^n + \alpha^{n+1},$$
$$\beta^{n+2} = \beta^n + \beta^{n+1},$$

справедливы не только для положительных, но и для любых целых значений n (для дробных значений n эти равенства тоже в известном смысле остаются в силе, но мы на этом не будем останавливаться). Отсюда легко получить, что формула Бине

$$u_n = \frac{\alpha^n - \beta^n}{\sqrt{5}}$$

имеет место для любого целого n.

Заметим в заключение, что и результат п. 17 можно (по индукции «назад») перенести на отрицательные значения номера:

$$\alpha^{-n} = u_{-n}\alpha + u_{-n-1}. \tag{110}$$

Это равенство переписывается как

$$(-1)^n \beta^n = (-1)^n u_n \frac{1}{\beta} + (-1)^n u_{n+1},$$

т. е.

$$\beta^{n+1} = u_{n+1}\beta + u_n.$$

Кроме того, (110) можно представить в виде

$$\alpha^{-n} = (-1)^{n-1} u_n \alpha + (-1)^n u_{n+1},$$

т. е.

$$(-1)^n \alpha^{-n} = u_{n+1} - u_n \alpha,$$

или, иначе,

$$\frac{u_{n+1}}{u_n} - \alpha = (-1)^n \alpha^{-n} \frac{1}{u_n}. \qquad (111)$$

27. Числа Фибоначчи могут составить основу своеобразной «фибоначчиевой» системы счисления, т. е., представления любого натурального числа a в виде некоторой последовательности «цифр» $\varphi_1\varphi_2 \ldots \varphi_r$. Эта последовательность может быть получена следующим (индуктивным!) образом.

Отнимем от заданного числа $a=a_0$ наибольшее из не превосходящих его чисел Фибоначчи u_n и, написав цифру $\varphi_1= 1$ и разность $a_1 =a_0 — u_n$ будем считать это первым шагом нашего построения.

Предположим, что k шагов построения уже выполнены, в результате чего появилась последовательность цифр

$$\varphi_1\varphi_2 \ldots \varphi_k \qquad (112)$$

состоящая из нулей и единиц, а также некоторое число n_k. Тогда k+1-й шаг построения будет состоять в следующем: сравним число a_k с числом Фибоначчи u_{n-k}, и если окажется $a_k < u_{n-k}$, то припишем к последовательности (112) $\varphi_{k+1} = 0$ и фиксируем число $a_{k+1} = a_k$, а если будет $a_k \geq u_{n-k}$, то припишем к (112) $\varphi_{k+1}=1$ и положим $a_{k+1} = a_k — u_{n-k}$. Мы выподним n — 1 шагов этого процесса, в результате чего, очевидно, придем к n—1-членной последовательности (112) и числу $a_{n-1} = 0$.

Фактически описанный процесс является последовательным выделением из числа a слагаемых, равных наибольшим возможным числам Фибоначчи, т. е. представлением a в виде суммы различных чисел Фибоначчи.

Окончательную соответствующую числу a последовательность (112) будем называть его *фибоначчиевой записью* и обозначать через Ф*(а)*. Составляющие Ф(а) нули и единицы назовем *фибоначчиевыми цифрами* числа *а*. Ясно, что если Ф(а) = $\varphi_1\varphi_2 \ldots \varphi_{n-1}$,то

$$a = u_n\varphi_1 + u_{n-1}\varphi_2 + \ldots + u_2\varphi_{n-1}. \qquad (113)$$

Поясним сказанное на примере. Пусть a = 19.

Тогда

$$u_n = 13\ (n=7),\quad \varphi_1 = 1,\quad a_1 = 19 - 13 = 6;$$
$$u_6 = 8 > a_1,\qquad \varphi_2 = 0,\quad a_2 = a_1 = 6;$$
$$u_5 = 5 \leqq a_2,\qquad \varphi_3 = 1,\quad a_3 = 6 - 5 = 1;$$
$$u_4 = 3 > a_3,\qquad \varphi_4 = 0,\quad a_4 = a_3 = 1:$$
$$u_3 = 2 > a_4,\qquad \varphi_5 = 0,\quad a_5 = a_4 = 1;$$
$$u_2 = 1 \leqq a_5,\qquad \varphi_6 = 1,\quad a_6 = 1 - 1 = 0.$$

Таким образом, Ф(19)= 101001, и 19 = $u_7 + u_5 + u_2$.

Ясно, что каждое число a имеет единственную фибоначчиеву запись Ф*(a)*. Однако не всякая начинающаяся с единицы последовательность нулей и единиц обязана быть фибоначчиевой записью Ф(a) для некоторого числа a. Например, в Ф(a) не могут стоять две единицы подряд.

Действительно, пусть в Ф(a) две единицы встречаются подряд после некоторого нуля:

$$\varphi_k = 0,\ \varphi_{k+1} = \varphi_{k+2} = 1.$$

Это значит, что

$$a_{k-1} < u_{n-k+1}, \tag{114}$$
$$a_{k-1} = a_k,\quad a_k - u_{n-k} = a_{k+1},\quad a_{k+1} \geqq u_{n-k-1}.$$

Но почленное сложение всех составляющих вторую строку соотношений дает нам

$$a_{k-1} \geqq u_{n-k} + u_{n-k-1} = u_{n-k+1},$$

что противоречит (114).

Значит, две единицы подряд могли бы встретиться в Ф*(a)* лишь в том случае, когда впереди них вовсе не было бы нулей, т. е. было бы $\varphi_1=\varphi_2=1$. Но тогда по определению процесса $a_1 = a_0 - u_n \geqq u_{n-1}$, так что

$$a_0 \geqq u_n + u_{n-1} = u_{n+1},$$

и u_n не является наибольшим числом Фибоначчи, не превосходящим a.

Вместе с тем ограничение, заключающееся в отсутствии двух стоящих рядом единиц, оказывается уже достаточным: всякая последовательность из n— 1 нулей и единиц, начинающаяся с единицы и не содержащая двух единиц подряд, есть фибоначчиева запись Ф*(a)* некоторого числа a, для которого

$$u_n \leqq a < u_{n+1}. \tag{115}$$

В этом можно убедиться, воспользовавшись, например, результатом задачи о прыгуне из п. 6. Пусть имеется n— 1 клетка, по которым прыгун прыгает доступными для него способами (т. е. на соседнюю

клетку или через клетку). После выполнения прыжков все клетки, в которых прыгун побывал, помечаются нулями, а остальные клетки — единицами. Так как всего возможно u_{n-1} способов выполнения прыжков, различных способов пометок будет тоже u_{n-1}.

Если к каждой из них приписать впереди единицу, то мы получим запись, которая может быть фибоначчиевым значением для числа *a,* удовлетворяющего (114). Но таких чисел ровно u_{n-1} и каждое из них должно иметь свою запись; поэтому каждой записи соответствует хотя бы одно число.

2.6. Аксиоматика множеств

Используя аксиоматический подход, формально построим теорию множеств на основании следующих аксиом.

Аксиома существования. *Существует по крайней мере одно множество.*

Аксиома объемности (экстенциональности). *Если множества* M_a и M_b *составленные из одних тех же элементов, то они совпадают (равны):*

$$M_a = M_b.$$

Аксиома объединения. *Для произвольных множеств* M_a *и* M_b *существует множество, элементами которого являются все элементы множества* M_a *и все элементы множества* M_b *и которое никаких других элементов не содержит.*

Из аксиом объемности и объединение следует, что для произвольных множеств M_a и M_b множество, которое удовлетворяет условиям аксиомы объединения, единственно. Действительно, если были бы два таких множества M_{c_1} и M_{c_2}, то они содержали одни и те же элементы (все элементы, которые принадлежат множеству M_a, и все элементы множества M_b) и потому, согласно аксиоме объемности, $M_{c_1} = M_{c_2} = M_c$. Назовем это единственное множество M_c *объединением* множеств M_a и M_b и будем писать

$$M_c = M_a \cup M_b.$$

Аксиома разности. *Для произвольных множеств* M_a *и* M_b *существует множество, элементами которой есть те и только те*

элементы множества M_a, которые не являются элементами множества M_b.

Аналогично, из второй и четвертой аксиом делаем заключення, что для произвольных множеств M_a и M_b существует в точности одно множество, которое содержит элементы множества M_a, не принадлежащие множеству M_b. Назовем это множество M_c *разностью* множеств M_a и M_b:

$$M_c = M_a \backslash M_b.$$

Аксиома степени. *Для каждого множества M существует семейство множеств $B(M)$ (булеан), элементами которого являются все подмножества M_i, $M_i \subset M$, и только они.*

Аксиома существования пустого множества. *Существует такое множество Ø, что ни один элемент ему не принадлежит.*

Если операции и понятия теории множеств были введены интуитивно, то аксиоматический подход позволяет формально на основании введенных **шести аксиом** определить эти операции и понятия теории множеств.

С помощью операций объединения и разности, используя введенные аксиомы, определим еще три операции на множествах.

Пересечение множеств M_a и M_b определяется формулой

$$M_a \cap M_b = M_a \backslash (M_a \backslash M_b).$$

Можно показать, что элементами пересечения $M_a \cap M_b$ есть те и только те элементы, которые принадлежат как множеству M_a, так и множеству M_b.

Дополнение $\overline{M}$ множества M определяется формулой

$$\overline{M} = 1 \backslash M.$$

Симметрическая разность множеств M_a и M_b определяется формулой

$$M_a \backslash M_b = (M_a \backslash M_b) \cup (M_b \backslash M_a).$$

На основании введенной аксиоматики можно доказать справедливость как приведенных выше законов, которые определяют свойства сигнатуры алгебры множеств (законы идемпотентности, коммутативностм, ассоциативности, дистрибутивности, действия с константами, двойного дополнения, законы де-Моргана), так и следующих законов:

закон дистрибутивности пересечения относительно разности

$$M_a \cap (M_b \backslash M_c) = M_a \cap M_b \backslash M_a \cap M_c;$$

закон коммутативности симметрической разности

$$M_a \backslash M_b = M_b \backslash M_a;$$

закон ассоциативности симметрической разности

$$M_a \backslash M_b \backslash (M_b \backslash M_c) = (M_a \backslash M_b) \backslash M_c;$$

закон дистрибутивности пересечения относительно симметрической разности

$$M_a \cap (M_b \backslash M_c) = M_a \cap M_b \backslash M_a \cap M_c;$$

законы склеивания

$$M_a \cap M_b \cup M_a \cap \overline{M}_b = M_a (M_a \cup M_b) \cap (M_a \cup \overline{M}_b) = M_a;$$

законы поглощения

$$M_a \cup M_a \cap M_b = M_a, \quad M_a \cap (M_a \cup M_b) = M_a;$$

законы Порецького

$$M_a \cup \overline{M}_a \cap M_b = M_a \cup M_b,$$

$$M_a \cap (\overline{M}_a \cup M_b) = M_a \cap M_b.$$

Используя эти законы, рассмотрим задачу минимизации представления множества M с помощью операций $\cup$, $\cap$, $^-$.

Под *сложностью представления множества* M будем понимать число символов M_i, $\overline{M}_i$ в задающем его выражении.

Пусть в пространстве **1**=$\{M_1, M_2, M_3\}$ задано множество вида

$$M(M_1,M_2,M_3) = \overline{M}_1 \cap \overline{M}_2 \cap \overline{M}_3 \cup \overline{M}_1 \cap \overline{M}_2 \cap M_3 \cup \overline{M}_1 \cap M_2 \cap \cap \overline{M}_3 \cup M_1 \cap \overline{M}_2 \cap \overline{M}_3 \cup M_1 \cap M_2 \cap \overline{M}_3 \cup M_1 \cap M_2 \cap M_3.$$

На основании законов идемпотентности, коммутативности и ассоциативности объединения получаем

$$M(M_1,M_2,M_3) = (\overline{M}_1 \cap \overline{M}_2 \cap \overline{M}_3 \cup \overline{M}_1 \cap \overline{M}_2 \cap M_3) \cup (\overline{M}_1 \cap \overline{M}_2 \cap \cap \overline{M}_3 \cup \overline{M}_1 \cap M_2 \cap \overline{M}_3) \cup (\overline{M}_1 \cap \overline{M}_2 \cap \overline{M}_3 \cup M_1 \cap \overline{M}_2 \cap \overline{M}_3) \cup \cup (M_1 \cap \overline{M}_2 \cap \overline{M}_3 \cup M_1 \cap M_2 \cap \overline{M}_3) \cup (M_1 \cap M_2 \cap \overline{M}_3 \cup M_1 \cap M_2 \cap M_3).$$

Используя законы коммутативности объединения и склеивания, имеем

$$M(M_1,M_2,M_3) = \overline{M}_1 \cap \overline{M}_2 \cup \overline{M}_1 \cap \overline{M}_3 \cup \overline{M}_2 \cap \overline{M}_3 \cup M_1 \cap \overline{M}_3 \cup \cup M_1 \cap M_2.$$

Согласно законам коммутативности объединения и пересечения и закону склеивания, имеем

$$M(M_1,M_2,M_3)= \overline{M}_1 \cap \overline{M}_2 \cup \overline{M}_3 \cup \overline{M}_2 \cap \overline{M}_3 \cup M_1 \cap M_2.$$

Согласно законам коммутативности пересечения и поглощения, имеем

$$M(M_1,M_2,M_3)= \overline{M}_1 \cap \overline{M}_2 \cup \overline{M}_3 \cup M_1 \cap M_2.$$

Сложность представления заданного множества уменьшилась от 21 до 5.

Последовательность применения законов будем называть *стратегией преобразований*. Сложность представления множества, получаемого в результате применения этих законов (каждый из которых определяет эквивалентное преобразование), зависит от используемой стратегии.

Найдем стратегию, которая всегда порождает минимальное выражение заданного множества.

Рассмотрим алгебру $A=<B(\mathbf{1}), \cup, \cap, \bar{}>$ и определим множества, которые могут быть рождены (образованы) из произвольных подмножеств $M_1,M_2,\ldots,M_n$, называемых порождающими или образующими пространства **1** с помощью операций $\cup, \cap, \bar{}$.

Множество

$$M_i^{\sigma_i} = \begin{cases} M_i & \text{при } \sigma_i = 1, \\ \overline{M}_i & \text{при } \sigma_i = 0 \end{cases} \quad i=1, 2,\ldots,n,$$

в дальнейшем будем называть *первичным термом*. Множество вида

$$\bigsqcap_{i=1}^{n} M_i^{\sigma_i} = M_1^{\sigma_1} \cap M_2^{\sigma_2} \cap \ldots \cap M_n^{\sigma_n}, \quad \sigma_i = 0, 1,$$

назовем *конституентой.*

Общее число различных конституент не превышает $2^{n.}$ Каждой конституенте можно сопоставить двоичный набор длины n, число этих наборов равно $2^{n.}$ Если некоторые конституенты равны $\varnothing$, то общее количество конституент меньше 2^n, при этом среди подмножеств найдутся хотя бы два такие, которые можно выразить одно через другое, т.е. зависимые. Например, если $n = 2$ и $M_2 = \overline{M}_1$, то существуют только две отличные от $\varnothing$ конституенты

$\varnothing = M^0{}_1 \cap M^0{}_2 = M^1{}_1 \cap M^1{}_2$, $C_1 = M^0{}_1 \cap M^1{}_2$, $C_2 = M^1{}_1 \cap M^0{}_2$.

Лемма 1. *Пересечение двух различных конституент пусто.*

Действительно, если конституенты

$$C_a = \prod_{i=1}^{n} M_i^{\sigma_i} \quad \text{и} \quad C_b = \prod_{i=1}^{n} M_i^{\sigma_i^*}$$

различны, то $\sigma_k \neq \sigma_k^*$ по крайней мере для одного k, $k \leq n$. Но тогда $M_k^{\sigma_k} \cap M_k^{\sigma_k^*} = \varnothing$ и, следовательно, $C_a \cap C_b = \varnothing$.

Лемма 2. *Объединение всех конституент равно* **1**.

Представим **1** в виде

$$\mathbf{1} = \prod_{i=1}^{n} \left(M_i^0 \cup M_i^1 \right)$$

и, раскрыв дужки, в правой части равенства получим объединение всех конституент.

Лемма 3. *Множество M_i - равно объединению конституент, каждая из которых содержит M_i^1.*

Согласно лемме 2,

$$1 = C_1 \cup C_2 \cup \ldots \cup C_i = \bigcup_{i=1}^{l} C_i$$

где C_i, i=l, 2,...,l, - конституента. Определим пересечение левой и правой частей этого выражения с M_i. Имеем

$$M_i = (M_i \cap C_1) \cup (M_i \cap C_2) \cup \ldots \cup (M_i \cap C_i).$$

Если C_j содержит в качестве аргумента пересечения M_i^0, то $M_i \cap C_j = \varnothing$. Если же C_j содержит M_i^1, то $C_j \cap M_i = C_j$. Следовательно, M_i - объединение тех конституент, которые содержат M_i^1 в качестве сомножителя.

Теорема 4. *Каждое непустое множество, которое образовано из множеств M_1, M_2,...,M_n с помощью операций $\cup, \cap, \bar{}$, является объединением некоторого числа конституент.*

Согласно лемме 3, теорема справедлива для множеств M_1, M_2,...,M_n. Следовательно, достаточно доказать, что если произвольные множества M_a и M_b представимы в виде объединения некоторого числа конституент, то и множества $M_a \cup M_b$, $M_a \cap M_b$ и $\overline{M}$, если они непустые, также можно представить в виде объединения конституент.

Пусть множества M_a и M_b представимы в виде объединения конституент $M_a = C_{a_1} \cup C_{a_2} \cup \ldots \cup C_{a_k}$ и $M_b = C_{b_1} \cup C_{b_2} \cup \ldots \cup C_{b_s}$ ... Тогда множество $M_a \cup M_b$, очевидно, можно представить в виде объединения конституент.

Согласно закону дистрибутивности,

$$M_a \cap M_b = (C_{a_1} \cap C_{b_1}) \cup \ldots \cup (C_{a_k} \cap C_{b_s}),$$

при этом если $C_{a_\alpha} \neq C_{b_\beta}$ то, согласно лемме 1, $C_{a_\alpha} \cap C_{b_\beta} = \varnothing$, в противном случае $C_{a_\alpha} = C_{b_\beta}$. Следовательно, пересечение $M_a \cap M_b$ либо пусто, либо представимо в виде объединения конституент. Докажем, что множество $\overline{M}$ также представимо в виде объединения конституент, если

$$M = C_1 \cap C_2 \cap \ldots \cap C_k.$$

Согласно закону де-Моргана,

$$\overline{M} = \overline{C_1 \cup C_2 \cup \ldots \cup C_k} = \overline{C_1} \cap \overline{C_2} \cap \ldots \cap \overline{C_k} =$$

$$= \overline{M_1^{\sigma_{11}} \cap M_2^{\sigma_{12}} \cap \ldots \cap M_n^{\sigma_{1n}}} \cap \overline{M_1^{\sigma_{21}} \cap M_2^{\sigma_{22}} \cap \ldots \cap M_n^{\sigma_{2n}}} \cap \ldots$$

$$\cap \overline{M_1^{\sigma_{k1}} \cap M_2^{\sigma_{k2}} \cap \ldots \cap M_n^{\sigma_{kn}}} = (\overline{M}_1^{\sigma_{11}} \cup \overline{M}_2^{\sigma_{12}} \cup \ldots \cup \overline{M}_n^{\sigma_{1n}}) \cap$$

$$\cap (\overline{M}_1^{\sigma_{21}} \cup \overline{M}_2^{\sigma_{22}} \cup \ldots \cup \overline{M}_n^{\sigma_{2n}}) \cap \ldots \cap (\overline{M}_1^{\sigma_{k1}} \cup \overline{M}_2^{\sigma_{k2}} \cup \ldots \cup \overline{M}_n^{\sigma_{kn}})$$

Раскрывая скобки и используя соотношения $M_\alpha \cap \overline{M}_\alpha = \varnothing$, $M_\alpha \cup \overline{M}_\alpha = \mathbf{1}$, а также добавляя в те пересечения, в которых отсутствует нижний индекс β, сомножитель $M_\beta \cup \overline{M}_\beta$, получаем, что множество $\overline{M}$ также представимо в виде объединения конституент.

Теорема 5. *Из n множеств в алгебре* $A = \langle B(\mathbf{1}), \cup, \cap, \bar{} \rangle$ *можно образовать не более чем* 2^{2^n} *множеств.*

Каждое множество M, согласно теореме 4, является объединением конституент, число которых не превышает 2^n; следовательно, число различных объединений не превышает 2^{2^n}. При этом если множества $M_1, M_2, \ldots, M_n$ независимы, т.е. все конституенты отличны от пустого множества, то число различных конституент равно 2^n и число множеств, которые образованы с этих конституент в виде их объединения, равно 2^{2^n} (с учетом пустого множества).

Введение понятия конституенты разрешает задавать множество M при фиксированных независимых подмножествах $M_1, M_2,...,M_n$ универсального множества **1** в виде объединения конституент:

$$M = \bigsqcup_{j} \prod_{i=1}^{n} M_i^{\sigma_i}$$

Каждое фиксированное множество $M_i \in$ **1** разбивает пространство на две частей: на собственно M_i и на $\overline{M}_i$. При независимых множиннах $M_i \in \{M_i / \ i = 1,...,n\}$ пространство разбивается на $\underbrace{2 \times 2 \times \ldots \times 2}_{n \text{ раз}} = 2^n$ областей. Каждая область является пересечением n множеств M_i или $\overline{M}_i, i=1,...,n$. Сопоставим этой области двоичный вектор $(\sigma_1,\sigma_2, ..., \sigma_n)$, в котором $\sigma_i = 1$, если в пересечение $C = \prod_i M_i^{\sigma_i}$ входит M_i , и $\sigma_i=0$, если входит $\overline{M}_i$, а также десятичный эквивалент

$$d(C) = \sum_{i=1}^{n} \sigma_i \cdot 2^{i-1}.$$

Любое множество M в пространстве **1** можно задать в виде объединения этих областей. Сопоставим множеству M двоичный вектор длины 2^n, в котором i-му разряду отвечает область с десятичным эквивалентом, равным i. Вектор, который определяет множество, представим в виде десятичного эквивалента:

$$d(M) = \sum_{i=0}^{2^n-1} c_i \cdot 2^i, \quad c_i = 0,1$$

Следовательно, множество M в пространстве может быть задано в виде соответствующего десятичного эквивалента.

Рассмотрим, например, в трехмерному пространстве **1**$=\{M_1,M_2,M_3\}$ множество $M(M_1,M_2,M_3)$ с десятичным эквивалентом $d(M)=217$. Имеем

$217= 1\cdot 2^7 + 1\cdot 2^6 + 0\cdot 2^5 + 1\cdot 2^4 + 1\cdot 2^3 + 0\cdot 2^2 + 0\cdot 2^1 + 1\cdot 2^0$.

Множеству M отвечает двоичный вектор (1,1,0,1,1,0,0,1), который определяет включение областей в множество M (рис. 33, *а*).

Кроме диаграммы Ейлера пространство может быть задано в виде гиперкуба или n-мерного куба (n-размерность пространства, равная числу фиксированных множеств).

Гиперкубом (n-мерным кубом) называется граф *H,* каждая вершина которого взаимно однозначно соответствует области пространства, и две вершины соединены ребром, если они соответствуют соседним областям (имеющим общую границу). Сопоставленные этим областям двоичные векторы отличаются в одном и только одном разряде.

Гиперкуб для рассматриваемого пример изображен на рис.33,*б* (вершины, которые отвечают конституентам множества *M,* заштрихованы).

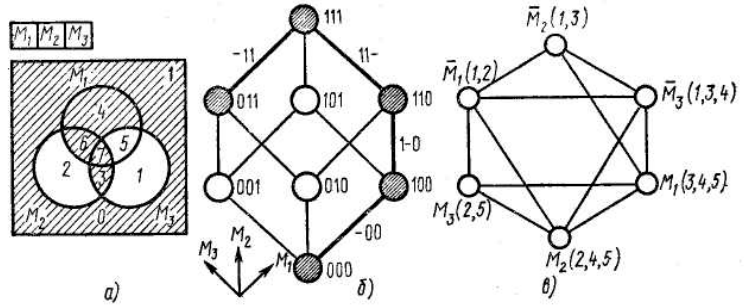


Рис. 33

Часто множество *M* задают в виде двоичной таблицы, каждой строке которой взаимно однозначно соответствует конституента. Множество строк таблицы линейно упорядочено по возрастанию десятичного эквивалента соответствующего двоичного набора. Столбцам соответствуют множества, которые образовывают пространство, последний столбец сопоставляется множеству *M,* и единица указывает на вхождение соответствующей конституенты в множество *M.* В данном случае имеем табл. 3.

Таблица 3.

d(C)	M_1	M_2	M_3	*M*
0	0	0	0	1
1	0	0	1	0
2	0	1	0	0
3	0	1	1	1
4	1	0	0	1
5	1	0	1	0
6	1	1	0	1
7	1	1	1	1

Аналитически множество *M* задается в виде

$$M = \overline{M}_1 \cap \overline{M}_2 \cap \overline{M}_3 \cup \overline{M}_1 \cap M_2 \cap M_3 \cup M_1 \cap \overline{M}_2 \cap \overline{M}_3 \cup M_1 \cap$$
$$\cap M_2 \cap \overline{M}_3 \cup M_1 \cap M_2 \cap M_3,$$

или в виде мографа

$$G^M = \langle V, S_3 \rangle, \quad V = \{M_i, \overline{M}_i / i = 1, 2, 3\}, \quad S_3 \subset V^3,$$

$$S_3 = \{\underbrace{\{\overline{M}_1, \overline{M}_2, \overline{M}_3\}}_{1}, \quad \underbrace{\{\overline{M}_1, M_2, M_3\}}_{2}, \quad \underbrace{\{M_1, \overline{M}_2, \overline{M}_3\}}_{3},$$

$$\underbrace{\{M_1, M_2, \overline{M}_3\}}_{4}, \quad \underbrace{\{M_1, M_2, M_3\}}_{5}, \quad \underbrace{\{M_1, \overline{M}_2, \overline{M}_3\}}_{3}$$ (рис. 33, *в*).

В рассмотренной алгебре $A=<B(\mathbf{1}), \cup, \cap, \bar{} >$ операции являются зависимыми. Действительно, согласно закону де-Моргана, любое множество из 2^{2^n} множеств может быть построено и с помощью алгебры $A=<B(\mathbf{1}), \cup, \bar{} >$. Равносильными в смысле порождения любого множества из 2^{2^n} множеств являются алгебры $A=<B(\mathbf{1}), \cup, \bar{} >$, $A=<B(\mathbf{1}), \cap, \bar{} >$, которые могут быть заменены соответственно алгебрами $A=<B(\mathbf{1}), \cup, \setminus, \mathbf{1} >$, $A=<B(\mathbf{1}), \cap, \setminus, \mathbf{1} >$ согласно формуле $\overline{M} = \mathbf{1} \setminus M$, где универсум $\mathbf{1}$ рассматривается как нуль-местная операция.

Алгебра $<B(\mathbf{1}), \cup, \setminus, \mathbf{1} >$ в силу равенств

$$M_a \cup M_b = M_a \setminus M_b \setminus (M_a \cap M_b),$$
$$M_a \setminus M_b = M_a \setminus (M_a \cap M_b)$$

может быть заменена алгеброй вида $<B(\mathbf{1}), \cap, \setminus, \mathbf{1} >$.

Рассмотрим задачу минимизации представления множеств в алгебре Кантора. Пересечение попарно различных множеств $\cap M_i^{\sigma_i}$ называется *элементарным.* Выражение, которое задает множество M^i в виде объединения различных элементарных пересечений, называется *нормальной формой Кантора* (НФК) *множества М.* Объединение конституент множества M называется *совершенной* НФК *множества М.*

Минимальной НФК *множества М* называется НФК этого множества, которое имеет минимальную сложность.

Рассмотрим *метод Квайна,* который будем использовать для получения минимальной НФК множества *М.* Этот метод заключается в последовательном выполнении таких этапов.

1. Выделение максимальных интервалов. *Интервалом множества* M называется множество конституент множества M, образующих гиперкуб (некоторой размерности).

Очевидно, что мощность интервала равна степени 2 (т.е. 2^0, 2^1 и т.д.).

Запишем, например, множество интервалов для рассмотренного выше примера: {000, 100, 110, 011, 111, -00, 1-0, 11-, -11}. Здесь и далее «-» означает, что множество, соответствующее этому разряду, в пересечении отсутствует, т.е. по этому множеству после объединения соответствующих конституент состоялось склеивание.

Например, интервал -00, соответствующий множеству конституент 000 и 100, получается в результате преобразования

$$\overline{M}_1 \cap \overline{M}_2 \cap \overline{M}_3 \cup M_1 \cap \overline{M}_2 \cap \overline{M}_3 = M_2 \cap M_3.$$

Интервал $\boldsymbol{I}_\alpha$ называется *максимальным интервалом* $\boldsymbol{I}_{\text{макс}}$ *множества* M, если не найдется другого интервала $\boldsymbol{I}_\beta$ этого множества, содержащего интервал $\boldsymbol{I}_\alpha$, $\boldsymbol{I}_\alpha \not\in \boldsymbol{I}_\beta$.

В данном случае имеем четыре максимальных интервала: — 00, 1—0, 11—, — 11; каждый из них образует гиперкуб размерности 1 (ребро).

Пересечения $\bigcap_i M_i^{\sigma_i}$, соответствующее максимальному интервалу множества M, называется *простой импликантой* этого множества.

Объединение простых импликант множества M называется *сокращенной* НФК *множества* M.

Количество первичных термов, образующих простую импликанту, называется *рангом простой импликанты*, а элементарное пересечение — *рангом соответствующего интервала.*

При выделении максимальных интервалов множества интервалов, которые имеют один и тот же ранг, разбивают на пояса, причем i-й пояс содержит интервалы, которым отвечают наборы с i единицами в каждом. Тогда выделение максимальных интервалов сводится к сравнению элементов только соседних поясов, номера которых отличаются на единицу. Если построенные интервалы не являются максимальными, то процесс сравнения продолжают.

Результаты сравнения для рассматриваемого случая приведены на рис.34.

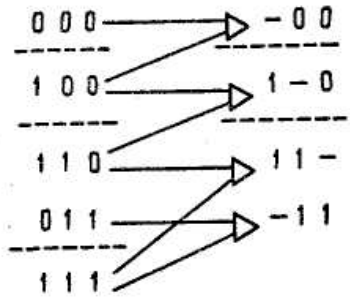


Рис. 34.

Сокращенная НФК множества $M(M_1, M_2, M_3)$ имеет вид

$$M(M_1, M_2, M_3) = \overline{M}_2 \cap \overline{M}_3 \cup M_1 \cap \overline{M}_3 \cup M_1 \cap M_2 \cup M_2 \cap M_3.$$

Построением сокращенной НФК множества M заканчивается первый этап метода.

Тупиковой НФК *множества* M называется такая НФК этого множества, которая при вычеркивании хотя бы одного первичного терма не определяет M.

Лемма 4. *Минимальная* НФК *множества* M *является тупиковой.*

Сложность минимальной НФК множества M нельзя уменьшить вычеркиванием первичного терма. Следовательно, эта форма есть тупиковой.

Лемма 5. *Тупиковая* НФК *множества* M *состоит из простых импликант этого множества.*

Если хотя бы одно пересечение соответствует интервалу множества M, не являющемуся максимальным, то это пересечение можно заменить простой импликантой вычеркиванием соответствующих первичных термов, не выходя из класса эквивалентных НФК (которые задают одно и то же самое множество) множества M, что противоречит определению тупиковой НФК.

Теорема 6. *Тупиковая* НФК *множества* M, *в том числе и минимальная* НФК**,** *содержится в сокращенной* НФК *этого множества.*

Тупиковая НФК множества M, в том числе и минимальная НФК, состоит, согласно лемме 5, из простых импликант. Сокращенная НФК множества M включает все простые импликанты. Следовательно, тупиковая

(минимальная) НФК множества M содержится в сокращенной НФК этого множества.

Согласно теореме 6, построение тупиковой НФК множества M сводится к покрытию двумерной таблицы.

Покрытием столбцов строками в двумерной таблице называется такое множество строк, при котором для каждого столбца найдется хотя бы одна строка из этого множества, на пересечении с которой этот столбец имеет единицу, причем при вычеркивании хотя бы одного элемента из этого множества строк указанное свойство не выполняется.

2. Построение и покрытия таблицы Квайна. *Таблица Квайна* - двумерная таблица, каждой строке которой взаимно однозначно отвечает максимальный интервал, столбцу - конституента, а на пересечении i-й строки и j-го столбца находится единица, если j-а конституента входит в i-й максимальный интервал, в противном случае клетку (i, j) не заполняют или ставят в ней 0.

Для рассматриваемого примера таблица Квайна имеет вид:

Таблица 4.

Максимальный интервал	Конституента				
	000	100	110	011	111
-00	1	1			
1-0		1	1		
11-			1		1
-11	•			1	1

Максимальный интервал называется *обязательным,* если найдется конституента, которая принадлежит ему и только ему. Множество обязательных интервалов образует *ядро покрытия.*

В данном случае ядром покрытия является множество { - 00, -11}, которое покрывает первый, второй, четвертый и пятый столбцы. Для образования покрытия можно взять либо вторую, либо третью строку. В результате получаем два покрытия:

{- 00, - 11, 1 - 0}, {- 00, - 11, 11 - };

каждое из них является минимальным и имеет сложность 6. Для определенности выберем первое из покрытий, которое соответствует минимальной НФК, задающей множество

$$M(M_1, M_2, M_3) = \overline{M}_2 \cap \overline{M}_3 \cup M_2 \cap M_3 \cup M_1 \cap \overline{M}_3.$$

В результате упрощения сложность $L(M)$ уменьшилась от 15 до 6.

Минимальная НФК находится в результате перебора всех покрытий, осуществляемого с помощью преобразования мультипликативно-аддитивноы формы в аддитивно-мультипликативную форму.

Для рассматриваемого примера идентифицируем четыре строки табл. 1.4 соответственно буквами *a*, *b*, c, *d*. Запишем множество строк, каждый элемент которого покрывает *j*-й столбец:

$$j=1\rightarrow A_1=\{a\},\ j=2\rightarrow A_3=\{a,b\},\ j=3\rightarrow A_3=\{b,c\},$$
$$j=4\rightarrow A_4=\{d\},\ j=5\rightarrow A_5=\{c,d\}.$$

Покрытием столбцов строками этой таблицы является множество строк, покрывающее все столбцы таблицы, и при удалении хотя бы одной из этих строк найдется непокрытый столбец. Следовательно, если каждое множество A_j представить в виде объединения ее элементов и найти пересечение всех множеств A_j, $\bigcap_j A_j$, то каждое пересечение в полученной аддитивной форме соответствует покрытию, а число всех покрытий равно числу различных пересечений в полученной аддитивно-мультипликативной форме:

$$\bigcap_j A_j = a\cap(a\cup b)\cap(b\cup c)\cap d\cap(c\cup d)=a\cap(b\cup c)\cap d =$$
$$=a\cap b\cap d\cup a\cap c\cap d.$$

Полученные пересечения $a\cap b\cap d$ и $a\cap c\cap d$ порождают два покрытия: {-00, 1-0, -11} и {-00, 11-, -11}; каждое из них соответствует минимальной НФК заданного множества *M*.

Дальнейшее уменьшение сложности выражения, которое определяет заданное множество, возможно, если из класса НФК перейти в класс скобочных форм Кантора (ДФК). Выражение, которое определяет множество *M,* называется *скобочной формой Кантора,* если кроме первичных термов и знаков операций объединения и пересечение у него входят скобки (,).

В рассмотренном примере сложность представления множества, равная 6, понижается до 5 в результате применения закона дистрибутивности пересечения относительно объединения

$$M(M_1, M_2, M_3) = \overline{M}_3\cap(M_1\cup\overline{M}_2)\cup M_2\cap M_3.$$

Преобразование мультипликативно-аддитивной формы в аддитивно-мультипликативную называется *методом Петрика,* который может быть определенный соответствующим *алгоритмом.*

Интуитивное (наивное) определение алгоритма. Совокупность правил, обладающих свойствами *массовости* (инвариантность относительно входной информации), *детерминированности* (однозначность применения этих правил на каждом шаге), *результативности* (получение после применения этих правил информации, являющейся результатом) и *элементарности* (отсутствует необходимость дальнейшего уточнения правил), называется ***алгоритмом***.

Индивидуальные тестовые задачи

Упражнение 1.

1. Построить функцию f: $A \to A$, где $A=\{0, 1\}$, которая не имеет обратной.

2. Какие из следующих функций есть отображениями:

а) f на R определяется таким образом: $\{(x, x^4): x \in R\}$;

б) f на R определяется таким образом: $\{(x^3, x): x \in R\}$;

в) f на R определяется таким образом: $\{(x, x^2): x \in R\}$;

г) f на R, f: x □ sin (x);

д) f на R, f: x а $1/x$;

е) f на Q, f: x а arcsin x;

ж) f: $A \to P(A)$ определяется таким образом: f: x а $\{x\}$;

з) f: $P(A) \to A$ определяется таким образом:

$f = \{(x, y): y \in$ I $\{a\}\}$, где a-фиксированный элемент из A?

3. Доказать, что если функция f инъективна, то существует f^{-1}.

4. Если функция f сюръективна, следует ли отсюда, что f^{-1} - отображение?

5. Построить пример, который показывает, что функция на $A=(-1, 0, 1)$, определенная как f: x а x^2, такая, что $f^{-1} \circ f \neq U_A$.

6. Пусть f: $A \to B$ и g: $B \to C$ — функции. Доказать, что:

а) если f и g инъективны, то $g \circ f$ инъективна;

б) если f и g сюръективны, то $g \circ f$ также сюръективна.

7. Пусть f: $A \to B$ и g: S→C — функции и g сюръективна.

Достаточно ли этого, чтобы обеспечить сюръективность $g \circ f$?

Упражнение 2.

1. Пусть f: $A \to B$ і g: $B \to C$; показать следующее:

а) если f сюръективна и g — отображение, то

$$G_{g^{-1}\square f} = C;$$

б) если f и g — биекции, то $(g^{-1} \circ f)^{-1} = f^{-1} \circ g$;

в) если $G_g \subseteq G_f$; то $(f \circ f^{-1} \circ g)C = G_g$.

Упражнение 3.

1. Построить биекцию между множествами Z×N и N×N.

2. Доказать, что если A и B — множества и $A \cup B$ конечна, то

$$|A \cup B| + |A \cap B| = |A| + |B|.$$

3. Доказать (от противного), что произвольная инъекция $N_m \to N_m$ ($m \in N$) является биекцией.

4. Пусть множество $\{0,1\}^N$ представляет собой последовательность $a_1, a_2, a_3, ..., a_n, ...$, где $a_i \in (0, 1)$. Доказать, что $|\{0,1\}^N| > |N|$.

5. Доказать (построением подходящих биекций и соображений по индукции), что если $A_1, ..., A_n$ — конечные множества, то

$$|A_1 \times ... \times A_n| = |A_1| \bullet ... \bullet |A_n|.$$

Упражнение 4.

1. Рассмотреть указанные ниже «определения» $\otimes$. Решить, правильно или нет каждое из них определяет бинарную операцию, и если так, то является ли операция коммутативной. Найти, если это возможно, единицу и обратный элемент к x. Предполагаются выполненными обычные свойства арифметики;

а) $x \otimes y = x - y$ на N;

б) $x \otimes y = (x \bullet y) - 1$ на Z;

в) $x \otimes y = \max(x, y)$ на N;

г) $x \otimes y = \sqrt{x^2 + y^2}$ на $\{x: 0 \le x, x \in R\}$;

д) $x \otimes y = x/y$ на $\{x: 0 < x, x \in R\}$.

2. Определить операцию φ на множестве $\{a, b, c\}$, как указано ниже. Проверить, что φ ассоциативна и коммутативна и найти единичный элемент.

φ	a b c
a	b c a
c	a b c
b	c a b

3. Предполагая обычные свойства операций +, -, • и / на R, доказать, что операция ψ, определеная на [1, ∞[следующим образом:

$$a \psi b = \frac{(a \cdot b) + 1}{a + b}$$

ассоциативна. Обосновать ответ.

Указание: не следует особенно обращать внимание на область определения.

4. Пусть $\otimes$ — ассоциативная операция на множества A с единицей e такая, что каждый элемент $a \in A$ обратим и обратный обозначается через a'. Показать, что

$$(a \otimes b)' = b' \otimes a',$$

5. Показать, что если $\otimes$ — ассоциативная операция на множества A с единицей e такая, что $a \otimes a = e$ для каждого $a \in A$, то $\otimes$ коммутативна.

6. Пусть $\otimes$ — ассоциативная операция на множестве A такая, что для любых $a,b \in A$, если $a \otimes b = b \otimes a$, то $a = b$. Показать, что каждый элемент A идемпотентен по отношению $\otimes$. Что можно сказать про $\otimes$, если операция имеет единицу?

Упражнение 5.

1. Доказать, что в совершенной НФК множества M при замене каждой операции объединения на симметрическую разность равенство не нарушается.

2. Установить, есть ли форма

$$M(M_1,M_2,M_3)=M_1 \square M_2 \cap \overline{M}_3 \cup M_1 \cap M_1 \cap M_3 \cup \overline{M}_2$$

совершенной.

3.Установить, является ли форма

$$M(M_1,M_2,M_3)=M_1 \cap \overline{M}_2 \cup M_1 \cap \overline{M}_3 \cup \overline{M}_2 \cap \overline{M}_3 \cup \overline{M}_1 \cap M_3$$

скращенной.

4. Выяснить, можно ли передавать последовательность символов в канале передач в виде мографа.

5. Минимизировать в классе нормальных форм Кантора множество M, заданое как объединение своих конституент:

$$M(M_1,M_2,M_3,M_4,)= \cup (0, 2, 7, 8, 11, 14, 15),$$

где десятичные числа являются числовыми эквивалентами двоичных векторов, которые определяют соответствующие конституенты этого множества.

6. Определить сложность минимальной скобочной формы множества M, заданного своей нормальной формой:

$M=M_l \square M_2 \mathrm{I}\ M_3 \mathrm{U}\ M_l \mathrm{I}\ M_2 \mathrm{I}\ M_3 \mathrm{U}\ M_1 \mathrm{I}\ M_3 \mathrm{I}\ M_4 \mathrm{I}\ M_6 \mathrm{U}\ M_2 \mathrm{I}\ \mathrm{I}\ M_4 \mathrm{I}\ M_5 \mathrm{I}\ M_6$.

7. Найти число тупиковых НФК множества

$M(M_1,M_2,M_3,M_4,)= \overline{M}_1 \mathrm{I}\ M_2 \mathrm{I}\ M_3 \mathrm{U}\ \overline{M}_1 \mathrm{I}\ \overline{M}_2 \mathrm{I}\ M_4 \mathrm{U}\ M_2 \mathrm{I}\ \mathrm{I}\ \overline{M}_3 \mathrm{U}\ \overline{M}_1 \mathrm{I}\ \overline{M}_4 \mathrm{U}\ M_1 \mathrm{I}\ \overline{M}_3 \mathrm{I}\ M_4$

8. Определить ранг (число конституент) множества

$M(M_1, M_2, ..., M_6)=(M_4 \mathrm{I}\ M_6 \mathrm{U}\ M_1 \mathrm{I}\ M_2) \mathrm{I}\ (M_1 \mathrm{I}\ M_3 \mathrm{U} M_5 \mathrm{I}\ M_6)$.

9. Найти минимальную НФК множества M, определенного в четырехмерном пространстве:

$M = \mathrm{U}$ (l - 00, -110, 0101, -0 -1, 0010, -01-, 0 - 0-).

10. Определить уменьшение мощности сигнатуры мографа G^M *(M),* определяющего мгожество M (M_1, M_2, M_3, M_4) = U (0, 4, 6, 7, 8, 9, 11, 13, 15), после минимизации в классе НФК**.**

3. 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.

3.1. Комбинаторные операции и функции

1. Имеются два основных типа операций над множествами. К первому типу относятся так называемые *алгебраические* операции, такие как объединение $A_1 \cup A_2 \cup \ldots \cup A_k = \bigcup_{i=1}^{n} A_i$ или пересечение $A_1 \cap A_2 \cap \ldots \cap A_k = \bigcap_{i=1}^{h} A_i$ нескольких множеств, а также теоретико-множественная разность $A \backslash B = \{x \mid x \in A,\ x \notin B\}$ 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 $A_1 \times A_2 \times \ldots \times A_k$, элементами которого являются всевозможные упорядоченные наборы вида $\langle a_1, a_2, \ldots, a_k \rangle$, где $a_1 \in A_1,\ a_2 \in A_2, \ldots,\ a_k \in A_k$ (иногда эту операцию рассматривают как ассоциативную, иногда — нет, это всегда ясно из контекста; в частости, использование $A^k \Rightarrow A \times A \times \ldots \times A$ $_k$ раз подразумевает обычно ассоциативный вариант) (в ассоциативном случае элементы $A_1 \times A_2 \times \ldots \times A_k$ называют также словами и записывают в виде $a_1 a_2 \ldots a_k$. Связную часть $a_i a_{i+1} \ldots a_{i+j}$ слова $a_1 a_2 \ldots a_k$ называют фрагментом этого слова; если $i = 1$, то префиксом, а если $i+j = k$, то суффиксом), булевская степень 2^A — множество всех подмножеств A: $2^A \Rightarrow \{X \mid X \subseteq A\}$ 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:

$$|A| = \sum_{x \in U} s_A(x) \tag{116}$$

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:

$$s_U^{(U)}(x) \equiv 1, \; s_\varnothing^{(U)}(x) \equiv 0, \tag{117}$$

$$s_{\bar{A}}^{(U)}(x) = 1 - s_A^{(U)}(x) = \overline{s_A(x)}, \tag{118}$$

$$s_{A \cap B}(x) = s_A(x) \cdot s_B(x) = s_A(x) \,\&\, s_B(x), \tag{119}$$

$$s_{A \setminus B}(x) = s_A(x) - s_{A \cap B}(x) = s_A(x) \,\&\, \overline{s_B(x)}, \tag{120}$$

$$s_{A \otimes B}(x) = s_A(x) + s_B(x) - 2s_{A \cap B}(x) = s_A(x) \oplus s_B(x). \tag{121}$$

Проверка соотношений (117) — (121) вполне тривиальна, как и переход к комбинаторным функциям на основе (116). 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 но *k*, этот результат обобщается на случай объединения любого конечного числа *k* множеств:

$$|A_1 \cup A_2 \cup \ldots \cup A_k| = \sum_{\sigma=1}^{k} (-1)^\sigma \sum_{(j_1, \ldots, j_\sigma)} |A_{j_1} \cap \ldots \cap A_{j_\sigma}|, \tag{122}$$

где внутренняя сумма распространяется на всевозможные σ-элементпые множества индексов. Метод подсчета на основе (122) называют *методом включения и исключения.*

4. Своеобразие обозначений для кардинальных операций находит оправдание в связи с соответствующими операциями над комбинаторными функциями:

$$|A_1 \times A_2 \times \ldots \times A_k| = |A_1| \cdot |A_2| \cdot \ldots \cdot |A_k|, \tag{123}$$

$$|2^A| = 2^{|A|}, \quad |A^B| = |A|^{|B|}. \tag{124}$$

Правило произведения (123) доказывается легко — при k=2 индукцией по мощности одного из множеств и затем индукцией по k. Заметим, что (123) 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 (124) 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 $C \subseteq B$ —

$$f(C) = \bigcup_{x \in C} \{f(x)\}. \tag{125}$$

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) Функция $\nu(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i \cdot 2^{n-i}$

устанавливает взаимно однотипное соответствие между множествами $A_n = \{0, 1\}^n$ и $B_n = \{0, 1, 2, ..., 2^n - 1\}$,

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) $\left|2^A\right| = \left|\{0,1\}^A\right| = \left|\{0,1\}^{|A|}\right| = 2^{|A|}$. 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 (124). Пусть $B = \{b_1, b_2, \ldots, b_n\}$, и пусть все функции A^B перечислены в левой части таблицы 1.3.

Таблица 1.3

x	b_1	$b_2 \ \ldots \ b_n$
$f_1(x)$	$f_1(b_1)$	$f_1(b_2) \ldots f_1(b_n) \to \langle f_1(b_1), f_1(b_2), \ldots, f_1(b_n)\rangle$
$f_2(x)$	$f_2(b_1)$	$f_2(b_2) \ldots f_2(b_n) \to \langle f_2(b_1), f_2(b_2), \ldots, f_2(b_n)\rangle$
$\ldots$	$\ldots$	$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$
$f_i(x)$	$f_i(b_1)$	$f_i(b_2) \ldots f_i(b_n) \to \langle f_i(b_1), f_i(b_2), \ldots, f_i(b_n)\rangle$
$\ldots$	$\ldots$	$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$
$f_m(x)$	$f_m(b_1)$	$f_m(b_2) \ldots f_m(b_n) \to \langle f_m(b_1), f_m(b_2), \ldots, f_m(b_n)\rangle$

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 $\left|A^B\right| = \left|A\right|^{|B|}$.

(г) Число размещений элементов $B = \{b_1, b_2, \ldots, b_n\}$ по k ящикам $A = \{a_1, \ldots, a_k\}$ 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 $A_i \Rightarrow A\backslash\{i\}$. 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

$$P_{n,n} = n! \Rightarrow n \cdot (n-1) \ldots 2 \cdot 1 \quad (0! \Rightarrow 1)$$

— одну из основных комбинаторных функций наряду с k^n, и для произвольного $k = \overline{1, n}$

$$P_{n,k} = n(n-1) \ldots (n-k+1) = \frac{n!}{(n-k)!}.$$

3.2. 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 $(\varnothing)n = \rho$. Пусть $\{\rho_1, \rho_2, \rho_3, \rho_4, \rho_5\}$ — отношения на $A = \{1, 2, 3\}$, графы которых изображены на рис. 35

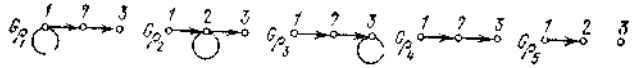


Рис. 35.

Если $R = \{\rho_1, \rho_2, \rho_3, \rho_4\}$, а $R_1 = R \cup \{\varnothing\}$, то имеем

$$(\varnothing)_R = \rho_4, \quad (\varnothing)_{R_1} = \varnothing, \quad (\rho_5)_{R1} = \rho_4.$$

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 $\rho' = s(\rho)$, и относительно s^*, если $\rho' = s^*(\rho)$.

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 $a, b \in A$ из $<a, b> \in \rho$ следует $<b, a> \in \rho$, и *антисимметричным,* если следует $<b, a> \notin \rho$; *транзитивным,* если для любых $a, b, c \in A$ 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. 1.6).

Мы будем рассматривать отношение строгого порядка «меньше», обозначаемое обычно символом «<» *(a < b).* Отношение нестрогого порядка определяется через строгий: «$\leq$» $\Rightarrow$ «<» $\cup$ « = ». 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.

Н е о б х о д и м о с т ь. 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»

$$\rho': \langle a, b\rangle \in \rho' \Rightarrow \text{«}\langle a, b\rangle \in \rho$$

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 *v* из примера (а.1.6) $\{0, 1\}^n$ на

линейно упорядоченное множество натуральных чисел позволяет продолжить частичный порядок до линейного (называемого *лексикографическим)* следующим образом:

$$\langle x_1, \ldots, x_n \rangle < \langle y_1, \ldots, y_n \rangle \Rightarrow v(x_1, \ldots, x_n) < v(y_1, \ldots, y_n).$$

(б) Пусть $E_k=\{0,1,\ldots,k\text{-}1\}$. 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

$$v(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_i \left(\prod_{j=i+1}^{n} m_j \right), \quad \text{где} \quad \prod_{j=n+1}^{n} m_j \Rightarrow 1,$$

и снова получим расширение до линейного порядка. Это прямое обобщение примера (а): там *v* 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.

3.3. 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 $\rho(a) \mid \rho(b) \neq \varnothing$, 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 $\langle b, d \rangle \in \rho$, 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 (см. доказательство теоремы 15.4) 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, $E_n = (0, 1, ..., n - 1)$. 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:

$$
\begin{aligned}
C_4^2 \qquad & P_{4,2}/\rho \\
\{0,1\} &\to \{\langle 0,1\rangle, \langle 1,0\rangle\}, \\
\{0,2\} &\to \{\langle 0,2\rangle, \langle 2,0\rangle\}, \\
\{0,3\} &\to \{\langle 0,3\rangle, \langle 3,0\rangle\}, \\
\{1,2\} &\to \{\langle 1,2\rangle, \langle 2,1\rangle\}, \\
\{1,3\} &\to \{\langle 1,3\rangle, \langle 3,1\rangle\}, \\
\{2,3\} &\to \{\langle 2,3\rangle, \langle 3,2\rangle\}.
\end{aligned}
$$

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 «биноминальный 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}:

$$\langle y_0, y_1, \ldots, y_{n-1}\rangle \to 0^{y_0}10^{y_1}1\ldots10^{y_{n-1}}1.$$

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.

$$(x_1 + \ldots + x_k)^n = \sum_{n_1+\ldots+n_k=n} \binom{n}{n_1 \ldots n_k} x_1^{n_1} \ldots x_k^{n_k},$$

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

$$B(n+1) = \sum_{i=0}^{n} \binom{n}{i} B(n-i). \qquad (126)$$

Соотношение (126) получается так. Пусть $\mathfrak{B}_{n+1}$ —множество всех разбиений E_{n+1} и $n \in A \subseteq E_n$. Через $\mathfrak{B}^A_{n+1}$ обозначим подмножество тех разбиений, у которых A есть класс разбиения, содержащий элемент n. Если $|A|=k$, то класс $\mathfrak{B}^A_{n+1}$ определяется выборкой (сочетанием!) из E_n по k—1 элементов, так как k-м элементом A является всегда n. Легко понять, что $\mathfrak{B}^A_{E_n}$ находится во взаимно однозначном соответствии с $\mathfrak{B}_{E_n \setminus A}$, а различные классы $\mathfrak{B}^A_{n+1}$ составляют в совокупности $\mathfrak{B}_{n+1}$:

$$\mathfrak{B}_{n+1} = \bigcup_{k=0}^{n} \bigcup_{A \in C_n^k} \mathfrak{B}^A_{n+1}, \tag{127}$$

откуда

$$|\mathfrak{B}_{n+1}| = \sum_{k=0}^{n} \sum_{A \in C_n^k} |\mathfrak{B}^A_{n+1}| = \sum_{k=0}^{n} \binom{n}{k} |\mathfrak{B}_{E_n \setminus A}|$$

и (126) доказано.

(д) Сходными рассуждениями получаются рекуррентные соотношения для комбинаторной функции $S(n, k)$ — числа разбиений E_n ровно на k непустых классов:

$$S(n+1, k) = \sum_{j=k-1}^{n} \binom{n}{j} S(j, k-1), \tag{128}$$

$$S(n+1, k) = S(n, k-1) + kS(n, k). \tag{129}$$

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:

$$S(n, k) = \frac{1}{k!} \sum_{j=0}^{k} (-1)^j \binom{k}{j} (k-j)^n. \tag{130}$$

Это можно доказать, например, с помощью (129) 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 A 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 A — функцию

$$F_A(x_1, \ldots, x_k) = \sum_{a \in A} x_1^{P_1(a)} \ldots x_k^{P_k(a)}. \qquad (131)$$

Функция (131) называется *производящей функцией* множества A по признакам $P_1, \ldots, P_k$. Переменные и операции, фигурирующие в (131), 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* (131),— *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

$$\langle a_1, a_2 \rangle \in \rho \Rightarrow \langle P_1(a_1), \ldots, P_k(a_1) \rangle = \langle P_1(a_2), \ldots, P_k(a_2) \rangle \qquad (132)$$

есть отношение эквивалентности на A. В последнем случае эквивалентным элементам A соответствуют одинаковые слагаемые в (131), поэтому после приведения подобных членов в (131) производящая функция может сохранить информацию о мощности классов эквивалентности (132) 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 $F_n(x) = \dfrac{F_{n-1}(x)}{(1-x)}$ и ввиду $F_1(x) = \sum\limits_{k=0}^{\infty} \binom{k}{0} x^k =$

$= \sum\limits_{k=0}^{\infty} x^k = \dfrac{1}{(1-x)}$ 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 (126), 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 $B'(x)=e^x \cdot B(x)$ с начальным условием $B(0)=1$. Интегрируя, получаем

$$B(x)= e^{(e^x-1)} = \frac{1}{e} \cdot e^{(e^x)} \qquad (133)$$

Разложим теперь (133) 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:

$$B(n) = \frac{1}{e} \sum_{k=0}^{\infty} \frac{k^n}{k!}. \qquad (134)$$

Представление о порядке роста $B(n)$ дают следующие оценки.

Теорема 1. $B(n) \le n!$

Доказательство. Доказательство проведем индукцией по n. Для $n=0$ имеем $B(0) = 1 = 0!$ Пусть для 1, 2,..., n утверждение справедливо. Тогда при $n \ge 2$ 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 $n > n_0$

$$B(n) > n^{n\left(1-\frac{1}{\log_m n}-\frac{\log_m \log_m n}{\log_m n}\right)}. \qquad (135)$$

Согласно (134) 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

$$f_i(n) = \sum_{j=1}^{m} \sum_{k=1}^{b_j} a_{ijk} f_j(n - d_{ijk}) \qquad (i = \overline{1, m}, d_{ijk} > 0) \qquad (136)$$

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 (136) на x^n и суммируя по $n = 1, 2,...$, получаем следующую систему линейных функциональных уравнений для производящих функций $F_i(x)$:

$$F_i(x) - c_i = \sum_{j=1}^{m} P_{ij}(x)\, F_j(x) \qquad (i = \overline{1, m}), \tag{137}$$

где $P_{ij}(x) = \sum_{k=1}^{b_j} a_{ijk} x^{d_{ijk}}$. 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 $q_j(t) \Rightarrow \sum_{l=1}^{k_j} \binom{l+t-1}{l-1} A_{jl}$ — 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

$$f_i(n) = c_i' \sum_{r=0}^{N} \alpha_r \cdot \sum_{j=1}^{s} q_j(n-r)\, x_j^{-(n-r)} = \sum_{j=1}^{s} \pi_j(n)\, x_j^{-n}, \tag{138}$$

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 (138) 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

$$f_1(n) = \frac{1}{833}\{(756n + 788)\cdot 2^n + 45\cdot 4^{-n}\},$$

$$f_2(n) = \frac{1}{49}(189n + 29)\cdot 2^n + \frac{20}{49}\cdot 4^{-n},$$

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 (1.231) 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 (131) взаимно однозначное.

Отметим еще две интерпретации (131) в классе функций алгебры логики, для которых также выполняется требование единственности представления в форме ∑П: *полиномы Жегалкина* и *функции Яблонского*. Для функций Яблонского правая часть (131) 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 $\langle a, b\rangle \in \rho \Rightarrow$ «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 $\langle s_1(x), ..., s_n(x)\rangle$, 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. 36 построены диаграммы Венна для n = 1, 2, 3.

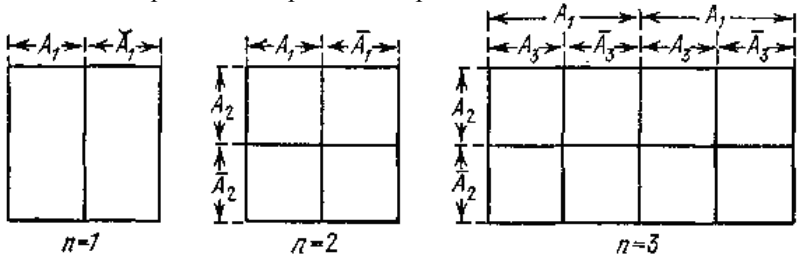


Рис. 36.

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 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. 37).

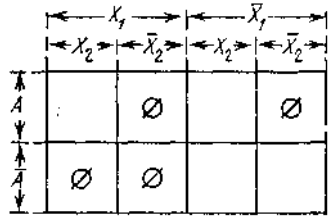


Рис. 37.

Так как условие $X \subseteq Y$ равносильно $X \mid \overline{Y} = \varnothing$, 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 $\overline{A}\ \overline{X}_1 X_2$ и $\overline{A}\ \overline{X}_1\ \overline{X}_2$. Число таких размещений равно $2^{|A|} \cdot 2^{|\overline{A}|} = 2^{|A|+|\overline{A}|} = 2^{|U|} = 2^n$.

3.4. 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. 38 (ферзь атакует любую клетку, расположенную с ним на одной линии — вертикали, горизонтали или диагонали).

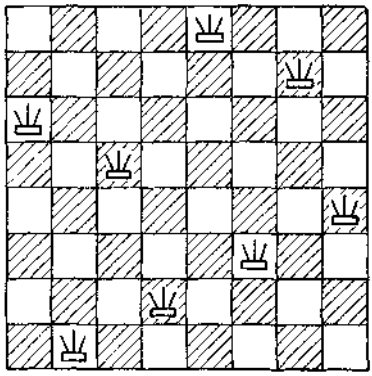


Рис. 38.

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 во просам 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 $G = \langle V, R \rangle$ —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* $\overline{W} = V \setminus W \in H'_T(G)$.

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 конечныхгиперграфов из теоремы 1 следует, что

$$\alpha(G) = |V| - \alpha'(G). \tag{139}$$

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 $M_i \Rightarrow \max\{x_i \mid X \in H'\} < \infty$ 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 $X \leq M < Y$, 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 $t = 1, 2, \ldots$, 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*

$$f_G(x_1, \ldots, x_n) = \underset{r \in R}{\&} \left(\bigvee_{v_i \in r} x_i \right) \tag{140}$$

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 (140) 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. 39, имеем

$$f_G(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) =$$
$$= (x_1 \vee x_2)(x_1 \vee x_4)(x_2 \vee x_3)(x_2 \vee x_7)(x_3 \vee x_4)(x_4 \vee x_5)$$
$$(x_4 \vee x_8)(x_5 \vee x_6)\&(x_5 \vee x_8)(x_6 \vee x_8)(x_6 \vee x_7) =$$
$$= x_2x_4x_5x_8 \vee x_2x_4x_6x_8 \vee x_2x_4x_5x_7x_8 \vee x_1x_2x_3x_5x_6 \vee$$
$$\vee x_1x_3x_5x_6x_7 \vee x_1x_3x_4x_5x_7x_8 \vee x_1x_3x_4x_6x_7x_8.$$

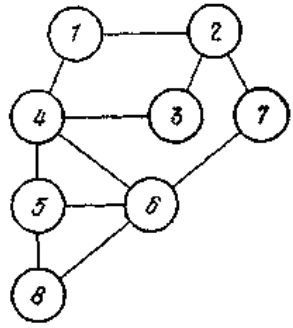


Рис. 39.

Таким образом, $\alpha(G) = 4$ и $|H_T| = |H'_T| = 7$. Множества H_T и H'_T перечислены в таблице 4.

Таблица 4

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 $X = \langle x_1,..., x_n\rangle$ 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

$$\sum_{k=0}^{n} h_k \binom{n}{k}^{-1} \leqslant 1. \tag{141}$$

А так как $\binom{n}{k}$ достигает максимума при $k = [n/2]$ (целая часть $n/2$), умножая (141) 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 $f_G(x_i,..., x_n)$:

$$|H_т(G)| = l_c(f_G) \leqslant \binom{n}{[n/2]} \sim \frac{2^n}{\sqrt{2^{-1}\pi n}}. \qquad (142)$$

Если G — граф, то оценка (142) может быть усилена:

$$|H_т(G)| = l_c(f_G) \leqslant (\sqrt[3]{3})^n, \quad \sqrt[3]{3} \approx 1{,}442\ldots \qquad (143)$$

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 $|W| \infty$.

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 $W=0 \cdot W_0 \cup \ldots \cup (n—1)W_{n-1}$, 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 $W = W_1 \cup W_2 \cdot A$, где $W_2 \cdot A \subseteq A^{d(W)}$, $d(W_1)<d(W)$. 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*

$$\sum_{w\in W} n^{-|w|} \leqslant 1. \qquad (144)$$

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 (144). Теорема доказана, так как для бесконечного W утверждение следует из его справедливости для любого конечного подмножества.

Пусть $W \in H(G)$ и $W=W_1 \cup \ldots \cup W_i, \ldots$, где $W_i = W \cap A^i$. Тогда спектр длин W можно задать в виде последовательности $D'(W) \Rightarrow <\sigma_1, \sigma_2, \ldots, \sigma_i \ldots>$, где $\sigma_i = |W_i|$, а неравенство Мак-Миллана переписать в форме

$$\sum_{i=1}^{\infty} \sigma_i n^{-i} \leqslant 1. \qquad (145)$$

Теорема 5. *Если выполнено* (145), *то существует* $W \in H(G)$, *удовлетворяющий условию* $D'(W) = <\sigma_1, \sigma_2, \ldots, \sigma_i \ldots>$.

Доказательство. Искомый префиксный код строится поэтапно: выбираем произвольно $W_1 \subseteq A$ так, что $|W_1| = \sigma_1$, и, после того как построены $W_1, W_2, \ldots, W_{k-1}$ на k-м этапе выбиваем произвольно $W_k \subseteq \Delta_k \Rightarrow A^k \setminus (W_1 \cdot A^{k-1} \cup \ldots \cup W_{k-1} \cdot A^1)$ 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 $k = 1, 2, \ldots$ Но мы имеем

$$|\Delta_k| = \left| A^k \right| - \sum_{i=1}^{k-1} |W_i| \cdot \left| A^{k-i} \right| = n^k - \sum_{i=1}^{k-1} \sigma_i \cdot n^{k-i} =$$

$$= n^k \left(1 - \sum_{i=1}^{k-1} \sigma_i n^{-i} \right) \geqslant \sigma_k$$

(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.

Т е о р е м а 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

$$W_2 \subseteq A^{d(W)} \setminus W_1 A^+ \text{ и } W' \Rightarrow W_1 \cup (A^{d(W)} \setminus W_1 A^+)$$

— 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*

$$\sum_{i=1}^{m} n^{-d_i} = 1. \qquad (146)$$

Доказательство. ***Необходимость***. Если $W' \in H_T(G)$,то по теореме 4 выполнено неравенство Мак-Миллана, которое мы возьмем в форме (145):

$\sum_{i=1}^{k} \sigma_i n^{-i} \leqslant 1$. 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, $W' = W_1 \cup \ldots \cup W_{k-1} \cup \Delta_k$ есть независимое расширение W, что противоречит предположению о тупиковости W. Поэтому выполняется (146).

Достаточность. Из (146) по теореме 5 следует существование W, для которого $W \in H(G)$ и $D(W) = \{d_1, \ldots, d_m\}$. Но, согласно теореме 4, $W \in H_T(G)$, так как всякое его расширение нарушает необходимое условие (144). Теорема доказана.

Описание структуры множеств $W \in H(G)$ дается так называемой структурной функцией

$$f_w(z_0, \ldots, z_{n-1}) \Rightarrow \sum_{w \in W} z_0^{|w|_0} z_1^{|w|_1} \ldots z_{n-1}^{|w|_{n-1}}, \qquad (147)$$

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 (147) 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 Густь $f=f_W$ 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

$$f_W - 1 = (z_0 + \ldots + z_{n-1} - 1) \cdot f_1 + (z_0 + \ldots + z_{n-1} - 1) f_\alpha = (z_0 + \ldots + z_{n-1} - 1)(f_{\pi(W)} + 1),$$

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 $g = \text{const}$, $f = (z_0 + ... + z_{n-1}) \cdot g - (g - 1)$ 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 $\deg(f) = 1 + \deg(\bar{g}) \leqslant t - 1$.

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 $\deg(\bar{f})<t$, по предположению индукции существует $\bar{W} \in H_T(G)$ такой, что $\bar{f} = f_{\bar{W}}$. Пусть $\bar{W} = \{w_1, ..., w_N\}$, а $\bar{f}$ 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

$$f_W = f_{\overline{W}} - f_{\overline{W}_h} + f_{\overline{W}_h \cdot A} = \bar{f} - h + (z_0 + \ldots + z_{n-1})\, h = f$$

и $g = 1 + f_{\pi(W)}$, так как $g = \bar{g} + h$ 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:

$$f = z_0 z_1 + 2z_0^2 z_1 + z_0 z_1^2 + z_0^3 + 2z_0 z_1^3 + z_1^4 + z_0^3 z_1^2 + z_0^2 z_1^3.$$

Выполнение (а), (б) очевидно. После деления f—1 на $z_0 + z_1 - 1$ получаем

$$f - 1 = (z_0 + z_1 - 1)\left(z_0^2 z_1^2 + z_1^3 + z_0 z_1^2 + \right.$$
$$\left. + z_1^2 + z_0^2 + z_0 z_1 + z_0 + z_1 + 1\right),$$

т. е. (в) тоже выполнено и

1) $h = h_0 = z_0^2 z_1^2$, а $\bar{f} - 1 = (z_0 + z_1 - 1)\left(z_1^3 + z_0 z_1^2 + \right.$
$\left. + z_1^2 + z_0^2 + z_0 z_1 + z_0 + z_1 + 1\right)$.

Следующие шаги выполняются по одному и тому же шаблону:

2) $h_1 = z_1^3 + z_0 z_1^2$, $\bar{\bar{f}} - 1 = (z_0 + z_1 - 1)\left(z_1^2 + z_0^2 + \right.$
$\left. + z_0 z_1 + z_0 + z_1 + 1\right)$;

3) $h_2 = z_1^2 + z_0^2 + z_0 z_1$; $\bar{\bar{\bar{f}}} - 1 = (z_0 + z_1 - 1)(z_0 + z_1 + 1)$;

4) $h_3 = z_0 + z_1$; $\bar{\bar{\bar{\bar{f}}}} - 1 = z_0 + z_1 - 1$.

Отсюда $\bar{\bar{\bar{\bar{f}}}} = z_0 + z_1 = f_{\{0,\,1\}}$ и $W_4 = \{0,\ 1\}$. Теперь остается от W_4 вернуться к искомому коду, проходя шагн 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) $f_W(z_0, \ldots, z_{n-1}) \leqslant 1$ 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) $f_W(z_0, \ldots, z_{n-1}) \leqslant 1$ 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

$$\chi(\overline{G}_\rho) = \alpha(G_\rho) = |A/\rho|. \qquad (148)$$

3.5. 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: xλ = λx = 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

$$\varphi(x_1 \circ x_2) = \varphi(x_1) * \varphi(x_2). \qquad (149)$$

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 *<x, y>* ∈ ρ следует, что для любого z: *<zz, xy>* ∈ ρ; *стабильным справа,* если из *<x, y>* ∈ ρ следует, что для любого z: *<xz, yz>* ∈ ρ. 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>* ∈ ε_φ ⇒ φ(*a*) =φ(*b*) *есть конгруэнтность. Если* ρ *есть*

отношение конгруэнтности на полугруппе П, то П/ρ есть полугруппа, в которой X ° 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 a, 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

$$\forall x_1, \ldots x_N \quad \left[\mathop{\&}_{i=1}^{k} \alpha_i(x_1, \ldots, x_N) = \beta_i(x_1, \ldots, x_N) \right] \to \to \alpha(x_1, \ldots, x_N) = \beta(x_1, \ldots, x_n), \qquad (150)$$

с помощью которых формулируются формальные, синтаксические правила вывода соотношений.

В любой полугруппе выполняется квазитождество

$$\forall x_1, x_2, y, z \qquad (x_1 = x_2) \to (yx_1z = yx_2z). \qquad (151)$$

Пусть *s* — правило вывода соотношений согласно (151), т. е. соотношение вида $yx_1z = yx_2z$ выводится из $x_1=x_2$ по правилу *s*. Если $s^*(\psi) = \rho$, то говорят, что $П = П<A \mid \psi>$ — представление полугруппы *П* образующими *A* 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'*

$$(yx = y'x' \;\&\; yz = y'z' \;\&\; uz = u'z') \to ux = u'x'. \qquad (152)$$

В дальнейшем мы будем использовать для соотношений вида α=β 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 (151), 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,

$$\frac{x}{x'} \leftarrow \frac{y}{y'} \rightarrow \frac{z}{z'} \leftarrow \frac{u}{u'},$$

подразумевая, что $\frac{\alpha}{\beta} \to \frac{\gamma}{\delta}$ 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,

$$\left[\frac{ab}{aba} = \left[\frac{c}{ca} = \frac{\lambda}{a}, \quad \left[\frac{ab}{ac} = \left[\frac{bb}{bc} = \frac{b}{c}.\right.\right.\right.\right.$$

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 $\alpha=(01)(010)(010)(01)$, где $\alpha=0101001001$, 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

$$J(V, \alpha) \Rightarrow \{j \mid \alpha = v_j\alpha_j\} \tag{153}$$

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:

$$F_V(\alpha) = \bigcup_{j \in J(V,\alpha)} b_j F_V(\alpha_j). \tag{154}$$

Например, для $V=\{01, 100, 010, 1001\}$ и $\alpha = 0101001001$ находим

$$F_V(\alpha) = b_1F_V(01001001) \cup b_3F_V(1001001) =$$
$$= b_1b_1F_V(001001) \cup b_1b_3F_V(01001) \cup$$
$$\cup b_3b_2F_V(1001) \cup b_3b_4F_V(001) = b_1b_2b_1F_V(001) \cup$$
$$\cup b_1b_3b_3F_V(01) \cup b_3b_2b_2F_V(1) \cup \{b_3b_2b_4\} =$$
$$= \{b_1b_3b_3b_1,\ b_3b_2b_4\}.$$

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 $|F_V(\alpha)| \le 1$). 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 $\alpha, \beta, \gamma, \delta \in V^+$ 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 $\alpha, \beta, \gamma, \delta \in V^+$. 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:

$$\text{для любого } x\text{: } V^+x \cap V^+ \neq \varnothing \text{ и } xV^+ \cap V^+ \neq \varnothing \text{ влечет } x \in V^*. \quad (155)$$

Рассмотрим еще два частных случая условия (155):

$$\text{если } \alpha, \beta, \gamma \in V^+,\ x\alpha = \gamma,\ \beta x = \gamma, \text{ то } x \in V^* \quad (156)$$

(эквивалентно: $xV^+ \cap V^+ \cap V^+x \neq \varnothing$ влечет $x \in V^*$), и

$$\text{если } \alpha, \beta, \gamma \in V^+,\ \alpha x = \beta,\ x\alpha = \gamma, \text{ то } x \in V^* \quad (157)$$

(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 (155) и может служить критерием единственности разложения на множители. Пусть, например, выполнено (157) и не выполнено (155): $\alpha, \beta, \gamma, \delta \in V^+$, $\alpha x = \beta$, $x\gamma = \delta$ 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 (157). Аналогично проверяется равносильность (156) и (155).

Условие префиксности кода V — достаточное для единственности разложения слов V^+ на простые множители. Это легко вывести непосредственно из (154) и формально, записав его в виде

$$Vx \cap V^+ \neq \varnothing \to x \in V^* \quad (158)$$

и сравнив с (155). Оно не является необходимым.

Теорема 1 и эквивалентные ей критерии (156), (157) 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 $A = \{a_1, \ldots, a_n\}$ — 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

$$f_{V^N}(z_1, \ldots, z_n) \leqslant (f_V(z_1, \ldots, z_n))^N. \tag{159}$$

Учитывая, что строгое неравенство в (159) 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 $B = \{b_1 .. ., b_m\}$ — алфавит имен образующих в том же порядке. Пусть f - естественный гомоморфизм B^+ на V^+:

$$f(b_i) = v_i \quad (j = \overline{1, m}), \quad f(b_{i_1} \dots b_{i_k}) = f(b_{i_1}) \dots f(b_{i_k})$$

$$\text{и} \quad R(V) = \left\{ \frac{\alpha}{\beta} \middle| f(\alpha) = f(\beta) \right\}.$$

R(V) есть отношение конгруэнтности на B^+, соответствующее гомоморфизму *f*, поэтому $П = П\langle B|R(V)\rangle$. Пусть

$$R_1(V) = R(V) \setminus R(V)^2,$$

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 $f\left(\frac{\alpha_1}{\beta_1}\right) > 0$ 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 $f(\alpha_1\alpha_2) = f(\alpha_1)f(\alpha_2) = f(\beta_1)f(\beta_2) = f(\beta_1\beta_2)$ 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. $П = П\langle B|R_1(V)\rangle$.

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:

$$f_i(x_1, \dots, x_m) = g_i(x_1, \dots, x_m), \quad i \in J, \tag{160}$$

— система уравнений в словах (f_i, g_i — слова в алфавите неизвестных). Набор слов $X_0 = \langle x^0_1, \dots, x^0_m \rangle$ называется решением системы (160), если при подстановке левая и правая части каждого уравнения совпадают графически. Набор слов называется *унтер-решением* системы (160), 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 F(ρ) — 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 ε(α) примитивный корень этого слова, т. е. кратчайшее из слов β, для которых верно $\alpha\{\beta\}^+$. В представлении $\alpha = \varepsilon(\alpha)^\tau$ натуральное число τ = τ(α) — показатель слова α. Слово α называется простым, если ε(α)=α, в противном случае — периодическим (τ(α) > 1). Очевидно, что для любого слова α: ε(ε(α)) = ε(α), т. е. ε(α) — простое (если бы $\varepsilon(\alpha) = \beta^i$, $i \geqslant 2$, то имели бы $\alpha = \beta^{ik}$, $|\beta| < |\varepsilon(\alpha)|$).

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 |δ| = *k*. При *k* = 1 очевидно, что α= β = γ. Пусть утверждение верно при *k* < *m* и |δ| = *m*. Если α≠ β, то α и β находятся в отношении префиксности, скажем, α=β•ω. 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:

$$\varepsilon(\alpha^i) = \varepsilon(\alpha) \text{ для любого слова } \alpha \text{ и натурального } i; \quad (161)$$

$$\text{если } \alpha = \beta^i, \text{ то } \beta = \varepsilon(\alpha)^{\tau(\alpha)/i}, \ \tau(\alpha) = i \cdot \tau(\beta). \quad (162)$$

(l.261) докажем индукцией по |α|. При |α| = 1 имеем $\varepsilon(\alpha^i)= \varepsilon(\alpha)= \alpha$.

Пусть |α| = *k*. Тогда $\alpha^i = \left(\varepsilon\left(\alpha^i\right)\right)^{\tau(\alpha^i)}$, $\alpha= \varepsilon(\alpha)^{\tau(\alpha)}$ из $\alpha^{i+1} = \alpha \cdot \alpha^i = \alpha^i\alpha$ и по теореме 1, учитывая, что $\varepsilon(\alpha^i)$ и ε(α) — простые слона, заключаем, что $\varepsilon(\alpha^i)= \varepsilon(\alpha)$. (161) доказано.

Если $\alpha = \beta^i$, то $\varepsilon(\alpha) = \varepsilon(\beta^i) = \varepsilon(\beta)$, откуда

$$\alpha = (\varepsilon(\alpha))^{\tau(\alpha)} = (\varepsilon(\alpha)^{\tau(\beta)})^i = \varepsilon(\alpha)^{i\tau(\beta)}$$

и (162) доказано.

Пусть $\alpha = a_1a_2 \ldots a_N$. Через $\alpha^{(i)} = a_ia_{i+1} \ldots a_{i+N-1}$ 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*

$$\varepsilon(\alpha^{(i)}) = (\varepsilon(\alpha))^{(i)}, \quad (163)$$

$$C(\alpha) = \{(\varepsilon(\alpha^{(j)}))^{\varepsilon(\alpha)} | j = \overline{1, |\varepsilon(\alpha)|}\}, \quad (164)$$

$$|C(\alpha)| = |\varepsilon(\alpha)|. \quad (165)$$

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 $i = 1, 2,\ldots,|V|$, 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* $V = \{v_1, \ldots, v_m\}$, $W = \{w_1, \ldots, w_n\}$ *и для* $i = \overline{1, m}$ $v_i = \varphi_i(w_1, \ldots, w_n)$, *причем каждое* x_j $(j = \overline{1, n})$ *входит по крайней мере в одно из слов* $\varphi_i(x_1, \ldots, x_n)$, $i = \overline{1, m}$. *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* $|v_1| = |v_2| = \ldots = |v_m| = 1$, *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 множества 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 $|\alpha|$. 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 пазывается 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* (160) *в* $(m - 1)$*-буквенном алфавите, а* $\mathfrak{V}$ *— множество всех ее унтер-решений в* m*-буквенном алфавите. Тогда множество всех решений* (160) *и множество всех ее унтер-решений являются соответственно производными от* $\mathfrak{D}$ *и* $\mathfrak{V}$.

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) α=β — пустое слово и $i \neq j$;

(2) α или β отлично от λ, тогда $f(\alpha)$ и $f(\beta)$ находятся в отношении префиксности, скажем, $f(\alpha) = f(\beta) \cdot \gamma$ 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

$$R_1(V) = \left\{ \frac{x}{z}\left(\frac{y}{z}\right)^i \frac{y}{x} \right\}_{i=0}^{\infty} \tag{166}$$

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 $\alpha = (ab)^i a$ для некоторого натурального числа *i* Индукцией по *i* находим $F_V\left((ab)^i a\right) = \{z^j xy^{i-j}\}_{j=0}^{i}$. Из этого множества можно выбрать, очевидно, только одну несократимую пару $xy^i = z^i x$, откуда следует (166).

2) Любое из соотношений (166) выводится из $xy = zxx$:

$$xy^{i+1} = zxy^i = z^2xy^{i-1} = \ldots = z^ixy = z^{i+1}x.$$

В результате получаем

$$V^+ = \Pi\langle\{x,\ y,\ z\} \mid xy = zx\rangle.$$

(б) Рассмотрим обратнуюзадачу: описать все решения уравнения

$$xy = zx. \tag{167}$$

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 $\langle\lambda, \lambda, \lambda\rangle$, $\langle\alpha, \lambda, \lambda\rangle$ и $\langle \lambda, \alpha, \alpha\rangle$; решения, соответствующие $x_1 = \lambda$ в первом случае и $z_1 = \lambda$ во втором случае — $\langle \alpha, \alpha, \alpha\rangle$, где α — произвольное слово).

В первом случае подстановка в (167) дает $zx_1y = zzx_1$ и после сокращения на z слева — $x_1y = zx_1$. Следовательно, решение, соответствующее этому случаю, получается из некоторого решения $\langle x_h\ y,\ z\rangle$ того же уравнения (167), но с меньшей суммой длин слов, операцией

$$\langle x_1,\ y,\ z\rangle \to \langle zx_1,\ y,\ z\rangle. \tag{168}$$

Во втором случае получаем аналогично $xy = xz_1x$, откуда $y = z_1x$, т. е. решение имеет вид

$$\langle x,\ z_1x,\ xz_1\rangle, \tag{169}$$

где— произвольные непустые слова. Проверкой убеждаемся, что (169) является решением (167) при любых $x,\ z_1$ следовательно, любое решение (167) получается из (169) с произвольными $x,\ z_1$ операциями (168):

$$\langle x,\ z_1x,\ xz_1\rangle \to \langle (xz_1)x,\ z_1x,\ xz_1\rangle \to \ldots$$
$$\ldots \to \langle (xz_1)^i x,\ z_1x,\ xz_1\rangle \to \ldots$$

Итак, общее решение можно записать в виде

$$\langle (\alpha\beta)^i\alpha,\ \beta\alpha,\ \alpha\beta\rangle,\quad \langle \alpha,\ \lambda,\ \lambda\rangle, \tag{170}$$

где α, β— произвольные слова, i — произвольное неотрицательное целое (< α, α, α > получается из (170) 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 (1.248) $\alpha\left(G_\rho^{(N,n)}\right) = \left|A^N/\rho\right|$. Получим явное выражение для $\alpha(G_\rho^{(N,n)})$.

Пусть $P_n(N)$ — множество всех простых слов A^N. Тогда согласно п. 5 имеем

$$A^N = \bigcup_{d\,|\,N} \left[\alpha^{N/d}\,\middle|\,\alpha \in P_n(d)\right] \tag{171}$$

(объединение по всем делителям d числа N, включая 1 и N). Так как каждое слово α представимо в виде $\alpha = \varepsilon(\alpha)^{r(\alpha)}$, где $\varepsilon(\alpha)$ — простое (примитивный корень), единственным способом, множества в правой части (171) 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 $\mu(N)$ вычисляется так: $\mu(1) = 1$, 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:

$$x_1 x_2 = x_3 x_1,$$

$$x_1 x_4 x_1 = x_3 x_2$$

— 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 2 x_4 быть не может: 01 и 10 входят в X, $x_4 \Rightarrow 11$ влечет $x_4 = x_1^2$, 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:

$$\begin{array}{llll} 000 - x_1x_4x_2 = x_3x_4x_1, & & 100 - x_4x_1 = x_3x_2, \\ 001 - x_1x_4 = x_3x_2, & & 101 - x_4 = x_1x_2, \\ 010 - x_1x_3 = x_3^2, & & 110 - x_4 = x_1x_3, \\ 011 - x_4 = x_2x_1, & & 111 - x_4 = x_1^3. \end{array}$$

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:

⟨1, 2, *i*, 2⟩,	⟨2, 2, 1, *i*⟩,
⟨1, *i*, 2, 2⟩,	⟨2, *i*, 1, 2⟩,
⟨2, 1, 2, *i*⟩,	⟨*i*, 2, 1, 2⟩,
⟨2, 1, *i*, 2⟩,	⟨2, 2, *i*, 1⟩,
⟨*i*, 1, 2, 2⟩,	⟨2, *i*, 2, 1⟩,
	⟨*i*, 2, 2, 1⟩.

Аналогичными рассуждениями можно убедиться, что ни один из них нереализуем. Таким образом, матрице 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 $h_i \Rightarrow f(a_1 \dots a_i) = w(a_1) + \dots + w(a_i)$ и $h(\alpha) = \{h_i\}_{i=1}^N$. Тогда

$$h_i\left(\alpha^{(j)}\right) = \begin{cases} f(a_j \dots a_{j+i-1}), & \text{если } j+i-1 \leqslant N, \\ f(a_j \dots a_N a_1 \dots a_{j+i-N-1}), & \text{если } j+i-1 > N, \end{cases}$$

$$= \begin{cases} f(a_1 \dots a_{j+i-1}) - f(a_1 \dots a_{j-1}), \\ f(a_1 \dots a_N a_1 \dots a_{i+j-N-1}) - f(a_1 \dots a_{j-1}) \end{cases}$$

$$= h_{i+j-1} - h_{j-1}.$$

Отсюда

$$h\left(\alpha^{(j)}\right) = \{h_{i+j-1} - h_{j-1}\}_{i=1}^N,$$

причем первое слагаемое при $i = 1, 2, \dots, N$ 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

$$f(i) = w(i) = 1 - i \cdot \frac{N}{k} \quad (i = \overline{0, n-1}), \quad \|\alpha\| \Rightarrow \sum_{i=1}^{N} a_i.$$

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

$$\frac{|\alpha_1|}{\|\alpha_1\|} = \frac{N}{k} = \frac{|\alpha_2|}{\|\alpha_2\|}, \quad |\alpha_i| = p_iN, \quad \|\alpha_i\| = p_ik \quad (i = 1, 2).$$

Тогда $p_i \geq 1$, $|\alpha| = (p_1 + p_2)N$, $\|\alpha\| = (p_1 + p_2)k$ — противоречие и теорема доказана.

3.6. 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 — наименьший класс языков, содержащий все конечные языки и замкнутый относительно основных комбинаторных операций над языками.**

2. Основной способ задания грамматик регулярных множеств — графический. Пусть $G = <Q, R>$ — 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 $\Gamma = <G, F, P>$ называется *источником,* $G = G(\Gamma)$, $F = F(\Gamma)$ и $P = P(\Gamma)$ — элементы источника Г. Вершины графа G обычно называют *состояниями* источника Г.

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.

$$\mathfrak{L}(\Gamma) = \{\alpha \mid \exists\, p \in P \quad F(p) = \alpha\},$$

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

$$\varphi_\Gamma(\alpha_1\alpha_2, \ Q_1) = \varphi_\Gamma(\alpha_2, \ \varphi_\Gamma(\alpha_1, \ Q_1)). \qquad (172)$$

Регулярный источник Г, для которого $|\varphi_\Gamma(a, q)| \leq 1$ 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 $<A, Q, \varphi_\Gamma, q_0, — >$ 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 A -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 случаях 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 $q_0 = 0$, $A = E_2 = \{0, 1\}$, $\varphi_\Gamma(a, q) = a + q + 1 \pmod 4$ и $Q' = \{q \mid q \equiv 3 \pmod 4\}$ задают регулярный

источник, представленный также таблицей 1.5 и графически — на рис. 40.

Таблица 1.5

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

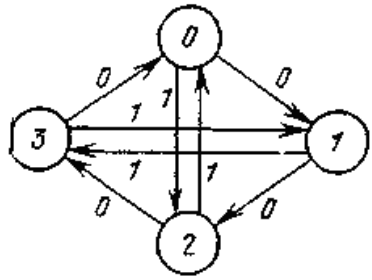


Рис 40.

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. 41. 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. 41.

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 $\Gamma = \langle G, F, P\rangle$ — произвольный простой регулярный источник. Полностью определенный *А*-источник $\Gamma_1 = \langle G_1, F_1, P_1\rangle$ зададим индуктивно:

$$Q_1 \subseteq \{q_M \mid M \subseteq Q\},\quad P_1 = P\left(q_{\{q_0\}}, \{q_M \mid M \cap Q' \neq \varnothing\}\right),\quad \varphi_{\Gamma_1}(a, q_M) = q_{\varphi_\Gamma(a, M)}.$$

Покажем, что для любого $\alpha \in A^*$ имеет место

$$\varphi_{\Gamma_1}\left(\alpha, q_{\{q_0\}}\right) = q_{\varphi_\Gamma(\alpha, q_0)}. \qquad (173)$$

Индукция по длине α. Если $|\alpha| \leq 1$, то (173) выполняется по определению. Пусть (173) справедливо для слов длины, меньшей *k*, и $|\alpha| = k > 1$: $\alpha = \beta a$, где $a \in A$, $|\beta| < k$. Тогда, используя (172), получаем

$$\varphi_{\Gamma_1}\left(\alpha, q_{\{q_0\}}\right) = \varphi_{\Gamma_1}\left(\alpha, \varphi_{\Gamma_1}\left(\beta, q_{\{q_0\}}\right)\right) = \varphi_{\Gamma_1}\left(a, q_{\varphi_\Gamma(\beta, q_0)}\right) =$$

$$= q_{\varphi_\Gamma(\alpha, q_{\varphi_\Gamma(\beta, q_0)})} = q_{\varphi_\Gamma(\beta a, q_0)} = q_{\varphi_\Gamma(\alpha, q_0)}$$

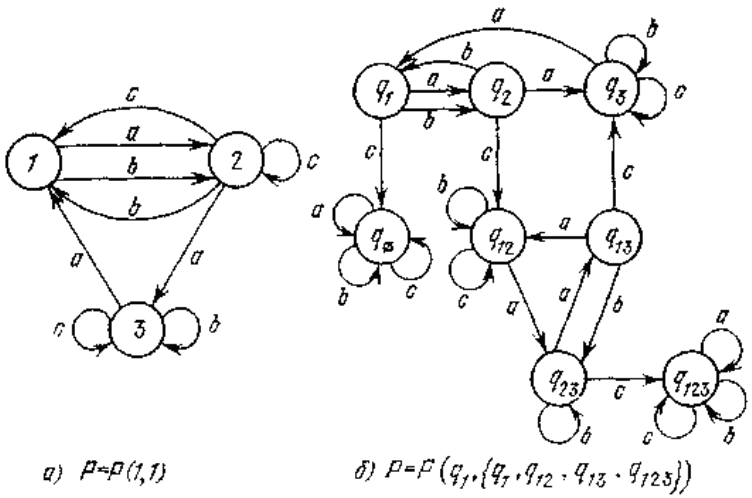
и (173) доказано. Но мы имеем

$$\alpha \in \mathfrak{L}(\Gamma) \leftrightarrow \varphi_\Gamma(\alpha, q_0) \cap Q' \neq \varnothing \leftrightarrow \varphi_{\Gamma_1}\left(\alpha, q_{\{q_0\}}\right) =$$

$$= q_{\varphi_\Gamma(\alpha, q_0)} \in Q_1' \leftrightarrow \alpha \in \mathfrak{L}(\Gamma_1)$$

и, следовательно, $\mathfrak{L}(\Gamma_1) = \mathfrak{L}(\Gamma)$. Теорема доказана.

Если число состояний Γ равно *N*, то число состояний эквивалентного ему *А*-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 А-источника.* При *N*= 3, *A = {a, b, c}* источник Лупанова и эквивалентный ему полностью определенный *А*-источник показаны на рис. 42, *а), б).*



а) $P=P(1,1)$ б) $P=P(q_1,\{q_1, q_{12}, q_{13}, q_{123}\})$

Рис. 42.

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 $A^*/\rho_{\mathfrak{L}} = \{\mathfrak{A}_0, \mathfrak{A}_1, \ldots \mathfrak{A}_t, \ldots\}$, где $\lambda \in \mathfrak{A}_0$. Заметим, что любой класс $\mathfrak{A}_t$ 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 $q_i = \mathfrak{B}_i$, $\lambda \in \mathfrak{B}_0$.

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

$$i(\alpha\beta) = j(i(\alpha), \beta). \tag{174}$$

Докажем, что

$$\rho_\Gamma \subseteq \rho_{\mathfrak{L}}. \tag{175}$$

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 (1.275) доказано.

Из (175) 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 A-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 Γ — полностью определенный A-источник, то $\Gamma_{\mathfrak{L}}$ получается из Γ факторизацией множества состояний по отношению эквивалентности, соответствующему $\rho_{\mathfrak{L}}$. Если число состояний Γ равното $N_{\mathfrak{L}}$, Γ совпадает с $\Gamma_{\mathfrak{L}}$ с точностью до обозначения состояний.*

Алгоритмические вопросы минимизации Л-источников решаются теми же методами, что и для конечных автоматов. (Пусть $Q(\Gamma_{\mathfrak{L}}) = Q(\Gamma)/\rho$ и $(q_i, q_j) \in \rho$, $\Rightarrow$ «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: $\rho_{N-1} = \rho_*$).

Если $|A| = 1$, то, как следует из теоремы 2, любой регулярный источник эквивалентен A-источнику, представляющему собой граф преобразования множества Q (рис. 43).

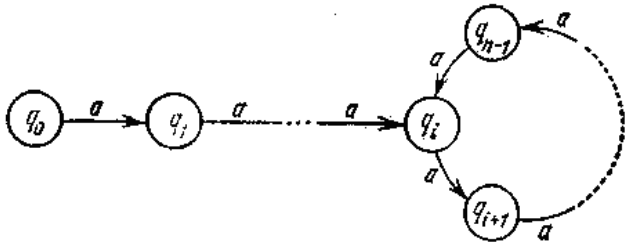


Рис. 43.

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, $\{a^{i^2} \mid i = 0, 1, 2, \ldots\}$ не регулярно — оно бесконечно и разность между соседними по возрастанию членами может быть как угодно велика.

Непосредственной проверкой можно убедиться, что

$$\rho_{A^* \setminus \mathfrak{L}} = \rho_{\mathfrak{L}}, \tag{176}$$

$$\rho_{\mathfrak{L}_1 \cap \mathfrak{L}_2} \supseteq \rho_{\mathfrak{L}_1} \cap \rho_{\mathfrak{L}_2}. \tag{177}$$

5. Здесь покажем, что некоторые естественные операции над регулярными языками приводят снова к регулярным языкам.

Из (176), (177) 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 $\mathfrak{L} \setminus \{\lambda\}$ регулярно; для любого регулярного $\mathfrak{L}$ множество $\mathfrak{L} \cup \lambda\}$ тоже регулярно.

Произведение языков $\mathfrak{L}_1$ и $\mathfrak{L}_2$ есть

$$\mathfrak{L}_1 \cdot \mathfrak{L}_2 \Rightarrow \{\alpha\beta \mid \alpha \in \mathfrak{L}_1, \beta \in \mathfrak{L}_2\}.$$

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 $\mathfrak{L}^+ \Rightarrow \mathfrak{L} \cup \mathfrak{L}^2 \cup \ldots = \bigcup_{i=1}^{\infty} \mathfrak{L}^i$.

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*

$$X = XS \cup T \tag{178}$$

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 (178), т. е. $X_0 = X_0S \cup T$. Подставляя это выражение для X_0 в правую часть (178) и применяя тождественные преобразований, получим

$$X_0 = (X_0S \cup T)S \cup T = X_0S^2 \cup TS \cup T. \qquad (179)$$

Следующую подстановку делаем в (179) 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. $X_0 \subseteq TS^*$. Таким образом, если (178) 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*

$$\begin{cases} X_0 = X_0S_{00} \cup X_1S_{10} \cup \cdots \cup X_{N-1}S_{N-1,0} \cup \{\lambda\}, \\ X_1 = X_0S_{01} \cup X_1S_{11} \cup \cdots \cup X_{N-1}S_{N-1,1}, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ X_{N-1} = X_0S_{0,N-1} \cup X_1S_{1,N-1} \cup \cdots \cup X_{N-1}S_{N-1,N-1} \end{cases} \qquad (180)$$

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* $\cup$, $\cap$, —, $\cdot$, *.

Это легко доказать, используя исключение неизвестных по теореме 1 и индукцию по N.

В силу теорем 1 — 3 в п. 5, формула $Y = \Phi(X_1, \ldots, X_n)$ алгебры регулярных множеств с операциями $\cup$, $\cap$, —, $\cdot$, * 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-источнником Г с множеством вершин $Q = \{q_0, q_1, \ldots, q_{N-1}\}$. Положим

$$X_i \Rightarrow \mathfrak{L}(\langle G, F, P(q_0, q_i)\rangle) \qquad i = \overline{0, N-1},$$

$$S_{ij} \Rightarrow \{a \mid a \in A, \exists r = \langle q_i, q_j\rangle \in R \qquad F(r) = a\},$$

$$i, j = \overline{0, N-1}.$$

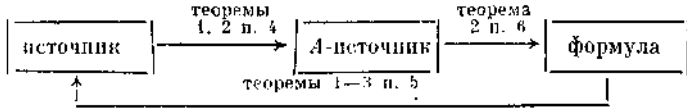
Тогда X_i и S_{ij} связаны системой уравнений (180), причем выполнены условия следствия 2 теоремы 1. Следовательно,

$$X_i = \Phi_i(S_{i0}, S_{i1}, \ldots, S_{N-1, N-1}), \qquad i = \overline{0, N-1},$$

и $\mathfrak{L} = \bigcup\limits_{q_i \in Q'} X_i$, ч.т.д.

Следствие (теорема Клини). *Класс регулярных множеств слов в алфавите А есть замыкание множества* $\{\{a\}, \ldots, \{a_n\}, \{\lambda\}\}$ *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 $A = \{a_1, \ldots, a_n\}$ 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* $m \to \infty$.

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)*-источников 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 $\left(\frac{l+1}{N+1} \right)^l \leqslant \left(\frac{N}{N+1} \right)^{N-1} \leqslant \left(1 - \frac{1}{N+1} \right)^{N+1} < \frac{1}{e}$, 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 в п. 1.18. 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 $\mathfrak{L}_1(\mathfrak{L}_2) = \mathfrak{L}_1 \setminus A^* \mathfrak{L}_2 A^*$, 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 $\mathfrak{L} = \{0^i\}_{i=0}^{\infty} \cup \{0^i 10^j\}_{i,j=0}^{\infty}$ — 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 $\mathfrak{L}_i \subseteq A^*$, $i \in J$:

$$\bigcap_{i \in J} A^*(\mathfrak{L}_i) = A^*\left(\bigcup_{i \in J} \mathfrak{L}_i\right).$$

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 Γ — сокращенный A-источиик,

порождающий $\mathfrak{L}$, $Q = \{q_0, q_1, \ldots, q_{N-1}\}$ — множество его состояний. Тогда для каждого $i = 0, 1, 2, \ldots, N-1$ существует слово β_i такое, что $\varphi_\Gamma(\beta_i, q_i) = \varnothing$.

Положим

$$\varphi_\Gamma(\beta_0, q_1) = q_{i(1)},$$
$$\varphi_\Gamma(\beta_0\beta_{i(1)}, q_2) = q_{i(2)},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\varphi_\Gamma(\beta_0\beta_{i(1)}\ldots\beta_{i(s-1)}, q_s) = q_{i(s)},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\varphi_\Gamma(\beta_0\beta_{i(1)}\ldots\beta_{i(N-2)}, q_{N-1}) = q_{i(N-1)},$$
$$\beta = \beta_0\beta_{i(1)}\ldots\beta_{i(N-1)}.$$

Возьмем произвольные слова $\alpha, \gamma \in A^*$. Пусть $\varphi_\Gamma(\alpha, q_0) = q_s$. Тогда имеем

$$\varphi_\Gamma(\alpha\beta\gamma, q_0) = \varphi_\Gamma(\beta\gamma, q_s) =$$
$$= \varphi_\Gamma(\beta_0\beta_{i(1)}\ldots\beta_{i(s-1)}\beta_{i(s)}\ldots\beta_{i(N-1)}\gamma, q_s) =$$
$$= \varphi_\Gamma(\beta_{i(s)}\ldots\beta_{i(N-1)}\gamma, q_{i(s)}) =$$
$$= \varphi_\Gamma(\beta_{i(s+1)}\ldots\beta_{i(N-1)}\gamma, \varphi_\Gamma(\beta_{i(s)}, q_{i(s)})) =$$
$$= \varphi_\Gamma(\beta_{i(s+1)}\ldots\beta_{i(N-1)}\gamma, \varnothing) = \varnothing.$$

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 $\alpha \to \alpha' \Rightarrow \binom{a_i}{a_j}\alpha$ 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 Γ — 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*

$$\alpha \neq \beta,\ \alpha' = \beta'\ \ u\ \ |\alpha| \leqslant |\beta| \leqslant 2N^3.$$

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 $\mathfrak{A}$ 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. 44).

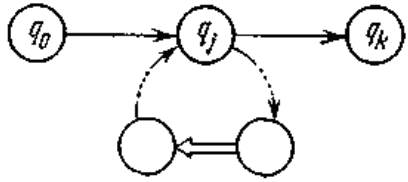


Рис. 44.

При возвращении к алфавиту A от пары слов, порожденной Γ_1, буквы b_i^{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. 45, соответствующие построения представлены на рис. 46 (и построении $\Gamma_1(a)$ нет необходимости, так как a — циклическая и, по лемме, существенная). Вывод: буквы a, b, d — существенные, c — фиктивная.

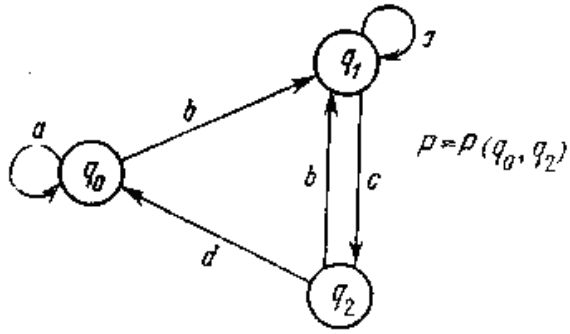

q_1
1
a
b
q_0
b
c
d
q_2
$p=p(q_0, q_2)$

Рис. 45.

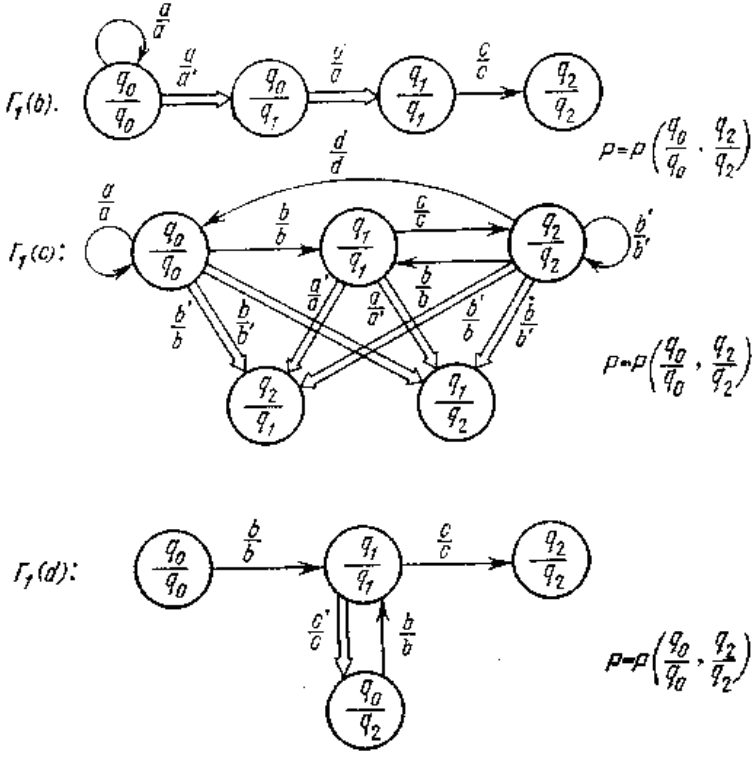

$\Gamma_1(b)$.
$p=p\left(\frac{q_0}{q_0}, \frac{q_2}{q_2}\right)$
$\Gamma_1(c)$:
$p=p\left(\frac{q_0}{q_0}, \frac{q_2}{q_2}\right)$
$\Gamma_1(d)$:
$p=p\left(\frac{q_0}{q_0}, \frac{q_2}{q_2}\right)$

Рис. 46.

Теорема 2. *Если a_i и a_j - контекстно неразличимы, то найдутся слова* $\alpha, \beta \in \mathfrak{A}$ *такие, что*

$$\alpha \neq \beta,\ \alpha' = \beta'\ и\ |\alpha| = |\beta| \leqslant 2N^2.$$

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 $\left\{\frac{a_k}{a_k} \mid k = \overline{1, m}\right\} \cup \left\{\frac{a_i}{a_j}, \frac{a_j}{a_i}\right\}$, 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. 45) любая пара букв, кроме *{a, b}*, контекстно различима. Для пары *{b, d}* подтверждающий это источник $\Gamma_1(b, d)$ изображен на рис. 47.

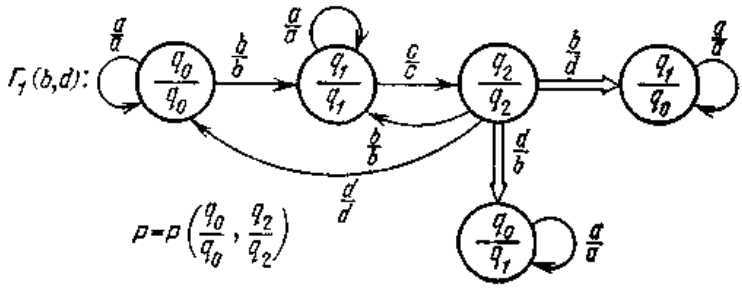


Рис. 47.

Другой источник изображен на рис. 48, *а)*.

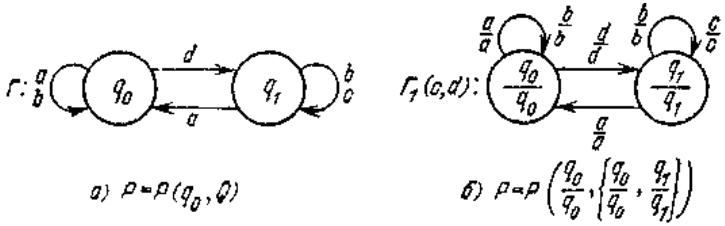


Рис. 48.

Легко проверить, что в порожденном им языке есть только две контекстно различимые буквы *c* и *d*, источник $\Gamma_1(c, d)$ показан на рис. 48, 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. 49.

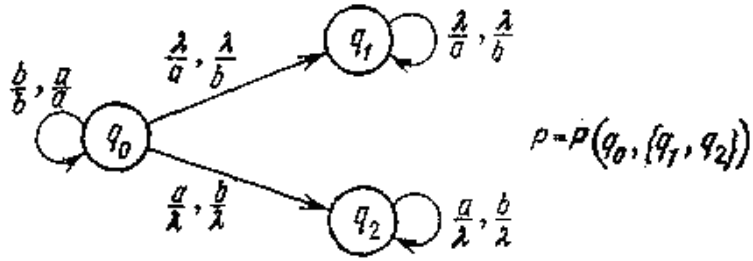


Рис. 49.

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 преобразования каждого ребра Г по правилу, показанному на рис. 50.

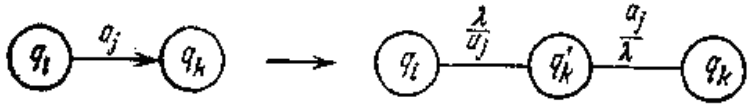


Рис. 50.

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

$$Q(\Gamma) = \{q_0, q_1, \ldots, q_{N_\Gamma - 1}\} \text{ и } Q(\Gamma_1) = \{r_0, r_1, \ldots, r_{N_R - 1}\}.$$

Возьмем $\left\langle q_0, \frac{r_0}{r_0} \right\rangle$ в качестве начального состояния Γ_2 и определим функцию следования посредством

$$\varphi_{\Gamma_2}\left(\frac{\lambda}{a}, \left\langle q_i, \frac{r_j}{r_k}\right\rangle\right) = \left\{\left\langle q_p, \frac{r_j}{\varphi_{\Gamma_1}(a, r_k)}\right\rangle \middle| q_p \in \right.$$

$$\left. \in \varphi_\Gamma\left(\frac{\lambda}{a}, q_i\right)\right\},$$

$$\varphi_{\Gamma_2}\left(\frac{a}{\lambda}, \left\langle q_i, \frac{r_j}{r_\kappa}\right\rangle\right) = \left\{\left\langle q_p, \frac{\varphi_{\Gamma_1}(a, r_j)}{r_k}\right\rangle \middle| q_p \in \right.$$

$$\left. \in \varphi_\Gamma\left(\frac{a}{\lambda}, q_i\right)\right\}.$$

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. 51. 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) $\{\langle 4, 5\rangle\}_R$.

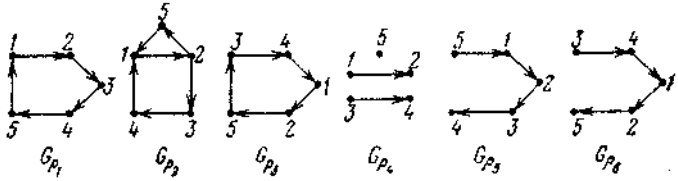


Рис. 51.

2. R — класс всех отношений эквивалентности на множестве $\{2, 3, 6, 8, 9, 12\}$, $\rho = \{\langle i, j\rangle \mid НОД\ (i, j) = 1\}$. 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 $\langle \alpha, \beta b\rangle \in \rho$, $\langle \alpha, \beta\rangle \in \tau \Rightarrow$ «α есть префикс $\{\beta\}$». Доказать, что $\tau = (=)_R$. Относительно какого правила отношение префиксности определяется отношением равенства синтаксически?

4. $X, Y \ldots$ — произвольные множества. Доказать о помощью диаграмм Венна, что:

а) $X = Y$ равносильно $X \otimes Y = \varnothing$;

б) $X \subseteq Y$ равносильно $X \mid Y = \varnothing$;

в) из условий $X \cap \overline{Z} = \varnothing$, $Y \cap \overline{W} = \varnothing$, $Z \cap W = \varnothing$, $X \cup Y = Z \cup W$ 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 $Y \subseteq X\overline{Z} \cup \overline{X}Z$, $YW \subseteq XZ \cup \overline{X}\overline{Z}$, $XY \subseteq Z \cup W$, $\overline{Y}\overline{Z} \subseteq X \cup W$ равносильна условию $Y = \varnothing$.

5. $A, B, C \subseteq U$, $|U| = n$. Доказать, что в универсе U

а) уравнение $AX \cup BY = \overline{B}X \cup AY$ имеет $2^{2n-|A \otimes B|}$ решений;

б) уравнение $X \cup Y = \overline{Y} \cup Z$ имеет 4^n решений;

в) система уравнений $\begin{cases} AX \cup B\overline{X} = C, \\ BX \cup A\overline{X} = \overline{C} \end{cases}$ имеет решение в том и только том случае, если $A = \overline{B}$ (если решение существует, то оно единственно и есть $X = A \otimes \overline{C}$).

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. 52 помечены символом $\varnothing$. Привести примеры функций каждого из остальных типов.

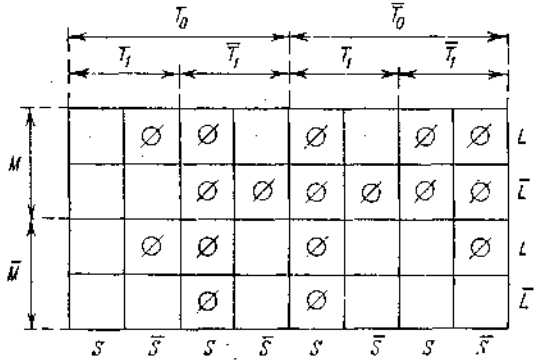


Рис. 52.

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, ..., 0^i1, ...},

б) {00, 10, 010, 0110, ..., 01^i0, ...}.

в) {0, 100, 10100, ..., v_i, v_{i+1}, $v_{i+2} = 1\, v_i v_{i+1}$, ...}.

12. Какие соотношения и какие тождества выполняются в полугруппе, порожденной преобразованием f с операцией суперпозиции преобразований?

а) $f = \begin{pmatrix} 1 \to 2 \\ 2 \to 3 \\ 3 \to 3 \\ 4 \to 1 \\ 5 \to 1 \end{pmatrix}$, 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 квазитождество Мальцева.

15. Найти все разложения слова α на множители из *V*, где:

а) $V = \{1, 10, 01\}$, $\alpha = 10100101$, $\alpha = (100)^21$;

б) $V = \{11, 110, 01, 101\}$, $\alpha = 110101$, $\alpha = 1^301^201$;

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 единицы.

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) $V = \{0^6, 0^9, 0^{17}\}$, $\Pi\langle\{x, y, z\} \mid xy = yx, xz = zx, yz = zy, x^8 = y^2, z^3 = x^4y^3\rangle = \Pi\langle\{x, y, z\} \mid xy = yx, xz = zx, yz = zy, x^3 = y^2, z^3 = x^7y\rangle$;

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 $x \cdot y$ понимается бесконечная последовательность *xy*. Пусть $\alpha = \beta\beta \ldots \beta \ldots = \beta^\infty$ — бесконечная периодическая последовательность. Доказать, что:

а) $\beta'\alpha = \alpha$ 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

$$x = (\alpha\beta)^l\alpha,\quad y = (\beta\alpha)^l\beta,\quad z = (\alpha\beta)^k,$$

где α, β — произвольные слова такие, что $\mathfrak{e}(\alpha) \neq \mathfrak{e}(\beta)$, $l, j, k \geqslant 0$. Показать, что среди его решений имеется бесконечное число попарно неизоморфных.

23. Найти общее решение следующих уравнений и систем уравнений в словах:

$$\text{а)}\ xyxz = zx^2y; \qquad \text{б)}\ z^3 = x^2yx^3; \qquad \text{в)}\ \begin{cases} xyz = yzx, \\ yx = ux. \end{cases}$$

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 $x^iy = yz^j$ при любом j.

28. Доказать, что:

а) если множество $\mathfrak{L}$ регулярно, то и его обращение $O(\mathfrak{L}) = \{x_1x_2 \dots x_i | x_i \dots x_2x_1 \in \mathfrak{L},\ x_j \in A\ (j \geqslant 1)\}$ 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 $\alpha' = x_1 \dots x_k$ и для множества слов $\mathfrak{L}$ пусть $\mathfrak{L}' = \{\alpha' | \alpha \in \mathfrak{L}\}$.

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 $\Phi(\mathfrak{L})$ 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 эквивалентности в случаях:

а) $\mathfrak{L} = \{ab, ca, bc, bb, ba\}$, $A = \{a, b, c\}$;
б) $\mathfrak{L} = \{ab, aba, bb\}$, $A = \{a, b\}$;
в) $\mathfrak{L} = \{a^3, a^4, \ldots\} = \{a^i \mid i \geqslant 3)\}$, $A = \{a, b\}$.

30. Пусть $\Gamma = \langle G, F, P\rangle$ — 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. 53.

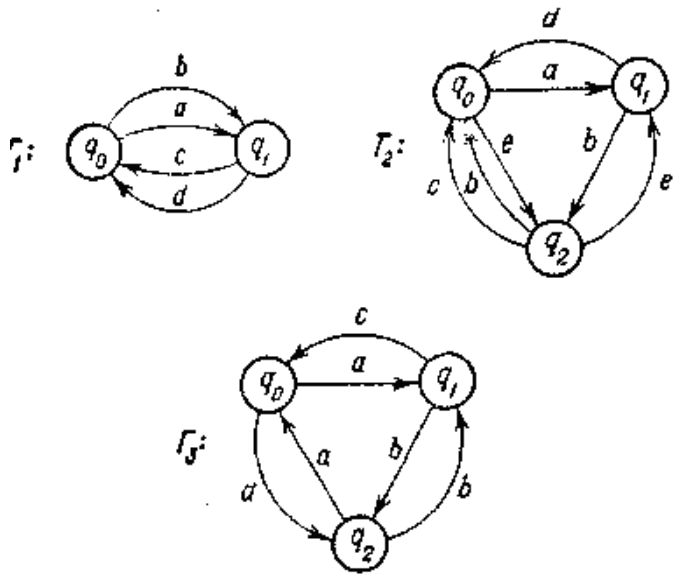


Рис. 53.

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).

4. Диаграммы Венна

Введение

Вероятно, нет такой науки, в которой логика применялась бы в большей мере, чем в математике. Все рассуждения в математике носят строго логический характер. Все предложения, если они не приняты за аксиомы, строго доказываются. Самый выбор аксиом также логически обосновывается: ищутся доказательства непротиворечивости, полноты, часто и независимости аксиом. В начале XX в. были сделаны даже попытки доказать, что математика вообще есть часть логики — что она сводится к логике (логицизм Г. Фреге и Б. Рассела). Однако знаменитая теорема Гёделя **о неполноте любых формальных систем** типа Principia Mathematica Б. Рассела и А. Н. Уайтхеда **доказала существование в языке каждой такой формальной системы предложения — в том числе даже содержательно истинного,— которое, тем не менее, не доказуемо и не опровержимо в этой системе.** Таким образом, нет оснований считать, что в соотношении между логикой и математикой первенствующая роль принадлежит логике.

В действительности эта ситуация уже была подготовлена предшествующим развитием математической логики. Математическая логика возникла, наоборот, из попытки перенести методы математики в логику. Такая попытка относилась прежде всего к методам алгебры. Нетрудно заметить, что **всякое решение алгебраических уравнений есть не что иное, как вывод логических следствий из посылок, записанных на языке алгебры в виде уравнений**. Естественно возникал вопрос о том, нельзя ли любые посылки выразить уравнениями и свести всякий логический вывод к решению уравнений. В действительности вопрос был поставлен так еще в XVII в. Лейбницем. Тот же Лейбниц размышлял и о других способах сведения логики к математике: к арифметике и геометрии. Но большинство из этих работ Лейбница стало известно науке только после 1903 г.

Первые попытки успешного построения алгебры логики были осуществлены в середине прошлого века прежде всего рядом английских логиков: Дж. Булем, А. де Морганом, У. С. Джевонсом. У Буля же были и попытки построения логики как некоторой арифметики с ее обычными четырьмя операциями. Но в логике уже издавна применялись фактические и геометрические методы. Прежде всего, в целях наглядности. Эти геометрические (точнее, графические) методы должны были облегчить усвоение силлогистики Аристотеля,

которая стала сложной и специальной наукой в трудах схоластов. Особенно широкой известностью уже в конце XVIII в. стала пользоваться графическая интерпретация силлогистики Аристотеля с помощью кругов Эйлера. Эйлер изложил эту интерпретацию в своих «Письмах к немецкой принцессе» (1768) с целью показать, что никаких трудностей в силлогистике Аристотеля на самом деле нет и что основные ее принципы не сложнее, чем утверждение, что если деньги у меня в кошельке, а кошелек в кармане, то деньги — в кармане.

Графические методы логики развивались и далее. Более детальную графическую формализацию элементарных предложений логики предложил Ж. Д. Жергонн (1771—1859). Другие способы их геометрической иллюстрации наметили И. Г. Ламберт (1728—1777) и Б. Больцано (1781—1848). Но особенного расцвета эти методы достигли в трудах английского логика Джона Венна, **построившего графический аппарат диаграмм**, фактически полностью эквивалентный развитой уже (в трудах Дж. Буля, А. де Моргана, У. С. Джевонса, Э. Шредера, П. С. Порецкого, Ч. Пирса и др.) логике классов.

К концу XIX в. в логике произошел подлинный переворот. Этот переворот был связан с тем, что в трудах немецкого ученого Готлоба Фреге **в логику впервые были в явном виде введены кванторы.** Основная трудность силлогистики Аристотеля состояла в том, что в ней нужно было уметь не только пользоваться правильными **модусами силлогизма, но и опровергать неправильные**. А это предполагало два разных вида отрицания (нужно было уметь сказать, например, «Неверно, что все *А* суть *не-В»).* Графическая интерпретация с помощью кругов Эйлера легко обнаруживала неправильные модусы без обращения к более сложному аппарату кванторов. Но она была пригодна только для рассуждений, не выходящих за пределы силлогистики Аристотеля, т. е. не была даже полностью эквивалентна исчислению одноместных предикатов. Не удивительно поэтому, что кванторная логика Фреге, переработанная затем (в «Principia Mathematica») Б. Расселом и А. Н. Уайтхедом, ознаменовала переход на новую ступень в истории логики. Логика классов, а вместе с ней и диаграммы Венна отошли на задний план и частично были забыты. Нельзя, впрочем, сказать, что совершенно: время от времени диаграммы Венна упоминаются в учебниках, а иногда и описываются, например, у Беннета и Бейлиса, К. Дюрра, Т. Котарбинского [18] и других математических логиков. Их использует в работах по бионике школа американского нейрофизиолога У. С. Мак-Каллока. И все же мы вряд ли ошибемся, если скажем, что большинство логиков знают о диаграммах только понаслышке. Об этом говорит, например,

обыкновение путать диаграммы Венна с кругами Эйлера (о различии между графическими методами Венна и Эйлера будет рассказано в далее).

Надо сказать, что и некоторые представители логики классов не слишком дружелюбно относились к диаграммам Венна. Л. Кутюра, например, говорил по поводу диаграмм Венна, что «этот схематизм, рассматриваемый как метод решения логических задач, имеет серьезные недостатки и является весьма мало полезным». Однако такое отношение к диаграммам представляется несправедливым и незаслуженным. Дело в том, что связанные с методом диаграмм Венна алгоритмы по существу не были.разработаны. Начиная с произведений Венна, в литературе описание графического аппарата сводится к демонстрации его на примерах. Цель настоящего раздела состоит не только в том, чтобы показать, что способы, относящиеся к диаграммам, могут быть доработаны до соответствующих алгоритмов, но и дать чисто геометрическое построение логики высказываний и логики одноместных предикатов в виде определенных операций над диаграммами Венна.

Не говоря уже о том, что задачи и методы классической алгебры логики конца XIX — начала XX в. и сейчас могут представлять интерес (тем более, что соответствующие работы остаются, в сущности, еще не проанализированными), нет необходимости ограничивать применимость диаграмм рамками логики классов. Правда, Венн применяет свои диаграммы исключительно для иллюстрации решения задач логики классов, и они работают при этом, как мы покажем, весьма успешно. Но диаграммы можно столь же эффективно использовать и в задачах, интересующих сегодняшних логиков — например, для получения логических следствий из заданной системы посылок.

Разумеется, предлагаемое геометрическое построение логики высказываний и логики одноместных предикатов эквивалентно обычному. Ценность и преимущество диаграмм состоит в их наглядности. Наглядность может оказаться существенным подспорьем при решении задач и доказательстве теорем, во всяком случае пренебрегать ею вряд ли стоит. Кроме того, аппарат диаграмм допускает обобщения, важные, например, при описании функционирования нейронных сетей и построении надежных систем из относительно мало надежных элементов.

Исходя из этих соображений, мы попытаемся дать строгое изложение метода диаграмм Венна и его приложений. Причем начнем с описания методов самого Венна, которые подробно рассматриваются в первой главе. Эти методы представляют интерес и сейчас и могут быть ис-

пользованы в курсах математической логики и в задачниках по этой дисциплине.

Во второй и третьей главах диаграммы Венна применяются при решении задач вывода логических следствий из посылок, выразимых на языке формул классического исчисления высказываний и классического исчисления одноместных предикатов. Предлагаются способы, позволяющие на диаграммах принципиально обозревать все возможные попарно неэквивалентные логические следствия из данных посылок, а также отличать такие логические следствия, которые можно считать в определенном смысле нетривиальными — так называемые простые логические следствия. Теоретически диаграммы Венна можно использовать, как будет показано, при любом числе переменных. Однако с ростом количества переменных наглядность диаграмм падает, и возникает необходимость обойти эту трудность — скажем, видоизменить диаграммы. В работе одним из способов преодоления указанной трудности является использование матричного аппарата, который, как и аппарат диаграмм, позволяет выходить за рамки логики классов и исчисления высказываний.

Во второй главе (в связи с такими, например, проблемами, как синтез надежных сетей из не вполне надежных элементов, построение управляющих и самоуправляющих систем и блочный анализ и синтез сложных устройств) аппарат диаграмм Венна в классическом исчислении высказываний расширяется в аппарат вероятностных диаграмм; вводятся сети диаграмм. В сетях понятие «диаграмма Венна» получает дальнейшее расширение — **диаграмма играет роль оператора.**

В четвертой главе освещаются вопросы, относящиеся к применению аппарата диаграмм Венна для описания функционирования формальных нейронов Мак-Каллока и сетей из них. В теории формальных нейронов, кроме диаграмм Венна в классическом исчислении высказываний и вероятностных диаграмм, используются пороговые и порядковые диаграммы n переменных. В этой главе решаются задачи синтеза формальных нейронов по заданным условиям их работы, описываются общие способы построения оптимальных нейронов. Рассматриваются формальные нейроны с обратными связями, показывается, что введение обратных связей позволяет увеличить количество функций, которые можно реализовать на одном формальном нейроне.

Прежде чем переходить к основному тексту, скажем несколько слов о создателе диаграммного метода.

Джон Венн (John Venn) родился 4 августа 1834 года в Драйпуле близ г. Халла в семье священника. Образование получил в одном из

колледжей Кембриджа (Гонвиль и Кийус колледж). В 1858 г. он начал (по семейной традиции) карьеру служителя церкви. Однако его интересы лежали в другой области, он стремился к научным занятиям в области философии и логики, В 1862 г. Венн вернулся в Кембридж, где был приглашен на только что созданное место лектора колледжа по моральным наукам. Несколько месяцев он совмещал еще свою работу с деятельностью священника в небольшой церкви, но вскоре отказался совсем от церковной деятельности; в 1883 г. он даже получил документ, утверждающий, что он не способен служить в церкви.

С 1862 г. он посвятил себя научной и педагогической деятельности. В 1883 г. в Кембридже он получил степень доктора наук и был избран членом Королевского общества. В 1903 г. Венн избирается президентом Гонвиль и Кийус колледжа. Занимаясь философией и логикой, Венн опубликовал много статей в различных журналах и несколько монографий, из которых наиболее значительными являются «Логика случая» («The Logic of Chance», 1866), «Символическая логика» («Symbolic Logic», 1881), «Принципы эмпирической логики» («The Principles of Empirical Logic», 1889). Много внимания Венн уделял также изучению истории и составлению биографических и исторических очерков, справочников и др.

В логике Венн не интересуется проблемами психологии; не занимается он и такими вопросами, которые в его время относились к области «метафизики». Свои философские позиции сам Венн характеризовал как «точку зрения опыта и здравого смысла»). Умер Венн 4 апреля 1923 г.

[Эти биографические сведения заимствованы из «Dictionary of National Biography», Oxford University Press, London, 1937, стр. 869—870; автором биографии Венна является его единственный сын, Джон Арчибальд Венн (избранный в 1932 г. президентом Королевского колледжа в Кембридже), который использовал неопубликованную автобиографию отца.]

4.1. Диаграммы Венна в логике классов

4.1.1. Круги Эйлера

Когда заходит речь о графических методах логики (не обязательно математической), обыкновенно вспоминаются так называемые круги Эйлера. Чтобы убедиться в этом, достаточно взять любой учебник традиционной (аристотелевской) логики и посмотреть разделы, посвященные объему понятия или категорическому силлогизму. Там

почти обязательно мы увидим картинки с изображениями кругов или прямоугольников. Объемы понятий или, выражаясь языком математической логики, классы изучаемых объектов принято обозначать фигурами, ограниченными простыми замкнутыми кривыми, обычно окружностями, расположенными на части плоскости так, что одна фигура включает в себя все объекты одного класса (скажем *A),* а другая — все объекты другого класса *(B).* Дополнение каждого класса обозначается внешней областью соответствующей замкнутой кривой. При этом некоторая часть плоскости представляет всю рассматриваемую предметную область.

Две такие фигуры на плоскости (для краткости вместо слов «часть плоскости» часто будем писать слово «плоскость») можно расположить, как легко проверить, **пятью различными способами**. На рис. 1 приведены все различные способы расположения двух кругов *A* и *B;* в первом случае плоскость делится на две части (ячейки): внутреннюю и внешнюю, во втором, третьем и пятом — на три, в четвертом — на четыре.

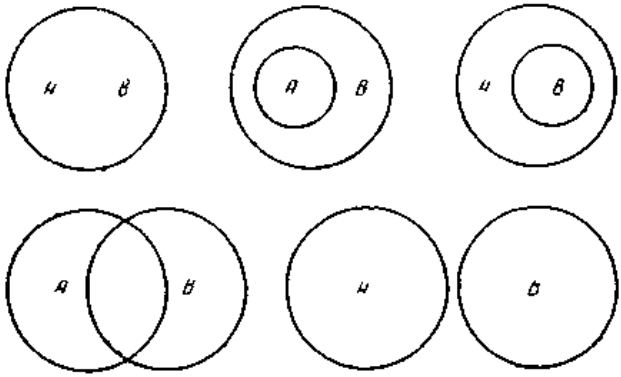


Рис. 1

Пересечение *A* с *B* в каждом случае включает в себя все объекты, принадлежащие как классу *A,* так и классу *B;* области на плоскости, лежащей вне *A* и *B,* соответствуют объекты, не принадлежащие ни *A,* ни *B;* наконец, областью, лежащей только внутри одной из фигур *A, B,* обозначаются объекты, относящиеся одновременно к соответствующему классу и к дополнению другого класса.

Такое представление классов как систем нескольких фигур на плоскости встречается уже в XIII веке. Например, в устройстве, предложенном Раймондом Луллием для машинизации процессов рассуждения, участвуют комбинации различных кругов на плоскости.

Ясно, что с ростом числа рассматриваемых классов, количество различных способов взаимного расположения соответствующих фигур на плоскости резко возрастает. В литературе показаны некоторые из трудностей, связанных с указанием всех способов расположения n (где $n > 2$) кругов на плоскости.

Наглядность геометрического представления классов в виде фигур на плоскости была использована Л. Эйлером при объяснении и решении задач силлогистики Аристотеля. О своих фигурах Эйлер писал, что они «очень подходят для того, чтобы облегчить наши размышления... и открыть нам все тайны, которыми похваляются в логике и которые доказывают в ней с большим трудом, между тем, как посредством этих знаков всякая ошибка сразу бросается в глаза». Здесь под словами: «тайны, которыми похваляются в логике» Л. Эйлер имеет в виду сложности схоластической разработки силлогистики Аристотеля. При этом Эйлер подчеркивает, что при предлагаемом им способе употребляют фигуры или пространства, «имеющие какую угодно форму, чтобы представить каждое общее понятие, и обозначают субъект предложения пространством, содержащим *A,* а предикат другим пространством, содержащим *B.* Природа самого предложения включает всегда либо то, что пространство *A* находится полностью в пространстве *B,* либо то, что оно находится в нем лишь частично, либо, что, по меньшей мере, какая-нибудь часть его находится вне пространства *B,* либо, наконец, что пространство *A* лежит полностью вне *B*».

Классической и вошедшей во все учебники логики иллюстрацией метода Эйлера является изображение (рис. 2) модуса barbara:

Всякое *M* есть *P,*
Всякое *S* есть *M,*
Следовательно, всякое *S* есть *P*.

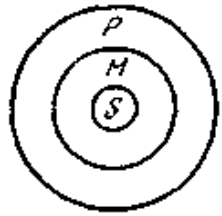


Рис.2

Эйлер предлагал при изучении модусов силлогизма учитывать все возможные случаи взаимного расположения кругов, соответствующих различным терминам в посылках данного модуса. Так, силлогизм

Всякое *A* есть *B,*
Но некоторые *C* суть *A,*
Следовательно, некоторые *C* суть *B.*

он иллюстрирует чертежами на рис. 3: в одном из них круг *C* расположен полностью внутри *B,* в другом — частично.

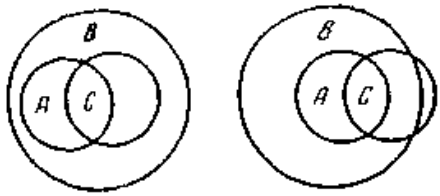


Рис. 3

Однако, следует заметить, что на диаграммах Эйлера, изображенных на рис. 3, не учитывается ряд других случаев (см. рис. 4; напомним, что термин «некоторые» понимается в аристотелевой логике в смысле «некоторые, а может быть и все»).

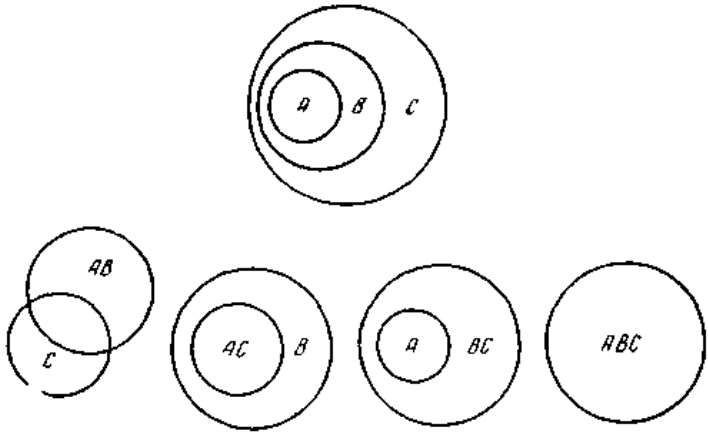


Рис. 4

Сам Эйлер указывал при этом, что его круги не совсем точно выражают информацию, содержащуюся в предложениях силлогистики. Так, по поводу фигуры, изображенной на рис. 5, где класс *A* соответствует деревьям, а класс *B* — грушам, он пишет, что ее можно выразить словами по-разному:

«1. Все груши — деревья.
2. Некоторые деревья — груши.
3. Некоторые деревья — не груши».

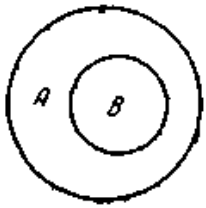


Рис. 5

Аналогично, по поводу фигуры, изображенной на рис. 6, он пишет, что здесь можно сказать:

«1. Некоторые *A* суть *B*.
2. Некоторые *B* суть *A*.
3. Некоторые *A* не суть *B*.
4. Некоторые *B* не суть *A*»

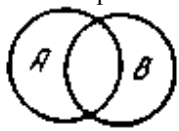


Рис. 6

Круги Эйлер использует также и для опровержения силлогизмов. Например, он пишет: «Если понятие *C* лежит полностью вне понятия *A,* то ничего нельзя сказать относительно понятия *B*. Может случиться, что понятие *C* лежит полностью вне *B* (как показано на рис. 7а.*)* или полностью внутри *B* (рис. 8.), или частично в *B* (рис. 9.*)*, так что ничего нельзя заключить».

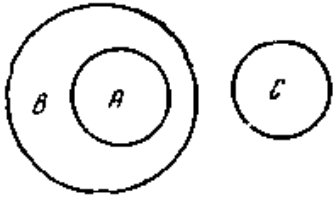


Рис. 7а

Рис. 7б

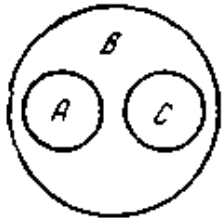


Рис. 8

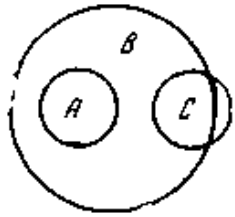


Рис. 9

Чтобы выразить, что некоторая часть пространства *B* находится в пространстве *C,* Эйлер отмечает эту часть звездочкой (рис. 10).

Рис. 10

Приведем пример неправильного силлогизма:

Некоторые *A* суть *B.* Ни одно *B* не суть *C.* Следовательно, некоторые *C* не суть *A.*

Этот силлогизм иллюстрируется диаграммами, приведенными на рис. 11—13.

Рис. 10

Рис. 11

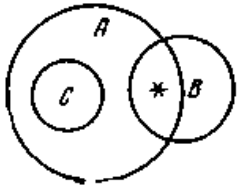


Рис. 12

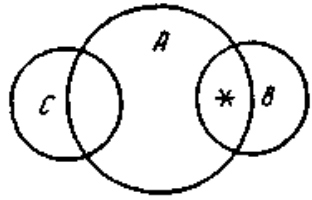


Рис. 13

Неправильность силлогизма следует из фигуры, построенной на рис. 12, где все *C* находятся внутри пространства *A.*Из диаграмм Эйлера для предложений силлогистики, вообще говоря, не всегда ясно, какую именно информацию они выражают. Так, предложения «Всякое *A* есть *B*» и «Ни одно *B* не есть *C*» выражаются обычно в виде диаграммы Эйлера, изображенной на рис. 7а. Если второе предложение точно

передается на диаграмме, то иначе обстоит дело с первым. Часть диаграммы с кругами *A* и *B* здесь понимается, собственно говоря, не в том смысле, какой следует из самого рисунка (т. е. как «Всякое *A* есть *B,* но некоторые *B* не суть *A»),* а в смысле «Всякое *A* есть *B,* но некоторые *B* не суть *A,* или всякое *A* есть *B,* и нет таких *B,* которые не были бы не *A»* (т. к. именно этот смысл вкладывается в предложение «Все *A* суть *B»),* Но этот смысл, конечно, никак нельзя вычитать из самой диаграммы, изображенной на рис. 7а: он содержится лишь в двух диаграммах— 7а и 76. Кроме того, добавление новой информации предполагает необходимость не просто детализировать уже имеющуюся диаграмму, но строить новую.

Последовательно проведенный метод Эйлера включает перебор всех возможных взаимных расположений классов. При рассмотрении модусов силлогистики Аристотеля этот перебор практически можно осуществить; в более общих случаях перебор становится необозримым. Круги Эйлера получили широкое распространение в литературе по математической логике. В качестве примера остановимся на «Алгебре логики» Э. Шредера. Прежде всего Шредер отмечает, что в действительности Эйлер не был первым, кто пользовался геометрическими иллюстрациями для выражения соотношений между субъектом и предикатом суждения,— уже у Людвига Вивеса (1555) употреблялись для этой цели уголки или треугольники. Шредер считает тем не менее, что фактически именно Эйлер должен считаться основателем метода диаграмм в логике. При этом он подчеркивает два обстоятельства. Во-первых, **эйлеровы круги можно рассматривать совершенно независимо от логики и построить соответствующее исчисление областей**. Во-вторых, в логике диаграммы Эйлера могут иметь существенное значение только в самом начале, когда речь идет об установлении некоторых принципов или аксиом логики, которые принимаются без доказательства и обосновываются только интуитивно. В качестве такого принципа Шредер использует, например, принцип силлогизма:

$$\frac{a \subset b,\ b \subset c}{a \subset c}$$

Читается:

«Если *a* включено в *b* и *b* включено в *c*,
то *a* включено в *c*».

Этот принцип Шредер иллюстрирует с помощью кругов Эйлера и поясняет тут же, что, если в начале курса логики геометрические иллюстрации оказываются более понятными, чем формальные логические выводы, то в дальнейшем соотношение между формальными преобразованиями и геометрическими иллюстрациями

меняется: **формальные преобразования становятся более убедительными, чем геометрические иллюстрации**. Диаграммы Эйлера употребляются обычно для иллюстрации «типичных» случаев; «вырожденные» случаи (например, когда все *a* суть *b* и в то же время все *b* суть *a)* при этом обычно опускаются.

Утверждение вроде: если для данных *a* и *b* существует *c* такое, что для всякого *x,* для которого

$$x \subset c \qquad | \qquad c \subset x$$

(здесь и далее вертикальные линии служат для параллельной записи двойственных теорем, относящихся к операциям пересечения (обозначается точкой, которая может и пропускаться) и объединения (обозначается знаком «плюс») классов)

имеет место также

$$x \subset a, \quad x \subset b, \quad | \quad a \subset x, \ b \subset x,$$

то мы имеем право сказать, что это эквивалентно утверждению

$$c \subset ab, \quad | \quad a + b \subset c,$$

Шредер иллюстрирует с помощью геометрических картинок (рис. 14, 15).

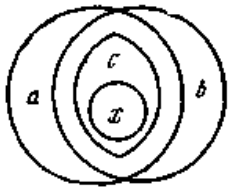


Рис. 14

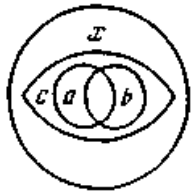


Рис. 15

Аналогичные картинки (рис. 16, 17) Шредер дает и для следующих двух двойственных теорем или определений:

Если для данных *a* и *b* существует такое *c*, что для всякого *x,* удовлетворяющего одновременно условиям

$$x \subset a, \quad x \subset b \quad | \quad a \subset x, \ b \subset x,$$

всегда верно также

$$x \subset c, \quad | \quad c \subset x,$$

то мы вправе сказать, что это эквивалентно тому, что

$$ab \subset c, \quad | \quad c \subset a + b.$$

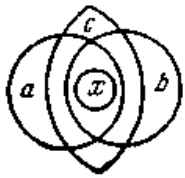


Рис. 16

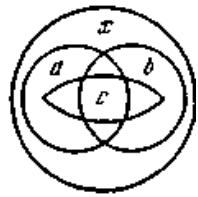


Рис. 17

Пересечение и объединение двух классов Шредер иллюстрирует с помощью кругов Эйлера, в которых заштриховывает общую часть (пересечение) или объединение. Ассоциативные законы для пересечения и объединения областей Шредер также иллюстрирует с помощью штриховки геометрических фигур или их частей.

Он приводит доказательство того, что из $ac \subset bc$ (соответственно $a + c \subset b + c$) еще не следует, что $a \subset b$, проводимое с помощью геометрической картинки (см. рис. 18, 19).

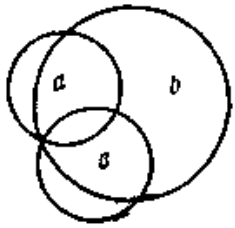


Рис. 18

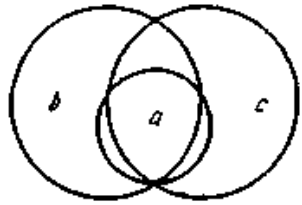


Рис. 19

Аналогичное графическое обоснование дается и для утверждения, что из *ac* = *be* (соответственно *a* + *c* = *b*+ c) еще не следует *a* = *b*, которое осуществляется с помощью рис. 20, 21; на обеих картинках заштрихованная область — это *c*.

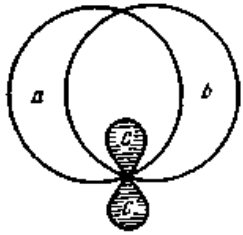


Рис. 20

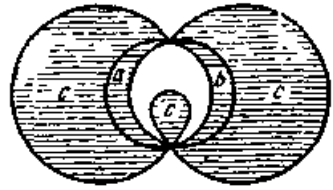


Рис. 21

Описанный выше метод опровержения тех или иных предложений с помощью графических примеров применяется в логике. В этой связи достаточно напомнить приводимое Я. Лукасевичем доказательство

неполноты некоторой аскиоматики, верной для любого числа кругов, принимаемых в качестве универсума, с помощью приведения примера предложения, верного для универсума из трех кругов и неверного для универсума из четырех. Аналогичный метод использован Ферисом в статье о силлогистике Ж. Д. Жергонна, представляющей интерес своей темой, поскольку речь идет о построении исчисления, элементарные предложения которого однозначно соответствуют диаграммам, изображенным на рис. 1. Здесь также неполнота некоторой аксиоматики доказывается с помощью графической интерпретации, для которой предложение, выразимое на языке Жергонна, оказывается ложным в области из четырех или большего числа кругов, но истинным в области, состоящей из трех кругов.

В заключение остановимся на истории эйлеровых кругов и графических методов вообще. Выше уже упоминалось замечание Э. Шредера о Л. Вивесе. Некоторые сведения об истории применения диаграмм в логике приводит Г. Шольц. Значительно обширнее специально посвященный истории графических методов раздел монографии Дж. Венна «Символическая логика». Венн начинает свой очерк с того, что опровергает некоторые распространенные в его время представления, относящиеся к этой истории. Так, он отказывается признавать графическими методами логики использование таких картинок, как, например, известное «древо Порфирия», поскольку с их помощью не решается какая-либо специфически логическая задача. Прежде всего, они не содержат анализа рассматриваемых предложений,— не выделяют в них субъект и предикат, не фиксируют их форму (e, a, i, o) ((e, a, i, o)— принятые в традиционной (аристотелевой) логике обозначения предложений силлогистики. Обозначают, соответственно, общеотрицательное («Ни одно *А* не есть *В»),* общеутвердительное («Все А суть В»), частноутвердительное («Некоторые *А* суть *В»)* и частноотрипательное («Некоторые *А* не суть В») предложения.)

Особенное неудовольствие Венна вызывают попытки У. Гамильтона приписать приоритет в изобретении геометрических методов силлогистики не Л. Эйлеру, а Алстеду или Вайзе. У Ланге, о котором И. Ламберт в 1771 году писал, что его труд [78] содержит большое число иллюстраций силлогизмов с помощью квадратов и кругов, Венн нашел, правда, в другой работе, только графические иллюстрации «древа Порфирия». Г. Шольц, со своей стороны, отмечает, что еще в 1584 г. Юлий Паций (Pacius) в своем комментированном издании аристотелева «Органона» пользовался фигурами (правда, это были не круги) для символизации отношений понятий с такой уверенностью, какая заставляет предполагать наличие у него предшественников.

Круги Шольцу встретились внервые у немецкого математика Иоганна Кристофа Штурма (1635—1703) в книге «Universalia Euclidea» (1661, Наад). Что же касается Ланге, то Шольц подтверждает, что его обработка книги Вайзе Nucleus Logicae (объем которой возрос при этом с 72 до 850 страниц) методически использует круги для наглядного представления силлогистики.

В 1903 г. были обнаружены наброски Лейбница, которые содержат графические способы выделения правильных модусов силлогизмов с помощью прямоугольников и кругов. Впрочем, Венн справедливо замечает, что идея такого способа сама по себе могла прийти в голову не только Л. Эйлеру; что и де Морган, да и сам Венн, пришли к этому способу еще до того, как обнаружили его в «Письмах к немецкой принцессе» Эйлера. Именно Эйлеру Венн и приписывает заслугу введения этих методов в логику, датируя последнее 14-ым февраля 1761 года, когда было написано первое письмо на эту тему (102-е из «Писем к немецкой принцессе»).

Специальное внимание в своем историческом очерке Венн обращает на то расширение метода кругов Эйлера, которое принадлежит Б. Больцано (1837). Здесь идет речь о диаграмматическом представлении предложения «Та часть *A,* которая есть *B,* совпадает с той частью *C,* которая есть *D*», т. е. предложения *AB = CD.* Больцано изображает это предложение с помощью диаграммы, помещенной на рис. 22, где действительно пересечение *A* с *B* совпадает с пересечением *C* с *D.*

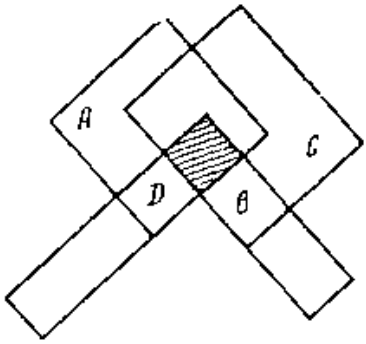


Рис. 22

По поводу этого изображения Венн замечал, что в нем не учитывается не запрещаемая условием задачи возможность того, что пересечение *B* с *D* не содержится ни в *A,* ни в *C,* т. е., что на диаграмме нет ячейки $\bar{A}B\bar{C}D$, которой,— если *A* и *C* пусты,— соответствует случай равенства *AB = CD* в силу того, что как *AB,* так и *CD* равны нулю.

(Диаграмма Венна, выражающая это предложение, приведена ниже, на рис. 35.) Сделаем два замечания:

1. $\bar{A}B\bar{C}D$ — есть обозначение пересечения классов $\bar{A}$, B, $\bar{C}$ и D, где черта над буквой означает дополнение к соответствующему классу.

2. Если включение одного класса в другой считать элементарной операцией, то равенство классов A и B $(A = B)$ определяется как $A \subset B$ и $B \subset A$: можно и наоборот, равенство считать элементарной операцией, тогда включение A в B (обозначение: $A \subset B$) можно определить так: существует класс v такой, что $A=vB$, в качестве v можно взять класс A, тогда $A = AB$.

Венн замечает, что кроме Больцано он больше ни у кого не встречал диаграммы четырех переменных.

У М. В. Дробиша и Э.Шредера Венн нашел то, что можно назвать трехкруговой диаграммой. Однако они только представляли комбинации или разбиепия классов, но не делали последующего шага для применения их в качестве основы для записи предложений.

Мих (1871) подошел близко к диаграмматическому представлению предложений; в одной из своих фигур он использовал три пересекающихся круга для представления разбиения класса; при этом некоторые из ячеек он штрихует. Правда, как отмечает Венн, штриховка не имеет у него пропозиционального значения,— не играет роли в диаграмматическом представлении предложений: концентрация все еще «эйлеровская».

Схемы Эйлера есть у Г. Плуке, который утверждает, что он изобрел этот метод независимо от Эйлера и раньше Эйлера; например, в 1759 году он представил модус barbara с помощью трех квадратов, последовательно включающих друг друга.

И. Кант и А. де Морган при диаграмматическом выражении предложений употребляли квадрат и круг, чтобы отличить субъект от предиката.

Латам в 1856 г. и Личман в 1864 г. использовали квадрат, круг и треугольник, чтобы различить три термина в силлогизме.

Эйлеровские круги получили широкую известность, но должно было пройти некоторое время, пока их стали использовать не только в качестве случайной иллюстрации. По-видимому, говорит Венн, первым логиком, который сознательно использовал круги Эйлера для представления отношений между предложениями, был Краузе. Венн читал работу Краузе 1828 года (первое издание работы вышло в 1803 году). Очевидно, пишет Венн, Краузе полностью понял отношение этих диаграмм к обычным предложениям.

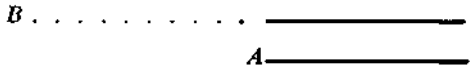
Для представления отношений между предложениями диаграммы использовались у Мааса (1793), представлявшего объемы понятий треугольниками с одной «подвижной» стороной. Венн замечает, что работа Мааса была не понята У. Гамильтоном, который писал о Маасе, что «это единственная попытка иллюстрировать логику не с помощью отношения геометрических величин, а с помощью отношений геометрических отношений — углов».

Кроме метода диаграмм, Венн рассматривает способ И. Ламберта, у которого терминам (классам) соответствуют отрезки прямых. Для того чтобы представить включение *С* в *В,* отрезок, соответствующий С, Ламберт рисует под отрезком, изображающим *В,* причем длина отрезка *С* меньше длины отрезка *В;* если классы не пересекаются, соответствующие отрезки изображаются рядом. Так, модус celarent:

Никакие *В* не суть *А*
Все *С* суть *В*
Следовательно, никакие С не суть *А*

Ламберт представляет схемой (очевидно, он предполагает, что $C \subset B$, т. е. пересечение дополнения *С* с *В не* пусто):

Если нам неизвестно, распределен ли предикат, то используются точки. Например, «Все *А* суть *В»* изображается



Это означает, что *А*, во всяком случае, покрывает часть *В* и может покрывать остаток, $A \subseteq B$.

Пунктирные линии используются также для изображения частных суждений. Так, Ламберт представляет «Некоторые *А* суть *В*» в виде

Венн, однако, считает, что схемы Ламберта явно хуже схем Эйлера.

В исторических очерках Венн отмечает также схему Велтона (1891), которая, с одной стороны, напоминает схему Ламберта, поскольку используются отрезки прямых, а с другой стороны,— в отношении интерпретации — метод Венна (для каждой возможной альтернативы оставляется место, которое заполняется тем или иным образом; о методе Венна далее).

Легко видеть, что все упомянутые авторы варьируют, в сущности, одну идею: все они в большей или меньшей степени приближаются к тому, что Э. Шредер называл **исчислением областей некоторого**

многообразия. Свойственные такому подходу недостатки уже обсуждались нами на примере кругов Эйлера. Эти недостатки, еще терпимые в рамках силлогистики, таковы, что к задачам математической логики, несравненно более сложным, чем получение модусов силлогизма, подход этот оказывается неприменимым. Поэтому в дальнейшем мы иметь с ним дела не будем; все это было сказано исключительно в пропедевтических целях.

Математическая логика должна была создать свой собственный графический метод; таковым и является метод диаграмм Венна.

Заметим, кстати, что в литературе имеется тенденция смешивать круги Эйлера и диаграммы Венна. Например, в книге Саппса есть раздел под названием «Диаграммы Венна», но начинается он с изображения эйлеровых кругов, причем попутно объясняется, что идея использовать круги принадлежит Эйлеру, Венн же лишь внес улучшения в его метод. Такая тенденция совершенно не оправдана, ибо **метод диаграмм Венна основывается на существенно иной идее, чем метод Эйлера, и общего у них только то, что оба являются графическими, оба связаны с геометрическим представлением объемов понятий, или классов, на части плоскости.**

Как мы видели, круги Эйлера возникают в традиционной силлогистике для ее нужд; к становлению математической логики они не имеют никакого отношения или весьма отдаленное. Диграммы же Венна создаются для обслуживания математической логики, и лежащая в основе их идея — **идея разложения на конституенты** — является одной из центральных в алгебре логики и вряд ли могла бы быть выдвинута без связи с этой последней. Во всяком случае, и И. Ламберт и Хр. Землер, у которых она впервые встречается, считаются (и по праву) прямыми предшественниками математической логики.

Но та математическая логика, в которой и для которой был создан метод венновских диаграмм, была по своим задачам и средствам существенно отлична от сегодняшней. Поэтому нам придется сделать довольно обширный экскурс в XIX в., чтобы выяснить, какие задачи ставились в алгебре логики — ибо такова была первоначальная форма математической логики,— как они решались и как при этом «работали» диаграммы Венна.

4.1.2. Постановка задач в алгебре логики XIX в. Способ решения логических уравнений по Булю

Как уже было отмечено, **математическая логика возникла первоначально как алгебра логики**. В действительности давно было

ясно, что **решение задач с помощью алгебраических уравнений состоит в том, что (1) информация, содержащаяся в условии задачи, выражается на специальном языке алгебры в виде некоторых равенств («уравнений»), (2) оперируя с которыми по определенным правилам, мы получаем из данной информации такую, которая дает ответ на интересующие нас вопросы, ответ, логически следующий из того, что нам дано.** Иными словами, было ясно, что **решение уравнений есть в действительности вывод логических следствий из информации, содержащейся в условии задачи**. Естественно, рано или поздно должен был возникнуть вопрос о том, нельзя ли вообще вывод логических следствий свести к решению уравнений, расширив соответственно возможности символического языка алгебры. По существу, именно так подходил к этому уже Лейбниц. Но первая довольно удачная попытка этого рода принадлежит Дж. Булю. Именно Буль истолковал впервые обычные операции арифметики в новом смысле: «качественном» (в противоположность количественному), как говорили логики конца прошлого века,— например П. С. Порецкий.

Поскольку термин «объем понятия» был уже достаточно употребительным, и с объемами (обозначая их буквами *a, b, . . ., u,* например) можно было оперировать, складывая их друг с другом, вычитая из некоторого объема его часть, ограничивая какой-нибудь объем частью, входящей одновременно в другой объем («умножая» их друг на друга), дополняя объем части до некоторого целого и т. п., «объемная» (теоретико-множественная) интерпретация логических операций, как объединения, пересечения, дополнения классов, стала вскоре наиболее распространенной, хотя у самого Буля это было еще не так. Буль разрешал, так сказать, «физически складывать» только классы, не имеющие общих элементов (в исчислении высказываний «сложение» по Булю соответствует так называемой «строгой дизъюнкции», или,— если обозначать «истину» цифрой 1, а «ложь» — цифрой 0,— сложению по модулю два, где сумма двух нечетных чисел есть четное число). Но умножение объемов *a* и *b* и у Буля соответствовало пересечению классов (т. е. классу элементов, общих как *a,* так и *b*).

Если мы будем обозначать, как это делалось уже в прошлом веке после Буля, операции объединения, пересечения и дополнения знаками +,- (или отсутствие знака) и (или'), то естественно придем к заключению,— исторически так оно и было,— что для объединения и пересечения имеют место обычные законы коммутативности и ассоциативности, а также законы дистрибутивности каждой операции по отношению к другой, законы идемпотентности *(a + a = a, aa = a),*

поглощения *(a + ab = a, a (a + b) = a)* и др. Естественно ввести, как это и сделал Буль, термины «1» и «О» для обозначения, соответственно, всего «мира речи» («универсума») и пустого множества, определить дополнение к объему *a*, т. е. $\bar{a}$, равенствами: $a + \bar{a} = 1, a\bar{a} = 0$. Заметим, что, для всякого *a, a*+1= *a*•1= *a, a* + 0 = *a, a*•0 = 0; если *a*+*b*=0, то *a* =1 и *b* = 0; если *ab* = 1, то *a* = 1 и *b* = 1. На всем этом мы позволим себе пока не останавливаться более подробно, тем более, что это уже широко известно. Отметим лишь, что, так как $x + \bar{x} = 1, 1 \cdot P = P$, то всякое выражение *P* можно представить в виде $P(x + \bar{x})$, или, пользуясь законом дистрибутивности, в виде $Px + P\bar{x}$ — вообще, в виде некоторой функции от *x,* такой, что $f(x) = Ax + B\bar{x}$. Полагая в этом равенстве $x = 1$, мы получим $f(1) = A$, и, полагая $x = 0$,— $f(0) = B$, т. е. $f(x) = f(1)x + f(0)\bar{x}$.

Аналогично, $f(x, y) = f(1, y)\, x + f(0, y)\, \bar{x}$, или, так как $f(1, y) = f(1,1)y + f(1, 0)\, \bar{y}, f(0, y) = f(0, 1)y + f(0, 0)\, \bar{y}$, то $f(x, y) = f(1, 1)\, xy + f(1, 0)\, x\bar{y} + f(0, 1)\, \bar{x}y + f(0, 0)\bar{x}\bar{y}$; $f(x, y, z) = f(1, 1, 1)\, xyz + f(1, 1, 0)\, xy\bar{z} + f(1, 0, 1)\, x\bar{y}z + f(1, 0, 0)\, x\bar{y}\bar{z} + f(0, 1, 1)\, \bar{x}yz + f(0, 1, 0)\, \bar{x}y\bar{z} + f(0, 0, 1)\, \bar{x}\bar{y}z + f(0, 0, 0)\, \bar{x}\bar{y}\bar{z}$ и т. д. Эти законы «разложения по конституентам» также были подмечены уже Булем, оперировавшим с равенствами, пользуясь правилом замены равным и вытекающими из него законами рефлексивности *(a = a),* симметричности (если *a = b,* то *b = a)* и транзитивности (если *a = b* и *b = c,* то *a = c)* равенства. Но Буль шел и дальше. Он полагал $\bar{a} = 1 - a$ и переносил в алгебру логики обычные правила решения арифметических уравнений. Так, уравнение

$$Ax + B\bar{x} = 0, \tag{1.1}$$

где *A* и *B* — выражения, не содержащие *x,* он решал следующим образом: заменяя $\bar{x}$ на 1 — *x,* получал $Ax + B(1 - x) = 0$ или $(A - B)x + B = 0$, откуда «заключал», что $x = \frac{B}{B - A}$. Смысл выражения $\frac{B}{B - A}$ раскрывался затем Булем с помощью разложения по конституентам:

$$\frac{B}{B - A} = f(1, 1)\, AB + f(1, 0)\, A\bar{B} + f(0, 1)\, \bar{A}B + f(0, 0)\, \bar{A}\bar{B},$$

где $f(i, j)$ есть $\frac{i}{j - i}$ (где $i = 0, 1$; $j = 0, 1$). Формальное вычисление этих коэффициентов дает

$$\frac{B}{B - A} = \frac{1}{0} AB + \frac{0}{-1} A\bar{B} + \frac{1}{1} \bar{A}B + \frac{0}{0} \bar{A}\bar{B}.$$

Лишенное смысла выражение $\frac{1}{0}$ Буль заменял нулем и, таким образом, получал ответ в виде

$$x = \bar{A}B + \frac{0}{0}\bar{A}\bar{B}, \qquad (1.2)$$

где $\frac{0}{0}$ толковалось им как означающее произвольный класс.

Не удивительно, что такого рода «решение» задачи не могло представляться достаточно убедительным. Почему $\frac{1}{0}$ равно 0? В каком смысле $x = \bar{A}B + \frac{0}{0}\bar{A}\bar{B}$ есть решение уравнения (1.1)? Что следует вообще понимать под делением *A* на *B?*

Если мы подставим в уравнение (1.1) вместо *x* выражениеи будем оперировать с полученным выражением

$$\bar{A}B + \frac{0}{0}\bar{A}\bar{B}$$

как обычно в арифметике, то будем иметь

$$A\left(\bar{A}B + \frac{0}{0}\bar{A}\bar{B}\right) + B\left(\bar{A}B + \frac{0}{0}\bar{A}\bar{B}\right) = A\bar{A}B + \frac{0}{0}A\bar{A}\bar{B} +$$

$$+ B\left(1 - \bar{A}B - \frac{0}{0}\bar{A}\bar{B}\right) = B - \bar{A}B = B\left(1 - \bar{A}\right) = BA.$$

Таким образом, если *AB* не есть 0, то найденное Булем выражение для *x* не удовлетворяет уравнению (1.1). Наоборот, если *AB* есть 0, то действительно при подстановке произвольного объема на место $\frac{0}{0}$ выражение (1.2) удовлетворяет уравнению (1.1).

Все это уже было известно Булю, который говорил о равенстве *AB* = 0, или «результате исключения *x*» из уравнения (1.1), как о необходимом и достаточном условии разрешимости уравнения (1.1). Однако на вопросы о том, что значит вообще разделить *A* на *B* и что дает право отбросить те или другие члены при таком делении, Буль не умел ответить. Все это очень интриговало логиков, знакомившихся с работами Буля, и все они старались получить ответ на эти вопросы. Таким образом, для проблематики алгебры логики конца прошлого века особенно существенными оказались вопросы:

1. Какие операции следует вводить в алгебре логики?
2. Имеют ли в ней, в частности, смысл обратные операции, и, если имеют, то какие?
3. Можно ли пользоваться в логике выражениями вида $\frac{0}{0}$? И, если да, то как именно?
4. Что значит, вообще, решить логическое уравнение и исключить неизвестное из него?

У нас нет возможности проследить здесь послебулевское развитие алгебры логики во всех подробностях. Важнейшие этапы таковы: У. С. Джевонс предложил в 1873 г. свой метод решения логических задач, основанный, по существу, на некотором переборе, использующем булевы конституенты (Джевонс именует их альтернативами). Затем в 1877 г. Э. Шредер дал первый строгий метод решения логических уравнений. Другой метод предложил в 1882 г. П. С. Порецкий; возникшая между ним и Э. Шредером дискуссия о том, что значит решить логическое уравнение, весьма любопытна; мы остановимся на ней и некоторых смежных вопросах в приложении. Остальные вопросы (1—3) относятся — все — к языку алгебры логики и, особенно, к ее операциям. Наибольшие трудности были связаны при этом с обратными операциями: «вычитанием» и «делением», хотя, как мы уже видели в связи с Булем, полной ясности не было даже насчет операции сложения. Желание перенести в логику по возможности большее число законов обычной арифметики должно было привести — и действительно привело Буля — к тому, что он отдал предпочтение истолкованию сложения как строгого «или». Как известно, построение логики — исчисления высказываний — как арифметики вычетов по модулю 2, т. е. как языка с операциями сложения (по Булю), умножения и константой «1», было выполнено с полной логической строгостью Жегалкиным и Стоуном только в двадцатых годах XX века. С другой стороны, **наличие двойственности между сложением и умножением при понимании их, соответственно, как объединения и пересечения (классов),** очень облегчавшее оперирование с логическими формулами и, особенно, простота геометрической интерпретации этих операций при таком их истолковании — вели к тому, что именно последнее и стало наиболее распространенным.

В истории алгебры логики интересно вообще, что начавшись с попыток перенести в логику все операции и законы арифметики, она постепенно приводила к тому, что логики начинали сомневаться не только в правомерности, но и просто в целесообразности такого переноса, начинали пользоваться более удобными, и более специфическими именно для логики операциями и законами. С этой точки зрения особенно интересна история обратных операций, которыми все логики XIX в.,— в том числе и такой крупный не только логик, но и алгебраист, как Э. Шредер,— еще очень много занимались.

В этой книге мы сосредоточим свое внимание на Джоне Венне. Дело не только в том, что он является создателем диаграммного метода. Венну принадлежит монография «Symbolic logic», которая систематизирует (хотя и не столь полно, как фундаментальный труд Э.

Шредера и не столь математично, как работа П. С. Порецкого) материал, накопленный в алгебре логики к концу XIX столетия. Достойно внимания также то обстоятельство, что Венн сумел ответить на вопросы 2 и 3, истолковав и булеву операцию деления, и символ $\frac{0}{0}$.

4.1.3. Символический язык Венна

Монография «Symbolic logic», первое издание которой вышло в 1881 году, второе — в 1894 году, является основным сочинением в серии работ Венна, посвященных обоснованию и развитию алгебры логики Буля. Термин «Symbolic logic», как отмечается в примечании к книге Чёрча, был, по-видимому, также впервые применен Венном, хотя еще Буль говорил о «Symbolical reasoning».

В журнале «Mind» за 1881 год об этом труде Венна мы читаем: «Цель этой выходящей вскоре в свет книги — дать полное представление о природе и объектах той системы логики, начало которой было положено в основном Булем. Книга содержит критическое разъяснение используемых принципов, выясняющих отношение этой системы к обычной логике и степень предполагаемого ею обобщения последней. Дается некоторое историческое освещение более ранних попыток этого рода, начиная со времен Лейбница, а также задач и употребления метода диаграмм в логике».

Сам Венн оценивает свое произведение следующим образом: он считает в нем «характеристическим и оригинальным» — «Тщательное исследование символической логики как целого, ее отношения к обычной логике и обычному мышлению и языку; установление и объяснение каждого общего символического выражения и правила на чисто логических основах вместо того, чтобы искать главным образом его формальное обоснование, и изобретение и употребление новой схемы диаграмматического обозначения, которое должно быть в истинной гармонии с нашими обобщениями».

«Символическая логика» Венна состоит из введения и двадцати глав (речь идет о втором, дополненном издании, в котором Венн дает ответы на некоторые замечания Э. Шредера и других авторов, относящиеся к первому изданию этой книги Венна). В каждый раздел книги включено большое число примеров.

Первой задачей символической логики по Венну является задача создания особого символического языка, который должен содействовать «расширению возможностей применения наших логических процессов при помощи символов». Однако прежде, чем заняться установлением такого языка, Венну нужно было решить

вопрос о том, что именно должно быть выразимо на этом языке. В этой связи Венн замечает, что **предложения, которые мы высказываем, могут иметь разный смысл. Они могут приписывать предмету или отрицать у него наличие какого-нибудь свойства. В предложении имеется субъект (класс предметов) и предикат, приписываемый предметам этого класса (или отрицаемый у них).** Именно так, т. е. посредством прецидирования, говорит Венн, и строятся предложения в силлогистике Аристотеля. **Четыре формы *a, e, i, o* элементарных предложений этой теории являются при этом хорошо выражающими смысл общих и частных, утвердительных и отрицательных предложений, и значение такой теории нельзя отрицать**. Наряду с ней возможна, однако, **другая интерпретация смысла предложений: объемная, как мы сказали бы теперь, которая состоит в том, что элементарное предложение рассматривается как выражение отношения взаимного включения или исключения двух классов объектов.** Но **такого рода отношений имеется пять, и они выражаются соответственно диаграммами** (рис. 1), которым приходится ставить **в соответствие в общем случае уже не элементарное, а сложное предложение в аристотелевом смысле**. Так, сложное предложение «Все *A* суть *B и* все *B* суть *A»* соответствует первой диаграмме: «Все *A* суть *B, но* некоторые *B* суть *не-A»* — второй диаграмме; «Все *B* суть *A, но* некоторые *A* суть не-*A*» — третьей диаграмме; «Некоторые *A* суть *B и* некоторые *A* суть *не-B и* некоторые *B* суть *не-A»* — четвертой диаграмме; и только пятой диаграмме соответствует элементарное предложение «Ни одно *A* не есть *B».* Наоборот, элементарному предложению «Все *A* суть *B*», даже в случае, когда рассматриваются только непустые классы, может соответствовать либо первая, либо вторая картинка. Элементарному предложению «Некоторые *A* суть *B*» — одна из первых четырех,— и неизвестно, какая.

Венн ставит перед собой задачу найти такой язык символической логики, на котором элементарные предложения силлогистики Аристотеля выражались бы также элементарными предложениями, и такую геометрическую интерпретацию для этих предложений, при которой всякому предложению, как простому, так и сложному, однозначно соответствовала бы некоторая картинка. При этом он сталкивается с рядом трудностей,— прежде всего, относящихся к истолкованию «универсума» («единицы») и дополнения, которые успешно преодолевает.

Даже в случае, когда мы имеем дело с хорошо известным нам понятием — таким, например, как «черный» — дополнение к нему, или, как Венн предпочитает говорить, его отрицание, недостаточно

определено. Действительно, что понимать под «не-черным»? — «Мы можем, — говорит Венн,— если угодно, определить этот класс как включающий все то, что не является черным:— включающий, например, геологический ледниковый период, притязания папства, последнее письмо Кларисы Харлоу и пожелания нашего отдаленного потомства. Но речь настолько ясно идет о цвете, что никакой разумный человек не станет сомневаться в том, что именно имеется в виду. И аналогично в большинстве случаев».

Иными словами, чтобы дополнение к классу имело точный смысл, нужно знать, что подразумевается под «целым» — «миром речи», или «универсумом». Никакого единого универсума в логике, по Венну, нет. Вопрос о выборе универсума,— как и тесно связанный с ним вопрос о геометрической интерпретации предложений,— относится уже не к логике, а к ее приложениям. Законы формальной логики,— особенно такие, которые относятся к дополнениям,— не действуют автоматически: если не позаботиться о надлежащем выборе универсума и соответствующем уточнении понятий, они потеряют смысл; «... когда мы говорим, что *«A* есть не-*B»,* и пытаемся рассматривать это не-*B* как атрибут, мы налагаем на наше понятие неопределенную широту его объема; и поэтому, когда мы не относим его к какому-нибудь ограниченному универсуму, мы должны признать, что наше суждение является в отношении к его предикату бесконечным или неопределенным».

Само по себе «все» очень неопределенное выражение и не имеет точного смысла. Чтобы можно было пользоваться логикой, его нужно уточнить, но это можно делать по-разному. «Мы можем ограничить наше применение терминов «хороший» и «не-хороший» лондонскими кэбами с нечетными номерами, и всякое логическое правило будет при этом оставаться столь же верным, как и в случае, если бц мы выбрали менее абсурдный род универсума». В логике классов существенно только, что «все, что исключается из *x,* включается в не-*x*. Только, когда мы начинаем исследовать, что имеется в виду под «все», возникает вопрос о границах (для не-а*),* и это практический вопрос, предполагающий интерпретацию наших данных».

Наоборот, после того как «универсум» (или 1; на мотивах, приводимых Венном в оправдание выбора единицы в качестве символического обозначения для универсума, мы позволим себе не останавливаться) уже уточнен, никакого существенного различия между положительными и отрицательными понятиями не остается: не-*X* можно принять за некоторое *Y,* и тогда *X* выразится как не-*Y*. Это обстоятельство широко используется Венном в его диаграммах, которые всегда относятся к некоторому специально уточненному универсуму. За последний при

этом удобно взять какую-нибудь простую геометрическую фигуру,— например, круг или квадрат. Вопрос о том, что лежит за пределами этого универсума, имеет для нас, по Венну, не больше смысла, чем вопрос о том, «что находится в поле нашего зрения *сзади* нашей головы». Но после того как универсум выбран, и дополнение таким образом определено, мы можем уже однозначно поставить в соответствие всякому элементарному предложению аристотелевой силлогистики его диаграмму. Поскольку Венн при этом вводит новую трактовку этих элементарных предложений, значительно более близкую к являющейся теперь обычной, и поскольку предлагаемая им интерпретация элементарных (в смысле Аристотеля) суждений для его метода является очень существенной, мы остановимся на этом более подробно.

Займемся прежде всего суждением вида «Все *A* суть *B*». Если за универсум выбрать круг (обозначим его через *U*)

и разбить двумя другими кругами *A*, *B* на четыре непересекающиеся друг с другом части AB, $\bar{A}B$, $A\bar{B}$, $\bar{A}\bar{B}$, то отношение включения *A* в *B* будет однозначно выражено тем, что пересечение *A* и $\bar{B}$ пусто: за пределами *B* нет никакого *A*. Если, следуя Венну, суждению «Все *A* суть *B*» будет однозначно поставлена в соответствие картинка, изображенная на рис. 23, то эта картинка будет одной и той же независимо от того, какой из случаев — первый или второй (рис. 1) — имеет место.

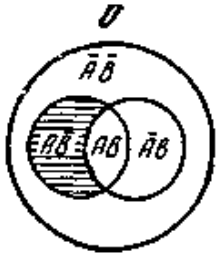


Рис. 23

Больше того, эта картинка будет правильно выражать суждение «Все *A* суть *B*» даже в том случае, когда класс *A* пуст.

В оправдание того, что общее утвердительное суждение «Все *A* суть *B*» он записывает в отрицательной форме $A\bar{B} = 0$, т. е. как «Не существует предметов, принадлежащих *A* и не-*B*», Венн приводит следующие соображения. Рассматриваемое предложение, говорит он, можно интерпретировать и утвердительно,— например, в форме Буля: $A = vB$, где v — неопределенный класс (иногда вместо v Венн пишет

символ $\frac{0}{0}$), и в форме Лейбница: $A = AB$. Но (как говорит Венн еще в первом издании книги) утвердительная формулировка имеет условный и сложный характер, в то время как отрицательная — проста и безусловна: когда мы говорим, что «Все *A* суть *B»,* то мы не знаем, существуют ли какие-нибудь *A* или *B;* но когда мы говорим, что нет таких *A,* которые были бы не-*B*, то наша интерпретация вполне определенна; это особенно важно в более сложных случаях, расхождение же с обычной речью не существенно — на то и символы, чтобы говорить на другом (не на обычном) языке.

То обстоятельство, что для пустого *A* оба суждения «Все *A* суть *B»* и «Все *A* суть не-*B*», или в соответствующей интерпретации для исчисления высказываний (которое, как и другие логики того времени, Венн трактует как частный случай исчисления классов): «Если *A,* то *B*» и «Если *A,* то не-*B*», должны считаться одновременно истинными, не смущает уже Венна, который пишет: «Мы должны допустить, что фраза «*x* влечет *y»* не предполагает, что известна какая-либо связь между рассматриваемыми фактами или что одно предложение формально выводимо из другого». И он поясняет это обстоятельство на примере диспута между логиками, имевшего место в связи с задачей, предложенной автором «Алисы в стране чудес» Льюисом Кэрролом, которую Венн называет поэтому «задачей Алисы». Эта задача состоит в следующем:

Имеются два предложения: *A* и *B*. И пусть установлено, что

(I) Если истинно *A,* то истинно *B*.

Есть еще одно предложение *C* такое, что

(II) Если истинно *C,* то, если *A* истинно, *B* неистинно.

Может ли при таких обстоятельствах *C* быть истинным?

Или та же задача в другой формулировке: В доме три человека: *A, B* и *C* (Allen, Brown, Carr), которые могут выходить и не выходить при условии, что

(I) *A* никогда не выходит без *B,*

(II) Если *C* выходит, то, если *A* выходит, *B* не выходит.

Спрашивается, может ли при таких условиях *C* выйти когда-нибудь.

На более привычном нам теперь языке исчисления высказываний эти условия можно записать так:

$$\text{(I)} \quad (A \supset B),$$

$$\text{(II)} \quad (C \supset (A \supset \neg B)).$$

И ясно, что, если *A* ложно, то оба высказывания (I) и (II) истинны, так как истинны и $(A \supset B)$, и $(A \supset \neg B)$.

Именно это и демонстрирует, пользуясь установленным им аппаратом, Венн.

Общее отрицательное суждение «Никакие *A* не суть *B*» Венн, естественно, также записывает в отрицательной форме *AB*=0 и сообщает, что в форме Буля это предложение имеет вид $A = v(1 - B)$. Этому суждению у Венна однозначно соответствует диаграмма, приведенная на рис. 24.

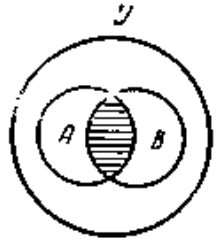


Рис.24

Частное суждение «Некоторые *A* суть *B*» Венн представляет в форме *AB* > 0 (знак > служит для выражения непустоты класса *AB)*. При этом он подчеркивает экзистенциональный характер такого предложения. В форме Буля, говорит, он, это предложение имеет вид: $\frac{0}{0}A = \frac{0}{0}B$, где $\frac{0}{0}$ — неопределенный, но не пустой класс.

Такое представление Венн считает особенно неудачным. Разбирая эту и другие формы записи частных предложений, Венн прежде всего показывает, что они недостаточно выражают экзистенциональный характер частных суждений. Так, в первом издании книги он пишет: «хотя *AB = AC* может быть прочитано как «Некоторые *B* суть *C*», но неверно, что всякое предложение «Некоторые *B* суть *C*» может быть сформулировано как *AB = AC*. Я не вижу никакой формы, которая покрывала бы все частные предложения, за исключением той, которая трактует их как утверждения существования, и ограничивается объявлением, что имеются такие *B,* которые являются *C*.

После того что уже было сказано, вряд ли требуется повторять, что $\frac{0}{0}$ не есть эквивалент для «некоторые». Быть может, наилучшая краткая формулировка его значения состоит в том, что оно представляет собой признание полного неведения в отношении термина, к которому $\frac{0}{0}$ приставлено».

Непустоту класса Венн изображает на диаграмме, помещая звездочку в соответствующей этому классу части диаграммы. Таким образом, суждению «Некоторые *A* суть *B»,* во всех возможных случаях соответствует одна единственная диаграмма (рис. 25).

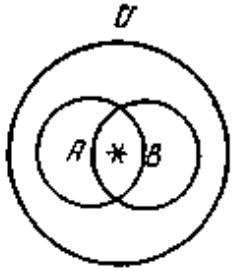


Рис. 25

Отрицательное частное суждение «Некоторые *A* не суть *B*» выражается — и опять-таки совершенно однозначно — диаграммой, представленной на рис. 26.

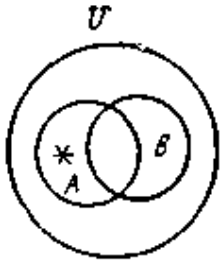


Рис. 26

Мы видим, таким образом, что Венну действительно удалось очень просто, по существу, справиться с задачей однозначно отобразить любое элементарное предложение соответствующей ему картинкой.

В качестве символов классов Венн употребляет буквы латинского алфавита. Операцию объединения он обозначает знаком плюс (+), операцию исключения (вычитания) одного класса из другого (при условии, что первый есть часть второго) — знаком минус (—), операцию ограничения (пересечения, или произведения классов) — знаком умножения (×) или без знака, операцию деления класса x на класс y он записывает в виде $\frac{x}{y}$. При описании операций подчеркивается как их аналогия с соответствующими операциями в математике (арифметике и алгебре), так и отличие от последних.

Таким образом, в логике классов Венна используются **выражения (термы)**, порождающие правила которых могут быть сформулированы следующим образом:

Пусть $a_1, \ldots, a_n$ — все графически различные буквы алфавита A_n; 0, 1 — буквы алфавита A.

1. Если $\beta \in A_n A$ (β есть буква, принадлежащая объединению алфавитов A_n и A), то β считается термом. Если $\beta \in A_n$, то β обозначает элементарный класс. Если β есть 1, то β — универсум. Если β есть 0, то β — пустой класс.
2. Если Q — терм, то $\bar{Q}$ считается термом.
3. Если Q, R—термы, то (QR), $(Q + R)$, $\frac{Q}{R}$, $(Q - R)$ считаются термами.

На классы, участвующие в операциях, Венн накладывает ограничения. Поэтому не всякий терм обозначает класс. Ниже мы остановимся на операциях над классами, причем будет показано, что операции вычитания и деления могут быть из логики классов исключены.

Пустоту класса E Венн обозначает $E = 0$, непустоту — $E > 0$. Равенство классов E и F Венн понимает как выражение $E\bar{F} + \bar{E}F = 0$, а неравенство $E \neq F$ — как выражение $E\bar{F} + \bar{E}F > 0$.

Такую форму записи предложений Венн называет отрицательной и неоднократно подчеркивает ее экзистен-циональный характер:

1. В предложении $E = 0$, где E — некоторый класс E, говорится о том, что не существует предметов, принадлежащих классу E.
2. Предложение $E > 0$ говорит, что существуют предметы, принадлежащие классу E.

Таким образом, знаки =, > Венн использует для построения формул логики классов. Порождающие правила при этом могут быть сформулированы (на современном языке) следующим образом:

1. Если A, B — классы, то $A = B$ считается формулой.
2. Если A — класс, то $A > 0$ считается формулой.

По Венну формула $A=B$ эквивалентна формуле

$$A\bar{B} + \bar{A}B = 0.$$

Формулы S и T Венн считает эквивалентными, если они говорят о пустоте (или непустоте) одного и того же класса.

Знак неравенства ≠ Венн фактически использует только для сокращения записи: предложение $A \neq B$ говорит, что класс $A\bar{B} + \bar{A}B$ не пуст, т. е. $A\bar{B} + \bar{A}B > 0$.

Венн формулирует правило для построения отрицаний: если $A = 0$, то отрицание этого предложения имеет вид: $A > 0$, и обратно. Например, отрицание $xy = 0$ есть $xy > 0$.

Объединение двух данных классов A и B возможно, как пишет Венн, например, в трех смыслах. Первый случай соответствует строгой дизъюнкции, во втором случае объединение понимается как в

современной теории множеств (в исчислении высказываний этому соответствует неисключающая дизъюнкция), в третьем случае элементы общей части классов *A* и *B* учитываются дважды. Эти различные подходы к объединению классов Венн иллюстрирует соответственно следующими примерами.

1. Амнистируются преступники, совершившие только одно из двух преступлений *A* и *B*. Класс амнистируемых в этом случае есть объединение классов преступников, совершивших *A* и совершивших *B,* в смысле строгой дизъюнкции.
2. Получают рождественские подарки почтальоны и служители прихода, при этом каждый человек может получить только один подарок; в этом случае мы имеем дело с неисключающей дизъюнкцией: обычным объединением классов почтальонов и служителей прихода.
3. В качестве примера третьего подхода к объединению Венн приводит случай, когда учитывается общее число оштрафованных, нарушивших правила *A* или *B,* причем лицо, нарушившее оба правила, штрафуется дважды. В классе, состоящем из всех оштрафованных, некоторые элементы учитываются дважды.

Венн отдает предпочтение второму случаю, то есть, если нет специальной оговорки, операция «сложения» соответствует у него операции объединения классов (или неисключающей дизъюнкции исчисления высказываний).

Анализируя операцию пересечения («ограничения»), Венн специально останавливается на случаях, когда пересечение двух классов фактически не ограничивает одного из них, например, случай совпадения *xy* с *y*.

Операция исключения (вычитания) классов, рассматриваемая Венном как обратная к операции объединения классов, определяется им только для случая, когда $x \subset z$; она рассматривается как дающая тот наименьший класс y_0, для которого $x + y_0 = z$. Поэтому, если для какого- нибудь *y*

$$x + y = z, \quad \text{то } y_0 = z - x = \bar{x}y = z\bar{x},$$

где $\bar{x}$ — дополнение класса *x* до универсума. Ясно, что, если ж и у не пересекаются, то y_0 совпадает с *y,* т. е. из $x+y = z$ можно заключить, что $y = z - x$.

Местами Венн совершенно некритически переносит свойства этой операции из арифметики в логику. Так, теорема Гаубера [которая в простейшем случае (у Венна рассматривается более общий случай) гласит: если род *A* подразделен на виды *x* и *y,* а также на виды α и β , не имеющие общих элементов, и мы знаем, что все *x* суть α, все *y* суть β,

то и наоборот, все α суть *x,* все β суть *y*] у Венна доказывается (в рассматриваемом здесь случае) следующим образом.
Пишется

$$A = \alpha + \beta,$$
$$A = x + y,$$

затем из верхнего равенства почленно вычитается нижнее, что записывается в виде

$$0 = (\alpha - x) + (\beta - y), \qquad (1.3)$$

и, так как нам дано, что *x* включено в α, а *y* — в β, откуда $\alpha - x = \alpha\bar{x}$, $\beta - y = \beta\bar{y}$, то (1.3) переписывается в виде $0 = \alpha\bar{x} + \beta\bar{y}$, откуда, далее, $\alpha\bar{x} = 0$, $\beta\bar{y} = 0$, т. е. все α суть *x* и все β суть *y*.
В своей «Алгебре логики» Э. Шредер критиковал Венна за такое «свободное» обращение со знаком минус в логике. Однако Венн и во втором издании «Символической логики» не счел нужным внести в это доказательство какие-либо существенные изменения. Правда, нетрудно заметить, что теорема Гаубера могла быть доказана Венном правильно, т. е. разработанные им методы (например, графические) позволяют из того, что $\alpha\beta = 0$, $x\bar{\alpha} = 0$, $y\bar{\beta} = 0, \alpha + \beta = x + y$, вывести $\alpha\bar{x} = 0$, $\beta\bar{y} = 0$, $xy = 0$ (решение этой задачи с помощью диаграмм Венна дано ниже).
Особое внимание Венн уделяет операции деления, которая у него является обратной логическому умножению (пересечению, или ограничению). Операция деления используется в логике Булем, но он «не делает никакой попытки объяснить свое употребление знака деления в логике», и Венн рекомендует обращаться к трудам «более философских математиков», среди которых называет прежде всего «Тригонометрию и двойную алгебру» А. де-Моргана, а затем «Арифметику и алгебру» Э. Шредера, «Лекции о комплексных числах» Г. Ганкеля и «Основы анализа» Р. Липшица. Венн стремится дать «объяснение принципов логического исчисления в полной зависимости от основ математического исчисления». Подчеркивая, что принципы логики должны иметь самостоятельное объяснение, Венн **ясно представляет связь между математикой и логикой.** Так, он пишет: «Я думаю, что Mr. Harley прав, полагая, что термин «математическая» употреблялся им (Булем) «в расширенном смысле как обозначающий науку о законах и комбинациях символов, и с этой точки зрения нет ничего нефилософского в трактовке логики как ветви математики вместо трактовки математики как ветви логики»»); и далее Венн продолжает:

«Буль и сам сказал то же самое: «Это просто факт, что последние законы логики,— те, на которых и можно только построить науку логики,— являются математическими по форме и выражению, хотя они не принадлежат к математике количества».

Результат операции деления *x* на *y* при условии, что класс *x* включается в класс *y*, т. е. при условии, когда «все *x* суть *y*», Венн представляет в виде $x + v\bar{x}\bar{y}$ (или в виде $xy + v\bar{x}\bar{y}$), где *v* — неопределенный класс. Если условие $x \subset y$ не выполнено, операция деления у Венна не определена, так как в этом случае класс *x* не может быть получен из класса *y* путем ограничения последнего каким-нибудь классом z. (Действительно, из *yz*=*x* следует непосредственно, что $\bar{y}yz = \bar{y}x$, т. е. — так как $\bar{y}y = 0$,— что класс $x\bar{y}$ пуст: «Все *x* суть *y*».)

Естественность такого представления Венн обосновывает не только тем, что произведение $x + v\bar{x}\bar{y}$ (или $xy + v\bar{x}\bar{y}$) на *y* равно *x* при условии, что *xy* = *x,* но и последующим анализом задачи.

В связи с вопросом об обратных операциях Венн приводит замечание Буля из книги «Дифференциальные уравнения», состоящее в том, что обратная **операция не есть в действительности операция в собственном смысле слова, поскольку <u>она не указывает последовательности шагов</u>, которые нужно выполнить.** Венн цитирует Буля: «Обязанность обратного символа состоит в том, чтобы предложить вопрос, а не в том, чтобы описать операцию. В его первоначальном смысле эта обязанность вопросительная, но не директивная».

Вопрос, который возникает в применении к делению, как обратной операции, Венн формулирует так: найти выражение для наиболее общего класса, который, если наложить на него ограничение, обозначенное через *y,* сведется в точности к *x.* Отсюда Венн, прежде всего, делает заключение, что весь класс *x* должен быть частью искомого класса, т. е., что задача может быть решена только в этом предположении. Так как, далее, слагаемое, содержащее множителем *y,* не может исчезнуть при умножении на *y,* то Венн заключает, что помимо *x* искомый класс может содержать еще только некоторую часть класса *xy.* Поскольку любая часть этого класса при добавлении к *x* дает класс, удовлетворяющий требованиям задачи, Венн вводит еще неопределенный множитель *v,* получая, таким образом, все возможные решения задачи. Результат деления *x* на *y* представляется у него поэтому в виде, как он пишет, «группы классов» $x + v\bar{x}\bar{y}$, где *v* — любой класс.

Ясно, что никаких других решений задача не имеет. В дальнейшем мы покажем, что решение этой задачи можно получить с помощью изложенного Венном общего метода решения логических уравнений.

Как и Шредер, Венн понимает при этом, что вводимая им «операция» деления в логике не нужна, так как при решении логических уравнений без нее можно обойтись и никакой логический процесс вообще не подсказывает необходимости в ней. Однако для аналогии с арифметикой Венн, в отличие от Шредера, сохраняет деление как особую операцию в теории классов.

В связи с этим подчеркнем, что не только операцию деления в логике вводить не обязательно, но и вообще можно обойтись без обратных операций: без деления и без исключения (вычитания). Операция вычитания может быть исключена в силу равенства $z - x = z\bar{x}$ (при $x \subset z$).

На протяжении всей книги Венн регулярно ссылается на историю математики и логики. Так, приведя высказывание Буля об обратных операциях, Венн добавляет: «Представляется удивительным, что автор, который так ясно сформулировал природу обратной операции в математике, никогда не предложил в явной форме какого-нибудь соответствующего объяснения в логике».

Встретившись с каким-нибудь вопросом, Венн не успокаивается, пока не разберется в нем с более общей точки зрения. Так, исследуя обратные операции, он ставит общий вопрос о том, как в паре противоположных операций выделить прямую и обратную, и предлагает при ответе на такой вопрос заведомо считать обратной ту (если таковая имеется), результат которой не вполне определен. В противном случае, т. е. если результат обеих операций однозначно определяется, выбор прямой и обратной операций можно делать по произволу.

Операции разделяются Венном на интерпретируемые и неинтерпретируемые. Венн, следуя Булю, рассматривает операцию деления как неинтерпретируемую. Венн приводит цитату из Буля в связи с неинтерпретируемостью деления: существует «цепочка вывода, ведущая нас через промежуточные шаги, которые нельзя интерпретировать, к конечному результату, являющемуся интерпретируемым», и далее Буль продолжает: «Употребление неинтерпретируемого символа $\sqrt{-1}$ в промежуточных процессах тригонометрии является иллюстрацией сказанного». Таким образом, наряду с О. Л. Коши и Дж. Буля можно считать предшественником идей, на которых базируется теория ε-символа у Д. Гильберта.

В дальнейшем, как пишет Венн, его точка зрения на деление изменилась,— результатом деления он считает некоторый, хотя и

неопределенный, класс. При этом Венн придерживается объемной интерпретации, которая ему представляется более ясной. В то же время он отмечает возможность интерпретации деления и «по содержанию», или по признаку. Такую интерпретацию дал Шредер в 1877 году. Как отмечает Венн, на самом деле она была уже у И. Ламберта, который говорил: «Отвлечем от понятия *A* частичное понятие *B* и назовем получающееся понятие *R*. Тогда мы имеем

$$R = \frac{A}{B} = ac$$

(например)». Здесь *A,* очевидно, предполагается имеющим признаки *a, b, c,* а *B* — имеющим признак *b*. Далее, однако, Ламберт замечает: «Но мы здесь занимаемся делением не в большей мере, чем занимаемся умножением при композиции».

Интерпретацию в терминах признаков Венн считает неприемлемой, так как неясно, что означает понятие «содержание», и он поэтому отдает предпочтение объемной интерпретации, которая особенно наглядна, когда речь идет о конечном числе объектов.

Опираясь на наглядное представление введенных им операций, Венн формулирует различные их свойства , такие, например, как

$$x(y + z) = xy + xz$$

$$(a + x)(b + x) = ab + x$$

Какой-либо попытки систематизировать эти законы, выбрав среди них более простые, с тем, чтобы, пользуясь ими, выводить другие,— в отличие от Шредера, который, как известно, уже заметил даже невыводимость дистрибутивных законов из, как мы сказали бы теперь, аксиоматики алгебраической структуры,— у Венна нет. По существу, он ограничивается только приведением примеров, иллюстрирующих эти законы. Мы позволим себе, поэтому не останавливаться подробнее на этих вопросах алгебры логики у Венна. Отметим лишь, что правила поглощения:

$$A = A + AB, \qquad A = A(A + B),$$

и правила выявления:

$$Ax + B\bar{x} = Ax + B\bar{x} + AB,$$

$$(A + x)(B + \bar{x}) = (A + x)\ (B + \bar{x})(A + B)$$

также имеются у Венна.

Особую роль в логике классов у Венна играет разложение выражений на конституэнты по данным переменным *x, y, z, . . ., u*. Венн делает это аналогично Булю, т. е., представляя выражение (для определенности отметим сучай трех переменных $x,\ y,\ z$) $f(x,\ y,\ z)$ в виде

$$f(x, y, z) = f(1, 1, 1)\, xyz + f(1, 1, 0)\, xy\bar{z} + f(1, 0, 1)x\bar{y}z +$$
$$+ f(1,\ 0,\ 0)\ x\bar{y}\bar{z} + f(0,\ 0,\ 1)\ \bar{x}\bar{y}z + f(0,\ 1,\ 1)\,\bar{x}yz +$$

$+ f(0, 1, 0)\ x\bar{y}z + f(0, 0, 0)\ x\bar{y}\bar{z},$ где $f(1,1, 1)$, $f(1. 1,0)$,..., $f(0, 0, 0)$ равны нулю или единице, а также в виде совершенной простой суммы (определяемой аналогично совершенной дизъюнктивной нормальной формуле в исчислении высказываний) или (двойственного) совершенного простого произведения (определяемого аналогично совершенной простой сумме с заменой знаков + и • друг на друга). Такие разложения Венн использует постоянно для доказательства своих утверждений. По существу, они же лежат в основе его графического метода диаграмм.

Здесь следует еще отметить, что помимо правил де Моргана для образования выражения, противоположного данному, Венн употребляет прием, состоящий в том, что противоположностью для выражения, разложенного по каким-нибудь переменным, например, для

$$Axy + Bx\bar{y} + C\bar{x}y + D\bar{x}\bar{y},$$

является выражение, получаемое заменой коэффициентов *A, B, C, D* соответственно на $\bar{A}, \bar{B}, \bar{C}, \bar{D}$, т. е. выражение

$$\bar{A}xy + \bar{B}x\bar{y} + \bar{C}\bar{x}y + \bar{D}\bar{x}\bar{y}.$$

Венн широко пользуется и тождеством Джонсона

$$(ac + \bar{a}d) = (a + d)(\bar{a} + c): \quad (1.4)$$

он иллюстрирует это тождество таблицей (табл. 1):

Таблица 1

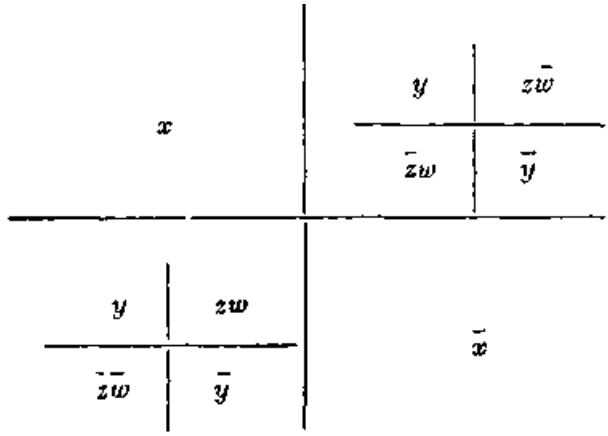
a	c
d	$\bar{a}$

по строкам члены в ячейках перемножаются, по столбцам — складываются. Алгорипмический характер табличного представления тождества (1.4) Венн демонстрирует на примере получения логических следствий посылок, заданных в положительной форме (т. е. в виде равенства единице). Пусть дано

$$xyz\bar{w} + x\bar{y}\bar{z}w + \bar{x}yzw + \bar{x}\bar{y}\bar{z}\bar{w} = 1. \quad (1.5)$$

С помощью таблицы (табл. 2), которая строится для формы (1.5) и читается как сумма (дизъюнкция) из произведений (конъюнкций) по строкам или как произведение (конъюнкция) из сумм (дизъюнкций) по столбцам:

Таблица 2



Венн легко получает равенство

$$xyz\bar{w} + x\bar{y}\bar{z}w + \bar{x}yzw + \bar{x}\bar{y}\bar{z}\bar{w} =$$
$$= (x + y + \bar{z}\bar{w})(x + \bar{y} + zw)(\bar{x} + y + \bar{z}w)(\bar{x} + \bar{y} + z\bar{w}). \quad (1.6)$$

В силу (1.5), правая часть (1.6) приравнивается также единице. Следовательно, все конъюнктивные члены (сомножители) правой части (1.6), записанные в положительной форме, являются логическими следствиями из (1.5). Этот метод может представлять интерес и сегодня, особенно при нахождении логических следствий из посылок.

4.1.4. Алгебраические методы решения логических уравнений и исключения неизвестных

Как уже было отмечено, логика Венна представляет собой логику классов, т. е. некоторый эквивалент исчисления одноместных предикатов. В этом исчислении имеется часть, в точности эквивалентная исчислению высказываний. Все логики второй половины XIX в., занимавшиеся алгеброй логики, выделяли прежде всего именно эту часть логики классов. Они делали так потому, что решение всех задач в этой части легко сводилось к решению уравнений. Здесь речь шла о решении таких задач, в которых среди посылок не было частных суждений. Частные суждения рассматривались Шредером, например, только во втором томе его трехтомной «Алгебры логики». Венн стал рассматривать частные

суждения более подробно лишь во втором издании своей «Символической логики», трактуя их как неравенства вида $A > 0$.

Центральной задачей логики Венн поэтому считает задачу решения уравнений. **Уравнением или системой уравнений при этом выражается информация, содержащаяся в посылках задачи. Решение уравнений у Венна состоит в таком преобразовании этой информации, которое позволяет дать ответ на какие-нибудь вопросы, относящиеся к терминам, фигурирующим в посылках.** Наиболее простым видом таких вопросов являются вопросы, предлагающие охарактеризовать любой из рассматриваемых терминов (или его дополнений) через остальные или часть остальных. Задача, состоящая в извлечении такой информации, которая относится только к части терминов, называется проблемой исключения и рассматривается как особая задача. **Обе задачи — и решения уравнений, и исключения неизвестных — оказываются неразрывно связанными между собой**. Они являются, конечно, некоторым обобщением алгебраических задач, относящихся к составлению и решению уравнений. В то же время Венн хорошо понимает, что даже так поставленные логические задачи не могут решаться простым переносом в логику обычных методов решения алгебраических уравнений. «Главное, чего тут следует опасаться,— пишет он,— это ошибок, происходящих из ассоциаций с математикой».

В этой связи он уделяет особое внимание методам решения логических уравнений, предложенным Булем, и «неуклюжим», как он их называет, формам выражения этих уравнений и их решений, встречающихся у Буля.

Всякое равенство

$$\mathfrak{A} = \mathfrak{B}$$

в алгебре классов, содержащее класс x, может быть приведено к виду:

$$Ax + B\bar{x} + C = 0. \qquad (1.7)$$

Это преобразование можно осуществлять следующим образом: во-первых, исходное равенство записать в отрицательной форме

$$\mathfrak{A}\bar{\mathfrak{B}} + \bar{\mathfrak{A}}\mathfrak{B} = 0 \qquad (1.8)$$

во-вторых, левую часть равенства (1.8) заменить эквивалентной ей простой суммой (являющейся в логике классов аналогом дизъюнктивной нормальной формы логики высказываний) и, наконец, в-третьих, в левой части последнего равенства выделить коэффициенты при x и $\bar{x}$.

(Из исходного равенства $\mathfrak{A} = \mathfrak{B}$ равенство (1.8) получается следующим образом. Умножив обе части равенства $\mathfrak{A} = \mathfrak{B}$ один раз

на $\bar{\mathfrak{x}}$, другой раз на $\bar{\mathfrak{y}}$, получим равенства: $\mathfrak{x}\bar{\mathfrak{x}} = \mathfrak{x}\bar{\mathfrak{x}}$, $\mathfrak{x}\bar{\mathfrak{y}} = \mathfrak{x}\bar{\mathfrak{y}}$, т. е. $\mathfrak{y}\bar{\mathfrak{x}} = 0$ и $\mathfrak{x}\bar{\mathfrak{y}} = 0$. Складывая почленно два последние равенства, поличим (1.8).)

Равенство (1.7) называют уравнением относительно буквы (неизвестной) *x*. В (1.7) *B, A, C* — выражения, не содержащие переменной *x*. Непосредственным следствием такого уравнения являются условия $C = 0$, $AB=0$, оказывающиеся, как нетрудно заметить, необходимыми и достаточными для того, чтобы уравнение (1.7) имело решение, т. е., чтобы существовал такой класс *x* (хотя бы и пустой), для которого оно выполняется.

В п.1.41.2 уже показывалось, как Буль решал такие уравнения. Его результат:

$$x = \bar{A}B + \frac{0}{0}\bar{A}\bar{B}. \tag{1.9}$$

Венн указывает, что этому решению можно придать рациональную форму. Действительно, из условия задачи, т. е. из уравнения (1.7), немедленно следует, что $Ax = 0$, $Bx = 0$, т. е., что $B \subset x \subset \bar{A}$. Иначе говоря, *x* есть некоторый класс, который лежит между *B* и $\bar{A}$, почему его и можно выразить в виде

$$B + \frac{0}{0}\bar{A} \tag{1.10}$$

или в виде

$$\bar{A}B + \frac{0}{0}\bar{A}\bar{B}$$

(из $B = AB + \bar{A}B$, $AB = 0$ вытекает $B = \bar{A}B$; в $B + \frac{0}{0}\bar{A}B + \frac{0}{0}\bar{A}\bar{B}$ член $\frac{0}{0}\bar{A}B$ поглощается членом *B),* т. е. то самое, что получил Буль, но теперь уже без переноса в логику необоснованных арифметических преобразований. При этом Венн обращает внимание на то, что в логике задача решения уравнения, и вообще вывода логических следствий, не имеет смысла, пока не уточнено, какую именно информацию мы хотим извлечь из данной нам в посылках.

В качестве примера Венн рассматривает уравнение Джевонса $Ax + B\bar{x} + C = 1$, которое он переписывает в отрицательной форме (т. е. в форме равенства нулю) $\bar{A}\bar{C}x + \bar{B}\bar{C}\bar{x} = 0$, откуда он получает решение уравнения Джевонса в форме $x = \bar{B}\bar{C} + \frac{0}{0}(AB + C)$ при условии, что $\bar{A}\bar{B}\bar{C} = 0$.

[Это решение в виде (1.10) легко получается из

$$\bar{B}\bar{C} + \frac{0}{0}(A + C),$$

если учесть, что $A + C = AB + A\bar{B} + C = AB +$

$+ A\bar{B}C + A\bar{B}\bar{C} + C = AB + A\bar{B}\bar{C} + C$, и в окончательном результате член $\frac{0}{0} A\bar{B}\bar{C}$ поглощается членом $\bar{B}\bar{C}$.]

Как и Буль, Венн формулирует для алгебры логики не только алгорифмы «решения» уравнений, но и алгорифмы исключения неизвестных. В отличие от Буля, он при этом подчеркивает, что исключение неизвестных в алгебре логики имеет смысл, заведомо отличный от того, который оно имеет в обычной алгебре. **Под исключением неизвестных в логике понимается извлечение максимальной информации, относящейся только к некоторым из терминов, содержащихся в посылках задачи.**

Исключение термина из одной посылки Венн осуществляет посредством замены его словом «некоторые» или символом $\frac{0}{0}$. Это правило он показывает на простом примере

$$w = xy + \bar{x}z; \tag{1.11}$$

исключение из (1.11) термина *y* дает

$$w = \frac{0}{0} x + \bar{x}z.$$

Исключение неизвестной, таким образом, увеличивает неопределенность посылки. Того же результата можно достичь по Венну и заменой в последнем уравнении термина $\frac{0}{0}$ термином *w*, т. е. заменой явного уравнения (1.11) неявным

$$w = wx + \bar{x}z.$$

Хотя Венн иллюстрирует разные способы исключения неизвестных на многочисленных примерах, однако точного определения того, что именно понимается им под «исключением неизвестных», он не дает. Во втором издании «Символической логики» он прямо ссылается на Э. Шредера, по которому результатом исключения *x* из уравнения

$$Ax + B\bar{x} + C = 0 \tag{1.7}$$

является уравнение

$$AB + C = 0,$$

поскольку выражение

$$AB + C$$

«содержит все члены, свободные от *x* или те, которые могут быть освобождены от *x*».

Ясно, однако, что, если трактовать уравнение (1.7) как утверждение

$$\exists x\,(Ax + B\bar{x} + C = 0)$$

[существует такое *x*, что $Ax + B\bar{x} + C = 0$], то для того чтобы это утверждение было верно, необходимо и достаточно, чтобы имело

место $AB = 0$, $C = 0$, т. е. $AB + C = 0$. Действительно, если существует такое x, что

$$Ax + B\bar{x} + C = 0,$$

то независимо от того, какое это x, $C = 0$. Поэтому уравнение (1.7) можно заменить уравнением

$$Ax + B\bar{x} = 0,$$

умножая обе части которого один раз на Ax, другой — на Bx, мы получим

$$AAx\bar{x} + AB\bar{x}\bar{x} = 0,$$
$$ABxx + BBx\bar{x} = 0,$$

откуда

$$AB\bar{x} = 0, \qquad ABx = 0,$$

или, складывая почленно,

$$AB(x + \bar{x}) = 0, \qquad \text{т. е. } AB = 0.$$

Наоборот, если

$$AB + C = 0,$$

то заведомо существует такое x, для которого (1.7) имеет место. Роль такого x может играть, например, B. Мы уже знаем, что в этом случае уравнение (1.7) имеет решение; им является любое x в интервале от B до A. Можно сказать поэтому, что уравнение

$$AB + C = 0$$

является самым сильным логическим следствием уравнения

$$Ax + B\bar{x} + C = 0, \tag{1.7}$$

не содержащим неизвестной x. «Самым сильным» в том смысле, что любое другое следствие $S = 0$ из (1.7), не содержащее неизвестной x, должно быть следствием самого сильного. Действительно, из истинности самого сильного следствия

$$AB + C = 0$$

следует резрешимость уравнения (1.7), т. е. истинность его для какого-то x; из последней же, в свою очередь, следует, что $S = 0$; таким образом $S = 0$ есть логическое следствие равенства

$$AB + C = 0.$$

То обстоятельство, что правила силлогизма представляют собой именно правила вывода самых сильных логических следствий, не содержащих среднего термина, из посылок силлогизма, т. е., что они могут трактоваться как результат исключения среднего термина из посылок силлогизма, специально подчеркивается Венном. Он отмечает при этом, что хотя в заключении силлогизма происходит некоторая потеря информации по сравнению с посылками, однако эта

потеря несущественная, если речь идет о соотношении между субъектом и предикатом, выявляемом только с помощью среднего термина, играющего вспомогательную роль.

В связи с решением уравнений здесь следует заметить, что ни Венн, ни Шредер не рассматривают уравнений вида

$$f(x_1, \ldots, x_n) = 0,$$

где $f(x_1, \ldots, x_n)$ есть терм, тождественно равный 1 или 0.

Однако Венн этого не замечает, Шредер же в определении уравнения специально подчеркивает, что $f(x_1, \ldots, x_n)$ не является ни тождественно равным 0, ни тождественно равным 1.

Вернемся теперь к операции деления. Результат операции деления *a* на *b*, как пишет Венн, есть такое *x*, что $a = bx$. Таким образом, *x* есть корень уравнения $a=bx$, т. е. уравнения в отрицательной форме

$$\bar{a}bx + a\bar{x} + a\bar{b} = 0.$$

Решением последнего является $x = a + v\bar{a}(a + \bar{b})$ при условии, что $a\bar{b} = 0$. После преобразования получаем $x = a + v\bar{a}\bar{b}$ при условии $a\bar{b} = 0$— для сравнения см. выше результат операции деления.

Приведем пример, при решении которого Венн использует обратные операции.

Известно: из некоторого класса *w* предметов удалена их часть, которая состоит из *x*, являющихся *z*, и *y*, являющихся не-*z*; из оставшейся части удалены *z*, которые есть *y*, и *x*, которые есть не *-y;* оставшаяся в результате часть составляет класс *z*, которые не есть *x*. Требуется найти условия, при которых можно выразить класс *w* через остальные.

Символически условие задачи Венн записывает в виде уравнения:

$$w(1 - xz - y\bar{z})(1 - yz - x\bar{y}) = \bar{x}z.$$

Полученное уравнение он преобразует к виду $w\bar{x}\bar{y} = \bar{x}z$,

$w = \frac{\bar{x}z}{\bar{x}\bar{y}}$. Исключаяо перацию деления, он находит

$$w = \bar{x}\bar{y}z + v(1 - xz)(1 - \bar{x}\bar{y}), \quad w = \bar{x}\bar{y}z + v(x + xy\bar{z})$$

при условии $xyz = 0$.

Заметим, что к решению уравнения

$$Ax + B\bar{x} + C = 0 \qquad (1.7)$$

сводится решение логического уравнения с несколькими неизвестными. Например, уравнение

$$Axy + Bx\bar{y} + C\bar{x}y + D\bar{x}\bar{y} = 0$$

(предполагаем, что $Axy + Bx\bar{y} + C\bar{x}y + D\bar{x}\bar{y}$ не есть тождественный нуль или тождественная единица; легко видеть, что к такому виду может быть приведено любое логическое уравнение с

неизвестными *x* и *y),* где *A, B, C* и *D*) не содержат *x, y,* перепишем в виде уравнения с одной неизвестной *x:*

$$(Ay + B\bar{y})\,x + (Cy + D\bar{y})\,\bar{x} = 0.$$

Полученное уравнение "имеет решение относительно *x* тогда и только тогда, когда

$$(Ay + B\bar{y})\,(Cy + D\bar{y}) = 0,$$

т. е., когда

$$ACy + BD\bar{y} = 0. \qquad (1.12)$$

Последнее имеет решение относительно *y* тогда и только тогда, когда *ABCD* = 0. Разрешая уравнение (1.12) относительно *y* и подставляя *y* в исходное уравнение, получим в результате уравнение с одним неизвестным *x,* решать которое мы уже умеем.

Чтобы решить систему логических уравнений, можно все уравнения записать в отрицательной форме и решить уравнение, являющееся дизъюнкцией (суммой) всех уравнений.

Как уже отмечалось, для записи частных предложений Венн во втором издании применил знак «<>». Этот знак он использовал при рассмотрении вопроса об исключении неизвестных в частных суждениях. Опираясь на то, что все, находящееся в *Ax* (соответственно в $B\bar{x}$), заведомо содержится в *A* (соответственно в *B)* и что из

$$A + B + C = 0$$

следует

$$Ax + B\bar{x} + C = 0,$$

Венн показывает, что неравенство

$$Ax + B\bar{x} + C > 0 \qquad (1.13)$$

имеет смысл тогда и только тогда, когда

$$A + B + C > 0, \qquad (1.14)$$

и называет последнее неравенство результатом исключения неизвестной *x* в (1.13)

Особый интерес представляет случай, когда среди посылок задач имеются как общие, так и частные, суждения. Венн ограничивается рассмотрением случая двух посылок:

$$\left.\begin{aligned} Ax + B\bar{x} + C > 0 \\ Dx + E\bar{x} + F = 0 \end{aligned}\right\}. \qquad (1.15)$$

Для этого случая Венн дает общее решение, которое иллюстрирует затем на примерах. Систему (1.15) нужно, очевидно, понимать как означающую

$$\exists x\,(Ax + B\bar{x} + C > 0 \;\&\; Dx + E\bar{x} + F = 0). \qquad (1.16)$$

Из того, что

$$Dx + Ex + F = 0,$$

следует

$$Dx = 0, \qquad Ex = 0, \ F = 0; \tag{1.17}$$

откуда, далее,

$$\frac{DEx = 0, \ DE\bar{x} = 0}{DE(x + \bar{x}) = 0},$$

$$DE \cdot 1 = 0,$$

$$DE = 0. \tag{1.18}$$

Из (1.17) следует: $x \subset \bar{D}$, $x \subset \bar{E}$. Поэтому $Ax \subset A\bar{D}$, $Bx \subset B\bar{E}$. Иесли $Ax + B\bar{x} + C > 0$, то тем более,

$$A\bar{D} + B\bar{E} + C > 0.$$

Таким образом, из (1.15) следует

$$\left.\begin{aligned} DE + F &= 0 \\ A\bar{D} + B\bar{E} + C &> 0 \end{aligned}\right\}. \tag{1.19}$$

И наоборот, из (1.19) следует (1.16), т. е. (1.15). Действительно, из (1.19) имеем $DE = 0$, $F = 0$, и, хотя бы один из $A\bar{D}$, $B\bar{E}$ или C больше 0. Если $A\bar{D} > 0$, то положим $x = \bar{D}$. Для этого x мы получим следующее

$$Ax + B\bar{x} + C = A\bar{D} + BD + C > 0;$$
$$Dx + E\bar{x} + F = D\bar{D} + ED + F = 0.$$

Если $B\bar{E} > 0$, то положим $x = E$. Для этого x будем иметь

$$Ax + B\bar{x} + C = AE + B\bar{E} + C > 0;$$
$$Dx + E\bar{x} + F = DE + E\bar{E} + F = 0,$$

т. е. (1.16) является верным.

Если же $C > 0$, то для любого x $Ax + B\bar{x} + C > 0$, и при $x = E$ $Dx + E\bar{x} + F = DE + E\bar{E} + F = 0$, т. е., опять-таки, (1.16) верно.

Итак, результат исключения, т. е. (1.19), есть «самое сильное следствие»,— уже хотя бы потому, что он просто эквивалентен системе посылок.

Венн замечает далее, что, если в случае универсальных высказываний исключение не всегда возможно (чтобы исключить x из

$$Ax + B\bar{x} = 0,$$

должно быть $A\,B = 0$), то в случае частных из

$$Ax + B\bar{x} + C > 0$$

всегда можно получить

$$A + B + C > 0.$$

Однако в ряде случаев такое исключение иллюзорно, например, если

$$A + B + C \equiv 1.$$

(A+B+C тождественно равно 1), то результат исключения не содержит никакой информации (1>0), т. е. ничего не дает. Впрочем, иллюзорным исключение может быть и для уравнения

$$Ax + B\bar{x} + C = 0,$$

если *AB* тождественно равно нулю (0 = 0). В этом случае Венн также говорит, что исключение невозможно. И он особо отмечает случай, когда $A = \bar{B}$, в котором исключение невозможно ни для

$$Ax + B\bar{x} > 0,$$

ни для

$$Ax + B\bar{x} = 0.$$

Если *A* и *B* представляют собой разложения по конституентам для одинаковых переменных, то, чтобы исключение было «возможно», нужно, — в случае уравнения $Ax + B\bar{x} = 0$, — чтобы *A* и *B* содержали общие члены (одинаковые слагаемые) с коэффициентами 1. Сумма этих слагаемых и есть в таком случае результат исключения *x*.

Так, для уравнения $(z\bar{u}v + \bar{z}u\bar{v} + zu\bar{v})x + (zuv + \bar{z}u\bar{v} + zu\bar{v})\bar{x} = 0$ результатом исключения *x* будет $\bar{z}u\bar{v} + zu\bar{v} = 0$ или $u\bar{v} = 0$.

Это соображение дает Венну возможность предложить удобный графический прием исключения неизвестных, который он иллюстрирует на ряде примеров (некоторые из них приводятся ниже), позволяющих очень наглядно ответить на оба вопроса:

1. Возможно ли исключение?
2. Если возможно, то получить его результат.

В случае неравенства $Ax + B\bar{x} > 0$ дело обстоит так, что в *A* + *B* не должны входить с коэффициентами «1» все члены разложения по конституентам для остальных переменных (кроме *x),* иначе будем иметь *A* + *B*≡1, где ≡ есть знак тождества. Отсюда также получается графический прием для исключения неизвестных, который мы осветим уже после того, как опишем графический метод Венна.

4.1.5. Графический метод Венна

Для простейшего случая двух классов диаграммы Венна фактически уже рассматривались в п. 4.1.3. Венн ввел диаграммы для наглядного представления заданной информации и для решения с их помощью некоторых задач символической логики. Построение диаграмм Венн начинает с разбиения части плоскости на 2^n ячеек с помощью *n* фигур,

где n — число переменных, данных в условии задачи. В дальнейшем предложенный Венном метод разбиения плоскости изменялся и усовершенствовался, делались попытки увеличения наглядности его для большего числа переменных. В настоящее время известно несколько способов деления плоскости на 2^n ячеек с помощью n фигур. Приведем некоторые из них. Начнем с метода Венна, дополняя его при $n = 1,\ldots,4$ усовершенствованиями Мак-Каллока.

Замкнутая кривая Ψ без самопересечений делит плоскость на две части (ячейки) — внутреннюю и внешнюю (предполагаем, что кривая Ψ — граница ячеек — не принадлежит ни одной из них), одну из ячеек (внутреннюю) обозначим a, другую — дополняющую a до плоскости — $\bar{a}$. Иногда в качестве Ψ удобно использовать прямую, которая также делит плоскость на две части (ячейки).

При $n = 1$ в качестве Ψ можно взять окружность произвольного, но фиксированного радиуса (см. первую диаграмму на рис. 1); или прямую (рис. 27).

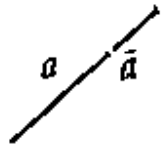


Рис. 27

При $n > 1$ замкнутые кривые $\Psi_1, \ldots, \Psi_n$ без самопересечений располагают на плоскости так, чтобы разделить ее на 2^n ячеек.

При $n = 2$ можно разделить плоскость на четыре ячейки двумя окружностями (см. четвертую диаграмму на рис. 1), или двумя прямыми (рис. 28).

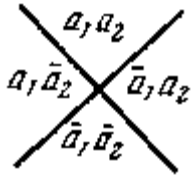


Рис. 28

При $n = 3$ еще можно воспользоваться тремя окружностями (рис. 29) или двумя прямыми и окружностью, как показано на рис. 30, но уже нельзя — тремя прямыми.

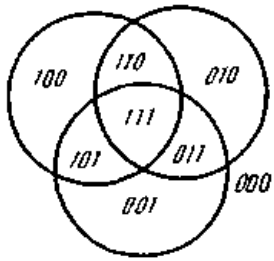


Рис. 29

Рис. 30

При $n = 4$ можно расположить на плоскости две прямые, окружность и эллипс так, что плоскость разделится на 2^4 ячеек (рис. 31); можно также ограничиться, как это делает Венн, четырьмя эллипсами (рис. 32). Следовательно, при $n = 1, 2, 3, 4$ плоскость можно разделить на 2^n ячеек с помощью n фигур, ограниченных кривыми без точек самопересечения.

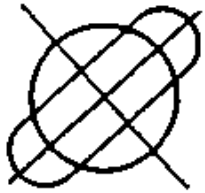


Рис. 31

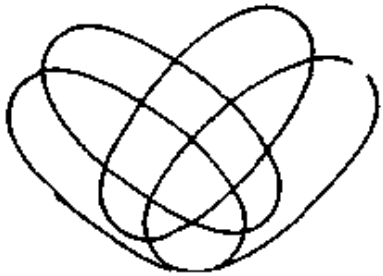


Рис. 32

Перейдем к вопросу об общем способе разбиения плоскости на 2^n ячеек с помощью n фигур.

Венн в своих работах не останавливается подробно на общем способе разбиения плоскости на 2^n ячеек, так как в разбираемых им задачах встречаются только случаи, когда $n \leq 5$, которые он исследует подробно [выше мы привели случаи n = 1, . . ., 4 (рис. 1 — первая и четвертая диаграммы, рис. 29, рис. 32); о случае n = 5 см. в п. 4.1.6]. Однако общий метод Венн представляет себе достаточно ясно. Так, он пишет: «Для чисто теоретических целей правило образования будет очень просто. Нужно начать с того, что нарисовать любую замкнутую фигуру и затем рисовать другие, подчиняя их только условию, чтобы каждая пересекала один и только один раз все существующие подразделения, произведенные теми фигурами, которые были проведены раньше». Сформулируем это правило Венна индуктивно.

1. При n = 1, 2, 3 способы разбиения окружностями указаны выше (первая и четвертая диаграммы, изображенные на рис. 1 и рис. 29).
2. Предположим, что при $n = k$, $k \geq 3$ указано такое расположение k фигур, что плоскость делится на 2^k ячеек. Тогда для расположения k+1 фигуры на этой плоскости достаточно, во-первых, выбрать незамкнутую кривую Жордана φ, имеющую с каждой из границ всех 2^k ячеек один общий кусок; во-вторых, обвести φ замкнутой кривой Жордана Ψ_{k+1} так, чтобы кривая проходила Ψ_{k+1} через все 2^k ячеек и пересекала границу каждой ячейки только два раза.

Таким образом получится расположение $n = k$ +1 фигур такое, что плоскость разделится на 2^{k+1} ячеек.

На рис. 29 в ячейках плоскости расположены последовательности из n (где n = 3) нулей и единиц; единица на i-ом месте последовательности означает принадлежность ячейки фигуре a_i нуль на j-ом месте — принадлежность ячейки дополнению фигуры a_j. Такие последовательности из n нулей и единиц можно воспринимать как числа в двоичной системе; эти числа, равные (в десятичной системе) соответственно $0, 1, \ldots, 2^n - 1$ (при n = 4 имеем 0, 1, . . ., 15), будем рассматривать как номера ячеек.

Другая нумерация ячеек встречается у Шредера; он нумерует их — например, при n = 3 в следующем порядке: abc получает номер 1, $ab\bar{c}$ — номер 2, $a\bar{b}c$ — номер 3, $a\bar{b}\bar{c}$ — номер 4, $\bar{a}bc$ — номер 5, $\bar{a}b\bar{c}$ — номер 6, $\bar{a}\bar{b}c$ — номер 7, $\bar{a}\bar{b}\bar{c}$ — номер 8 . Э. Шредер предлагает для решения логических задач иметь напечатанные штампы (с проставленными на них номерами ячеек) в большом числе и для выражения посылок вычеркивать на них соответствующие номера.

Известны и другие способы разбиения плоскости на 2^n ячеек. Например, метод Минского — Сэлфриджа. Очевидно, что с ростом n

наглядность картинок уменьшается. Для больших n поэтому удобнее пользоваться таблицами, состоящими из 2^n ячеек — таблицами Венна n переменных.

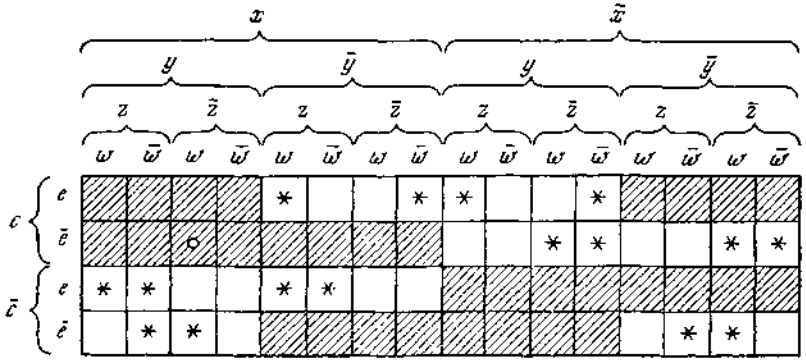


Рис. 33

На рис. 33 приведена таблица Венна шести переменных. Она содержит 64 ячейки. На левой стороне таблицы указаны все возможные комбинации фигур c, e и их дополнений $\bar{c}$, $\bar{e}$, на верхней стороне — все возможные комбинации фигур x, y, z, w и их дополнений $\bar{x}$, $\bar{y}$, $\bar{z}$, $\bar{w}$. Следовательно, для каждой ячейки таблицы можно однозначно определить, принадлежит она некоторой фигуре или ее дополнению; так, ячейка, расположенная на пересечении третьего слева столбца и второй сверху строки (на рис. 33 вэтой ячейке поставлен кружок), принадлежит c, $\bar{e}$, x, y, $\bar{z}$, w.

Между таблицами и картинками Венна легко устанавливается взаимно однозначное соответствие — по таблице можно построить картинку и, наоборот, по картинке восстановить таблицу. Продемонстрируем это на примерах для n = 1, . . ., 5—по таблицам 3—7 строятся диаграммы, приведенные на рис. 34.

Таблицы 3 — 7

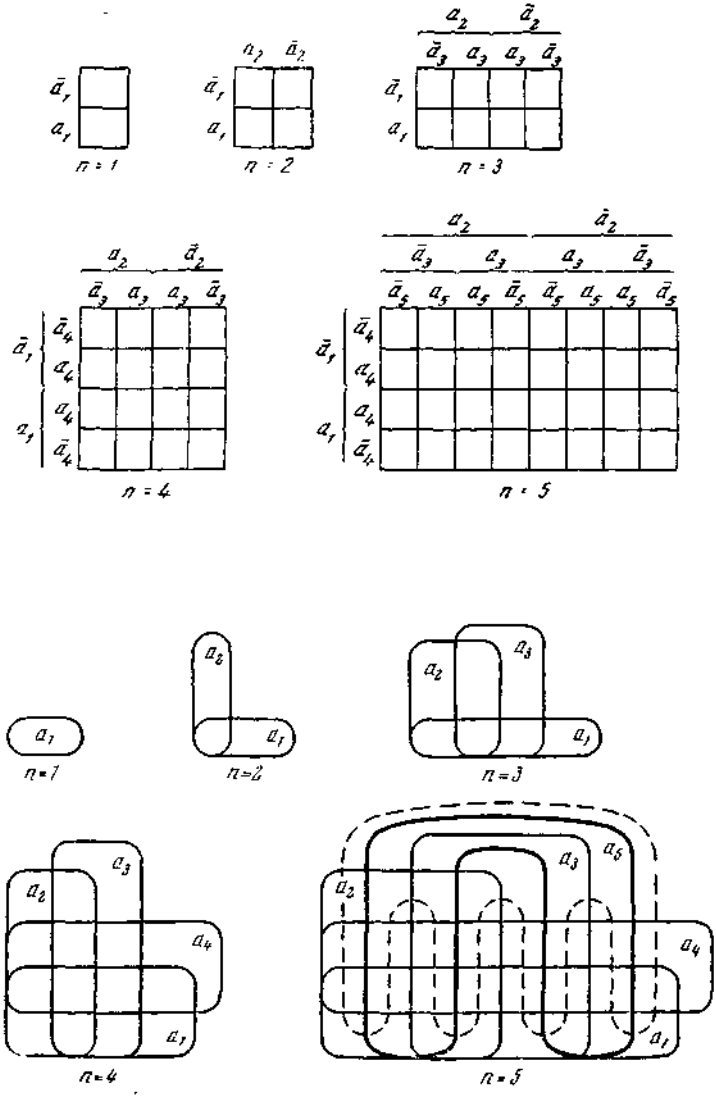


Рис. 34

Дальнейший рост n (переход от n к $n+1$) можно произвести удвоением таблицы Венна *(n* переменных) по горизонтали (на последней

диаграмме рис. 34 пунктиром отмечена граница фигуры, соответствующей переменной $a_в$ и имеющей вид «гребенки»).

Для дальнейшего не существенно, каким образом получены 2^n ячеек — разбиением плоскости или построением таблицы. Поэтому введем понятие символа Венна. Символом Венна n переменных будем называть часть плоскости, состоящую из 2^n ячеек, или таблицу Венна n переменных.

Ячейки символа Венна будем нумеровать числами 0, 1, . . ., 2^n — 1, двоичная запись каждого из которых совпадает с соответствующей последовательностью из n единиц и нулей (единица на i-ом месте последовательности означает принадлежность ячейки фигуре a_i, нуль на j-ом месте — принадлежность ячейки дополнению фигуры a_j.

Методы разбиения плоскости на 2^n ячеек с помощью n фигур и построения таблиц были предложены Венном, когда он занимался проблемой преобразования информации, заданной в виде одного или нескольких предложений. Имея посылки, он старался извлечь из них информацию о том, какие из 2^n ячеек (n — число графически неравных переменных, содержащихся в посылках) пусты, какие не пусты и какие пусты или не пусты в зависимости от того, пусты или не пусты какие-либо другие ячейки. Из числа предшественников этих методов (формулировавших их не графически) Венн отмечает Хр. А. Землера и У. С. Джевонса, которые предлагали перечислять все возможные комбинации терминов классов и их дополнений и вычеркивать те из них, которые пусты в силу условий задачи.

Свой графический метод Венн иллюстрирует на многочисленных примерах, не давая, однако, общего определения понятия «диаграмма». На основании анализа этих примеров попытаемся воспроизвести определение диаграммы по Венну.

Диаграммой n переменных по Венну можно назвать символ Венна n переменных, одни из ячеек которого могут быть заштрихованы, другие могут быть пустыми, а в третьих могут быть поставлены звездочки.

Например, на рис. 35 приведена диаграмма Венна четырех переменных, в которой заштрихованы ячейки

$A\bar{B}\bar{C}\bar{D}$, $AB\bar{C}\bar{D}$, $A\bar{B}C\bar{D}$, $A\bar{B}\bar{C}D$, $\bar{A}BCD$, $\bar{A}\bar{B}CD$; для построения символа Венна использованы эллипсы; диаграмма непосредственно выражает предложение

$$A\bar{B}\bar{C}\bar{D} + AB\bar{C}\bar{D} + A\bar{B}C\bar{D} + A\bar{B}\bar{C}D + \bar{A}BCD + \bar{A}\bar{B}CD = 0,$$

которое свидетельствует о том, что нет таких AB, которые не были бы и C и B, и таких CD, которые не были бы и A и B; ячейка $\bar{A}B\bar{C}\bar{D}$ диаграммы отмечена кружком для сравнения с соответствующей

диаграммой Больцано (рис. 22), на которой этой ячейки нет. Другие примеры диаграмм Венна см. на рисунках 10, 23—34.

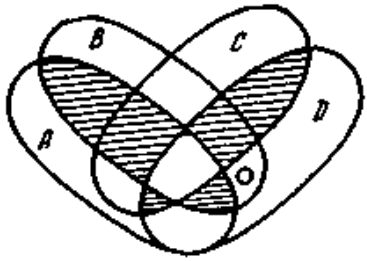


Рис. 35

Следует отметить, что звездочки у Венна встречаются только в одном примере. Диаграммы, не содержащие звездочек, изоморфны формулам исчисления высказываний. Звездочки на диаграммах появляются для выражения частных суждений, не формулируемых на языке формул исчисления высказываний. Однако введение звездочек не позволяет еще использовать диаграммы для выражения произвольных предложений логики классов. В третьей главе предлагается такое изменение диаграмм, которое позволяет выражать в них информацию, задаваемую произвольными формулами исчисления одноместных предикатов.

4.1.6. Некоторые задачи логики классов, их решение с помощью диаграмм Венна

1. Задача Буля. Это — одна из самых сложных задач, имеющихся в книге Буля «Исследование законов мысли, на которых основаны математические теории логики и вероятностей».

Задача гласит:

Представим себе, что некто сказал нам, что наблюдение некоторого класса явлений (естественных или искусственных, например, каких-нибудь веществ) привело к таким общим результатам:

α) Если одновременно отсутствуют признаки *A* и *C,* то обнаруживается признак *E* вместе с одним из признаков *B* или *D*, но не с обоими.

β) Всюду, где встречаются одновременно признаки *A* и *D* при отсутствии *E,* либо обнаруживаются оба признака *B* и *C,* либо оба отсутствуют.

γ*)* Всюду, где имеет место признак *A* вместе с *B* или *E* или вместе с обоими, обнаруживается также один и только один из признаков *C* и *D*. И наоборот, всюду, где наблюдается один и только один из признаков

C и *D,* обнаруживается также признак *A* вместе с *B* или *E* или же с обоими.

Предполагая эту информацию правильной, требуется, во-первых, выяснить, какие заключения в каждом случае можно вывести из наличия признака *A* относительно признаков *B, C* и *D;* во-вторых, решить вопрос о том, нет ли между признаками *B, C, D* каких-нибудь отношений, имеющих между ними место независимо от наличия или отсутствия остальных признаков (и если да, то каких именно?); в-третьих, аналогичным образом ответить на вопрос о том, что следует из наличия признака *B* относительно признаков *A, C* и *D* (равно как и наоборот, когда из наличия или отсутствия признаков этой последней группы можно сделать заключение о наличии или отсутствии признака *B);* в-четвертых, констатировать, что следует для признаков *A, C, D* самих по себе (т. е. независимо от остальных).

Формализация условий этой задачи не представляет трудностей. Вообще, по поводу задач этого раздела (эквивалентного классическому исчислению высказываний) Шредер пишет, например, что они даже не могут быть трудными, поскольку у нас есть общий метод (алгоритм) их решения.

Чтобы выполнить формализацию задачи Буля, обозначим высказывания о наличии признаков *A, B, C, D, E* через *a, b, c, d, e* соответственно; высказывания об их отсутствии — так же, но с чертою $\bar{a}$ означает, таким образом, отсутствие признака *a,* вообще — неверность (отрицание) *a*]*;* союзу «если..., то» поставим в соответствие знак импликации: $\supset$ (т. е. фразу вида «Если *X,* то *Y*» будем записывать как: $(X \supset Y)$**;** союзу «или» (неразделительному, т. е. говорящему только, что хотя бы одно из двух высказываний верно, но не исключающему того, что, может быть, оба верны) — знак «+»; союзу «и»—отсутствие знака (т. е. будем читать *XY* как *X* и *Y).*

На этом языке условия нашей задачи легко выразятся так:

$$\alpha)\ \ (\bar{a}\bar{c} \supset e\,(b+d)\,\bar{b}\bar{d})$$

(Фразу «но не с обоими» мы читаем здесь как «и неверно, что имеют место оба: *b* и *d»).*

$$\beta)\ \ (ad\bar{e} \supset (bc+\bar{b}\bar{c}))$$

(в этом случае различие между неразделительным и разделительным «или» исчезает, так как оба члена вместе невозможны);

$$\gamma)\ \ (a\,(b+e) \supset (c\bar{d}+\bar{c}d))\,((c\bar{d}+\bar{c}d) \supset a\,(b+e))$$

(напомним, что *(b+e)* означает «*b* или *e* или оба вместе» и что для выражения $(c\bar{d}+\bar{c}d)$ неразделительное «или» совпадает с разделительным, так как $c\bar{d}$ не может быть вместе с $\bar{c}d$**).**

Теперь напомним еще, что на языке алгебры логики Венна импликация $(X \supset Y)$ (т. е. «если *X*, то *Y*») выражается как $X\overline{Y} = 0$. (Здесь 0 можно понимать как знак для «лжи», $X\overline{Y} = 0$ читать как: «неверно, будто *X* и не-*Y*»). Равенство же *X* = *Y* эквивалентно совокупности двух импликаций (конъюнкции их, т. е. соединению союзом «и»): $(X \supset Y)$ и $(Y \supset X)$.

Посылки α, β, γ выражаются поэтому на языке Венна равенствами так:

$$\alpha^*) \quad \bar{a}\bar{c}\,\overline{(e\,(b+\bar{d})\,\overline{b\bar{d}})} = 0,$$

$$\beta^*) \quad ad\bar{e}\,\overline{(bc+\bar{b}\bar{c})} = 0,$$

$$\gamma^*) \quad a\,(b+e) = (c\bar{d}+\bar{c}d).$$

Хотя посылка γ (конъюнкция двух импликаций вида $(X \supset Y)$ и $(Y \supset X)$) выражается при этом особенно просто, для получения диаграммы Венна ее лучше трактовать как пару посылок: как две импликации; преобразование каждой из этих импликаций дает:

$$\gamma_1^*) \quad a\,(b+e)\,\overline{(c\bar{d}+\bar{c}d)} = 0,$$

$$\gamma_2^*) \quad (c\bar{d}+\bar{c}d)\,\overline{(a\,(b+c))} = 0.$$

Пользуясь теперь правилами де Моргана для образования противоположности (т. е. заменяя отрицание конъюнкции на дизъюнкцию отрицаний, отрицание же дизъюнкции на конъюнкцию отрицаний), мы выразим наши посылки так:

$$\alpha^0) \quad \bar{a}\bar{c}\,(\bar{e}+\bar{b}\bar{d}+bd) = 0,$$

$$\beta^0) \quad ad\bar{e}\,(\bar{b}+\bar{c})\,(b+c) = 0,$$

$$\gamma_1^0) \quad a\,(b+e)\,(\bar{c}+d)\,(c+\bar{d}) = 0,$$

$$\gamma_2^0) \quad (c\bar{d}+\bar{c}d)\,(\bar{a}+\bar{b}\bar{e}) = 0.$$

Теперь нам остается только воспользоваться законом дистрибутивности (т. е. «открыть скобки»), чтобы получить возможность выразить графически всю информацию, содержащуюся в наших посылках в виде диаграммы Венна. В результате мы получим:

$$\alpha^1) \quad \bar{a}\bar{c}\bar{e} + \bar{a}\bar{b}\bar{c}\bar{d} + \bar{a}\bar{b}\bar{c}d = 0,$$

$$\beta^1) \quad a\bar{b}\,cd\bar{e} + ab\bar{c}d\bar{e} = 0,$$

$$\gamma_1^1) \quad ab\bar{c}\bar{d} + a\bar{c}\bar{d}e + abcd + acde = 0,$$

$$\gamma_2^1) \quad \bar{a}cd + \bar{a}\bar{c}d + \bar{b}cd\bar{e} + \bar{b}\bar{c}d\bar{e} = 0.$$

Это значит, что в диаграмме Венна должны быть пусты (заштрихованы) все ячейки, в названия которых входят приведенные здесь комбинации: $\bar{a}\bar{c}\bar{e}$, $\bar{a}\bar{b}\bar{c}\bar{d}$, $\bar{a}\bar{b}\bar{c}d$, $a\bar{b}cd\bar{e}$, $ab\bar{c}d\bar{e}$, $ab\bar{c}\bar{d}$ и так

далее. Стандартная диаграмма Венна для пяти переменных имеет вид, изображенный на рис. 36 (на штриховку пока не следует обращать внимания).

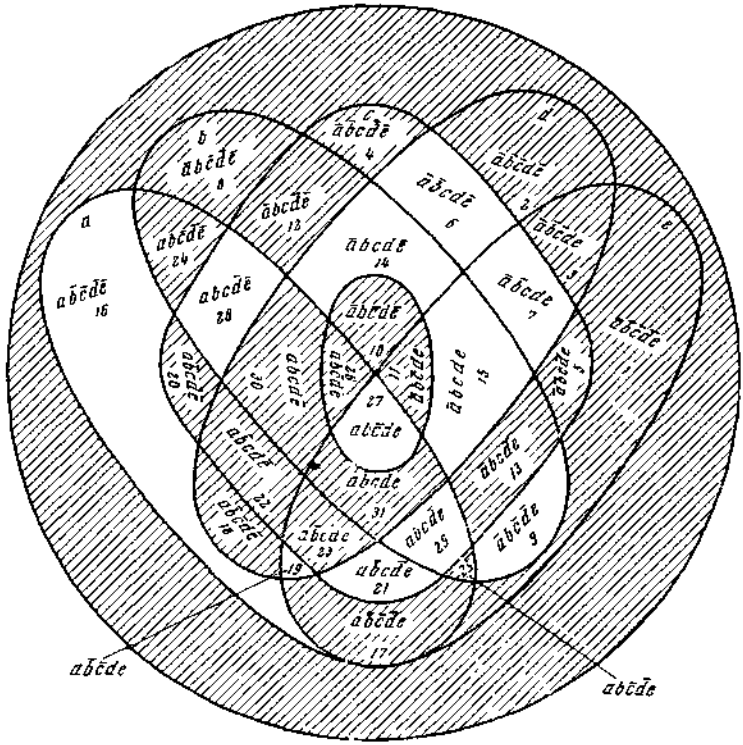


Рис. 36

Здесь четыре овала *a, b, d,* ей фигура *c*, имеющая вид кольца, внутри которого расположены ячейки $abc\bar{d}e$, $ab\bar{c}de$, $\bar{a}b\bar{c}d\bar{e}$, $\bar{a}b\bar{c}de$. Всего ячеек $2^5 = 32$. Они все выписаны на рисунке, где занумерованы так, что если записать номер в виде пятизначного числа в двоичной нумерации, сопоставить его первому знаку слева переменную *a* или $\bar{a}$ в зависимости от того, стоит ли на первом месте 1 или 0, второму знаку, аналогично, переменную *b* или $\bar{b}$, и так далее, то переводом номера будет имя ячейки. Так, число 9 в двоичной нумерации будет в этом случае иметь вид 01001; это значит, что ему соответствует ячейка $\bar{a}b\bar{c}\bar{d}e$; числу 10 соответствует ячейка 01010, или $\bar{a}b\bar{c}d\bar{e}$; наружная ячейка $\bar{a}\bar{b}\bar{c}\bar{d}\bar{e}$ будет иметь номер 0, ячейка $\bar{a}\bar{b}\bar{c}\bar{d}e$ номер 1, ячейка *abcde* — номер 11111, т. е. 31.

Если мы теперь заштрихуем все пустые в силу данной нам информации ячейки (ничего, кроме некоторого внимания, для выполнения этой операции не требуется), то вся предварительная обработка этой информации будет закончена, и ответ на поставленные выше вопросы можно будет искать на полученной картинке.

Начнем с того, что попытаемся ответить на последний из поставленных вопросов: что можно сказать о признаках *A, C, D* самих по себе? — Ответить на этот вопрос с помощью нашей картинки совсем не трудно, хотя и довольно канительно. Действительно, с этой целью достаточно выписать сначала все восемь различных комбинаций ячеек *a, c* и *d* и их дополнений, т. е. все выражения вида $\tilde{a}\tilde{c}\tilde{d}$, где волна над буквой означает, что над нею может иметься (или не иметься) черта. А затем к каждой из этих комбинаций добавить все четыре аналогичные комбинации по буквам *b, e* и посмотреть по нашей диаграмме, какие из восьми различных четверок заштрихованы в ней полностью. Эти четверки дадут нам требуемый ответ, как нетрудно проверить по следующей таблице (табл. 8), сличая ее строки с нашей картинкой; в результате мы получим тризаштрихованные полностью четверки в строках

$$acd,\ \bar{a}c\bar{d},\quad \bar{a}\bar{c}d.$$

Если номер в таблице помещен в квадрате, значит, соответствующая ячейка на диаграмме не заштрихована; значит, соответствующей строкой можно больше не заниматься.

Таблица 8

*	acd	$abcde$	31	$abcd\bar{e}$	30	$a\bar{b}cde$	23	$a\bar{b}cd\bar{e}$	22
	$ac\bar{d}$	$abc\bar{d}e$	[29]						
	$a\bar{c}d$	$ab\bar{c}de$	[27]						
	$a\bar{c}\bar{d}$	$ab\bar{c}\bar{d}e$	25	$ab\bar{c}d\bar{e}$	24	$a\bar{b}\bar{c}de$	17	$a\bar{b}cde$	[16]
	$\bar{a}cd$	$\bar{a}bcde$	[15]						
*	$\bar{a}c\bar{d}$	$\bar{a}bc\bar{d}e$	13	$\bar{a}bc\bar{d}\bar{e}$	12	$\bar{a}\bar{b}c\bar{d}e$	5	$\bar{a}\bar{b}c\bar{d}\bar{e}$	4
*	$\bar{a}\bar{c}d$	$\bar{a}b\bar{c}de$	11	$\bar{a}b\bar{c}d\bar{e}$	10	$\bar{a}\bar{b}\bar{c}de$	3	$\bar{a}\bar{b}\bar{c}d\bar{e}$	2
	$\bar{a}\bar{c}\bar{d}$	$\bar{a}b\bar{c}\bar{d}e$	[9]						

Как видно из таблицы, во всех строках, кроме отмеченных звездочкой, имеются не вычеркнутые на диаграмме ячейки. Таким образом, все, что мы можем сказать о признаках *A,C,D* независимо от остальных признаков, состоит в том, что

$$acd + \bar{a}c\bar{d} + \bar{a}\bar{c}d = 0;$$

или, иначе говоря, что все три признака *A, C, D* вместе не могут встретиться, но достаточно присутствовать признаку *C* (или *D*), чтобы один из двух остальных обязательно имел место.

Чтобы ответить на второй вопрос, можно поступить аналогично. Но только теперь, составив соответствующую таблицу (см. табл. 9), мы увидим, сличив ее с нашей диаграммой, что ни одной строки, отмеченной звездочкой, в ней нет, т. е. что на этот вопрос нужно дать отрицательный ответ: отношений, о которых в нем идет речь, нет.

Таблица 9

bcb	$abcde$	31	$abcd\bar{e}$	30	$\bar{a}bcde$	[15]
$bc\bar{d}$	$abc\bar{d}e$	[29]				
$b\bar{c}d$	$ab\bar{c}de$	[27]				
$b\bar{c}\bar{d}$	$abc\bar{d}\bar{e}$	25	$ab\bar{c}d\bar{e}$	24	$\bar{a}b\bar{c}de$	[9]
$\bar{b}cd$	$a\bar{b}cde$	23	$a\bar{b}cd\bar{e}$	22	$\bar{a}\bar{b}cde$	[7]
$\bar{b}c\bar{d}$	$a\bar{b}c\bar{d}e$	[21]				
$\bar{b}\bar{c}d$	$a\bar{b}\bar{c}de$	[19]				
$\bar{b}\bar{c}\bar{d}$	$a\bar{b}\bar{c}\bar{d}e$	17	$a\bar{b}\bar{c}\bar{d}\bar{e}$	[16]		

Для ответа на первый вопрос к числу посылок надо добавить еще *a* или классически эквивалентную ей информацию, состоящую в том, что $\bar{a} = 0$, т. е., что все ячейки, не входящие в *a*, пусты. Ясно, что при наличии признака *A* все различие возможностей для признаков *B, C* и *D* надо искать в пределах ячейки *a* нашей диаграммы. А в таком случае для ответа на первый вопрос нам придется только выяснить, что можно

сказать относительно $\tilde{b}, \tilde{c}, \tilde{d}$ (при наличии *a*) независимо от *e,* т. е., в применении к каким из $\tilde{b}\tilde{c}\tilde{d}$ пусты обе ячейки $\tilde{b}\tilde{c}\tilde{d}e$ и $\tilde{b}\tilde{c}\tilde{d}\bar{e}$ (из находящихся внутри *a).* Ответ на этот вопрос дает нам табл. 10.

Таблица 10

*	bcd	$bcde$	31	$bcd\bar{e}$	30
	$bc\bar{d}$	$bc\bar{d}e$	29		
	$b\bar{c}d$	$b\bar{c}de$	27		
*	$b\bar{c}\bar{d}$	$b\bar{c}\bar{d}e$	25	$b\bar{c}\bar{d}\bar{e}$	24
*	$\bar{b}cd$	$\bar{b}cde$	23	$\bar{b}cd\bar{e}$	22
	$\bar{b}c\bar{d}$	$\bar{b}c\bar{d}e$	21		
	$\bar{b}\bar{c}d$	$\bar{b}\bar{c}de$	19		
	$\bar{b}\bar{c}\bar{d}$	$\bar{b}\bar{c}\bar{d}e$	17	$\bar{b}\bar{c}\bar{d}\bar{e}$	16

Мы видим из нее, что пусты ячейки bcd, $b\bar{c}\bar{d}$ и $\bar{b}cd$, т. е., что при наличии признака *A* хотя бы один из двух признаков *C* и *D* должен отсутствовать: если же, кроме *A*, имеется в наличии и *B,* то оба — и *C* и *D* — отсутствовать не могут одновременно.

Действительно, из

$$bcd + \bar{b}cd = 0$$

следует $cd = 0$, т. е. *C* и *D* несовместны друг с другом; из того же, что $b\bar{c}\bar{d} = 0$ вытекает, что, при наличии *B,* $\bar{c}\bar{d} = 0$, т. е. (классически) имеет место *C* или *D.* На первый взгляд несколько труднее дать ответ па третий вопрос, особенно на вторую его (помещенную в скобки) часть. В действительности, ответить на него не более трудно, чем на остальные вопросы. Ответ дает табл. 11, из которой видно, что внутри *b* (при наличии признака *B,* иначе говоря) пусты ячейки $a\bar{c}\bar{d}$, $ac\bar{d}$, $\bar{a}c\bar{d}$, $\bar{a}\bar{c}d$, т. е. все три признака *A, C* и *D* не могут одновременно присутствовать, но, если какие-нибудь два из них отсутствуют, то отсутствует и третий.

Ответ на вторую часть вопроса также дают отмеченные звездочками строки табл. 11, но только теиерь уже обеих ее частей: и верхней, и нижней.

Таблица 11

*	$abcd$	$abcde$	31	$abcd\bar{e}$	30
	$abc\bar{d}$	$abc\bar{d}e$	[29]		
	$ab\bar{c}d$	$ab\bar{c}de$	[27]		
*	$ab\bar{c}\bar{d}$	$ab\bar{c}\bar{d}e$	25	$ab\bar{c}\bar{d}\bar{e}$	24
	$\bar{a}bcd$	$\bar{a}bcde$	[15]		
*	$\bar{a}bc\bar{d}$	$\bar{a}bc\bar{d}e$	13	$\bar{a}bc\bar{d}\bar{e}$	12
*	$\bar{a}\bar{b}cd$	$\bar{a}\bar{b}cde$	11	$\bar{a}\bar{b}cd\bar{e}$	10
	$\bar{a}\bar{b}c\bar{d}$	$\bar{a}\bar{b}c\bar{d}e$	[9]		
*	$a\bar{b}cd$	$a\bar{b}cde$	23	$a\bar{b}cd\bar{e}$	22
	$a\bar{b}c\bar{d}$	$a\bar{b}c\bar{d}e$	[21]		
	$a\bar{b}\bar{c}d$	$a\bar{b}\bar{c}de$	[19]		
	$a\bar{b}\bar{c}\bar{d}$	$a\bar{b}\bar{c}\bar{d}e$	17	$a\bar{b}\bar{c}\bar{d}\bar{e}$	[16]
	$\bar{a}\bar{b}cd$	$\bar{a}\bar{b}cde$	[7]		
*	$\bar{a}\bar{b}c\bar{d}$	$\bar{a}\bar{b}c\bar{d}e$	5	$\bar{a}\bar{b}c\bar{d}\bar{e}$	4
*	$\bar{a}\bar{b}\bar{c}d$	$\bar{a}\bar{b}\bar{c}de$	3	$\bar{a}\bar{b}\bar{c}d\bar{e}$	2
*	$\bar{a}\bar{b}\bar{c}\bar{d}$	$\bar{a}\bar{b}\bar{c}\bar{d}e$	1	$\bar{a}\bar{b}\bar{c}\bar{d}\bar{e}$	0

Из восьми возможных различных случаев, могущих иметь место по отношению к признакам *A, C, D,* т. е. случаи первый, шестой и седьмой невозможны (что отмечено нами знаком «—» слева), так как из первой и девятой строк таблицы следует, что $abcd + a\bar{b}cd = 0$, т. е., что $acd = 0$. Аналогично из 6-ой и 14-ой строк и 7-ой и 15-ой

Таблица 12

—	1.	$a = c = d = 1,$
0	2.	$a = c = \bar{d} = 1,$
0	3.	$a = \bar{c} = d = 1,$
+	4.	$a = \bar{c} = \bar{d} = 1,$
0	5.	$\bar{a} = c = d = 1,$
—	6.	$\bar{a} = c = \bar{d} = 1,$
—	7.	$\bar{a} = \bar{c} = d = 1,$
+	8.	$\bar{a} = \bar{c} = \bar{d} = 1,$

строк получается, что $\bar{a}c\bar{d} = 0$ и $\bar{a}\bar{c}d = 0$; в случаях 2-ом, 3-ем и 5-ом таблица не дает нам никакой информации (это отмечено знаком «0» слева), так как в соответствующих строках таблицы нет звездочки; в 4-ом случае, т. е. при наличии *A* и отсутствии *C* и *D,* признак *B* отсутствует, т. к. $a\bar{b}\bar{c}\bar{d} = 0$; в последнем (8-ом) случае, т. е. при отсутствии всех трех признаков *A, C* и *D,* признак *B* имеет место, т. е. $\bar{a}b\bar{c}\bar{d} = 0$ (случай 4-й, 8-й отмечены нами знаком «+» слева от таблицы). Таким образом, только в случаях 4 и 8 мы можем с уверенностью сказать, присутствует или отсутствует признак *B,* т. е. и на этот вопрос мы получаем полный и однозначный ответ, перечисляющий все случаи, когда из информации о наличии (или отсутствии) признаков *A, C* и *D* мы имеем возможность с уверенностью делать заключение о том, имеется или отсутствует признак *B.*

В связи с задачей Буля Шредер заметил, что, хотя формулировка задачи и не содержит логического противоречия, однако в ней имеются

некоторые неувязки, изобличающие задающего вопросы в том, что он говорит неправду о каких-то проведенных им «наблюдениях», поскольку в условиях задачи антецедент в посылке β ложен: случай, когда встречаются одновременно признаки *A* и *D* при отсутствии признака *E,* наблюдать фактически невозможно. В диаграмме Венна это замечание Шредера также иллюстрируется наглядно: на ней непосредственно видно, что ячейка *ade,* состоящая из ячеек с номерами 18, 22, 30, 26, пуста. Венн делает отсюда вывод, что задача Буля вообще плохо сформулирована: ее формулировку можно упростить, заменив посылки α, β посылками:

α′) Если одновременно отсутствуют признаки *A* и *C,* то обнаруживается признак *B* вместе с *E.*

β′) Всюду, где встречаются одновременно признаки *A* и *D,* встречается и признак *E.*

Третью посылку менять не нужно.

Нетрудно убедиться, что этим посылкам соответствует в точности та же диаграмма, которая изображена на рис. 36, т. е., что они эквивалентны посылкам Буля, хотя значительно проще их по формулировке и не содержат уже каких-либо неувязок. Уже история одной только задачи Буля, таким образом, достаточно поучительна.

В связи с этой историей интересно также, что Венн не ограничивался каким-нибудь одним из способов решения: графическим или аналитическим, а применял оба, проверяя, таким образом, один метод другим. Он замечает при этом, что бывают случаи, когда графический метод оказывается более удобным.

В случае задачи Буля ответ на четвертый вопрос представляет собою результат исключения *b* и *e* из посылок задачи, который мы теперь получим по правилам, изложенным в п.1.4.1.2, 1.4.1.4.

С этой целью мы, прежде всего, заменим наши посылки α^1, β^1, γ_1^1, γ_2^1 эквивалентным утверждением вида:

$$Xbe + Yb\bar{e} + Z\bar{b}e + U\bar{b}\bar{e} = 0,$$

а затем получим результат исключения в виде

$$XYZU = 0.$$

Эта процедура даст нам — после выполнения некоторых поглощений — для *X, Y, Z, U* соответственно выражения:

$$acd + \bar{a}\bar{c}d + \bar{a}c\bar{d} + a\bar{c}\bar{d},$$

$$ac + ad + \bar{a}\bar{c} + \bar{a}\bar{d},$$

$$acd + \bar{a}\bar{c} + \bar{a}\bar{d} + \bar{c}\bar{d},$$

$$acd + \bar{a}\bar{c} + c\bar{d} + \bar{c}d,$$

«перемножив» которые, мы получим

$$acd + \bar{a}\bar{c}d + \bar{a}c\bar{d},$$

т. е. тот самый результат, который дала нам таблица 8.

На первый взгляд, этот алгоритм представляется значительно более легким, чем составление таблицы, но, в действительности, они требуют примерно одних и тех же «затрат».

Совершенно аналогично производится исключение *a* и *e* из наших посылок для ответа на второй вопрос задачи. Теперь коэффициентами при ae, $a\bar{e}$, $\bar{a}e$, $\bar{a}\bar{e}$ будут соответственно

$$\bar{c}\bar{d} + cd,$$

$$\bar{b}c + b\bar{c} + cd + \bar{b}d,$$

$$\bar{c}d + \bar{b}\bar{d} + c\bar{d},$$

$$\bar{c} + \bar{d}.$$

«Перемножая» эти выражения, мы получим ноль, т. е. результат исключения имеет вид: 0 = 0. Но это и означает в алгебре логики (см. п.1.4.1.4), что исключение невозможно. Иными словами, мы опять получили аналитически тот же результат, который дала нам таблица 9.

Для ответа на первый вопрос нужно произвести исключение *e* после того, как мы добавим к посылкам еще одну посылку: *a,* или $\bar{a} = 0$. В качестве коэффициентов при *e* и $\bar{e}$ мы получим при этом (выполнив возможные поглощения):

$$a\bar{c}\bar{d} + acd,$$

$$acd + ab\bar{c} + \bar{b}c\bar{d} + \bar{b}\bar{c}d.$$

«Перемножая» эти выражения, будем иметь

$$ab\bar{c}\bar{d} + acd,$$

т. е. при наличии *A* получим

$$b\bar{c}\bar{d} + cd = 0.$$

Именно этот результат дала нам таблица 10.

Аналогично можно получить ответ на первую часть третьего вопроса (мы предоставляем читателям убедиться в том, что при этом получится тот же ответ, который дала нам верхняя часть таблицы 11). Для того чтобы получить ответ на вторую часть этого вопроса, проще всего поступить так: добавить один раз к посылкам посылку b $(\bar{b} = 0)$, а

другой раз — $\bar{b}$ $(b = 0)$ и исключить и в том и в другом случае только *e.* Результат исключения и здесь даст нам то же самое, что дала таблица 11.

В качестве примеров того, что бывают случаи, когда графический метод диаграмм Венна быстрее ведет к цели, чем аналитический, Шредер в своей «Алгебре логики» приводит следующие две задачи.

2. Задача Джевонса. Требуется упростить посылки:

$$a = b + c, \quad b = \bar{d} + \bar{c}, \quad \bar{c}\bar{d} = 0, \quad ad = bcd.$$

Если $a = b + c$, то все те части в *b* и *c*, которые не входят в *a*, должны быть вычеркнуты. Должны быть вычеркнуты также те части *a*, которые не входят ни в *b,* ни в *c.* Так как $b = \bar{c} + \bar{d}$, то из *b* нужно вычеркнуть то, что входит в *c* и *d* одновременно, а из $\bar{c}$ и из $\bar{d}$ все, что не входит в *b.* В силу $\bar{c}\bar{d} = 0$, нужно зачеркнуть вообще всю общую часть $\bar{c}$ и $\bar{d}$. Наконец, так как *ad = bed,* то значит, *ad* входит и в *b* и в *c*, т. е. ячейки $a\bar{b}d$ и $a\bar{c}d$ — пустые. Наша диаграмма имеет, следовательно, вид, изображенный на рис. 37, где может быть не пустой только ячейка $abc\bar{d}$.

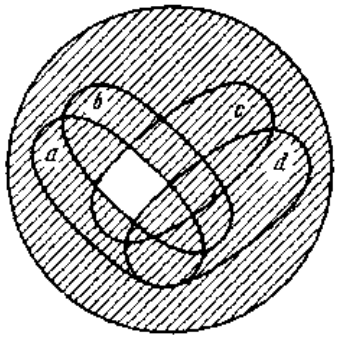


Рис. 37

Но так как на диаграмме изображен универсум (единица), то мы можем сказать, что этой ячейкой исчерпывается универсум, т. е., что пересечение фигур *a, b, c,* $\bar{d}$ равно единице. А это возможно только в том случае, когда все они равны единице, т. е. когда $a = b = c = 1$, $d = 0$. Так как вся информация, заключенная в посылках задачи Джевонса, учтена в диаграмме на рис. 37, которой непосредственно соответствуют равенства $a = b = c = 1$, $d = 0$, то эти равенства и дают ответ на задачу. Этот же ответ можно получить, конечно, и аналитически; в том числе и по правилам алгебры логики Шредера и Венна, но только более громоздко.

Действительно, преобразуя посылки, мы получим:

$$a(\overline{b+c}) + \bar{a}(b+c) + bcd + \bar{b}(\bar{c}+\bar{d}) + \bar{c}\bar{d} +$$
$$+ ad(\bar{b}+\bar{c}+\bar{d}) + (\bar{a}+\bar{d})bcd = 0.$$

Левую часть этого равенства остается привести к сокращенной дизъюнктивной нормальной форме,— что дает $\bar{a} + \bar{b} + \bar{c} + d = 0$, т. е. уже полученный нами с помощью диаграммы ответ. Это приведение требует, однако, довольно большого числа выкладок (хотя и очень простых). [«Открыв» скобки, его можно осуществить, в частности, применив в общей сложности восемь раз закон «выявления»: $AB + \bar{A}C = AB + \bar{A}C + BC$ и 15 раз закон «поглощения»: $A + AB = A$.] К этому же результату можно прийти, конечно, и приведя просто дизъюнкцию посылок к совершенной дизъюнктивной нормальной форме, в которой не будет заведомо равен нулю только член $abc\bar{d}$.

Диаграмматическое решение задачи Джевонса Шредер заимствовал у Венна. В связи с этой задачей Джевонса Шредер приводит и критические замечания Венна в адрес Джевонса, выражая при этом полное согласие с Венном. Поскольку в этих замечаниях речь идет о вопросе, много занимавшем логиков, начиная еще с Аристотеля (который считал недопустимыми вообще в науке рассуждения с пустыми предикатами), мы здесь немного остановимся на этих замечаниях Венна. В них речь идет о том, что, вопреки Джевонсу, обнаружение пустоты некоторого «простого» класса, обозначаемого одной буквой, или его дополнения, само по себе еще никак не может рассматриваться как свидетельствующее о противоречивости посылок; что оно становится таковым только в случае, когда мы молча предполагаем, что «простой» класс не может быть пуст. Почему, однако? Ведь и Джевонс не возражает против того, что пересечение двух простых классов *a, b* (т. е. класс *ab)* может быть пусто. Но чем отличаются классы *a* и *b* от класса *ab*? Чем отличается «простой» класс от пересечения двух или большего числа классов? И Венн говорит, что введение таких ограничений на «простые» классы, т. е. на классы, обозначаемые одной буквой или буквой с чертою, является «самоубийством» для символической логики.

То обстоятельство, что Джевонс считает всякое высказывание, из которого следует пустота какого-нибудь «простого» класса или его дополнения, противоречивым, Шредер квалифицирует как «фундаментальную ошибку» Джевонса и сочувственно цитирует в этой связи слова Венна о «самоубийственности» такого подхода к символической логике.

Теперь мы знаем, что формализовать можно и аристотелеву логику, определив достаточно точно «простые» (элементарные) предикаты, на

которые наложено требование непустоты. Если, однако, понятие «простого» класса не уточняется, то с критикой Венна можно согласиться.

3. **Задача из статьи Венна «О диаграмма-мическом и механическом представлении предложений и рассуждений».**

Требуется упростить посылки:

I. $y \subset x\bar{z} + \bar{x}z$,

II. $wy \subset xz + \bar{x}\bar{z}$,

III. $xy \subset w + z$,

IV. $yz \subset x + w$

($\subset$ — знак включения одного класса в другой). «Класс A включен в класс 5», $A \subset B$,— это значит, что класс, представляющий пересечение классов A и B, пуст, $A\bar{B} = \bar{0}$. На диаграмме включение класса A в класс B обозначается штриховкой (т. е. указателем пустоты) ячейки, соответствующей $A\bar{B}$ (рис. 23).

Возьмем в качестве A класс y, в качестве B класс $x\bar{z} + \bar{x}z$. Тогда на диаграмме (рис. 38) в фигуре, соответствующей классу y, пусты все ячейки, не являющиеся подъячейками фигуры, изображающей класс $x\bar{z} + \bar{x}z$ (на рис. 38 в заштрихованных ячейках поставлен номер посылки — I).

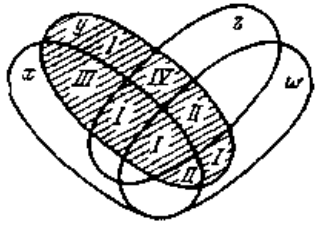


Рис. 38

Аналогично поступаем с остальными посылками.

В результате получаем диаграмму (рис. 38), на которой фигура, обозначающая класс y, полностью заштрихована. Следовательно, посылки I—IV эквивалентны предложению: $y = 0$.

4. Задача об уставе клуба. Этой задачей Венн начинает главу, посвященную разнообразным примерам, с помощью которых он хочет разъяснить начинающим трудные для них абстрактные принципы символической логики, и поэтому подробно ее разбирает.

В задаче требуется обсудить с точки зрения их непротиворечивости и простоты следующие правила клубного устава:

1. Финансовый комитет должен избираться из состава общего комитета.

2. Никто из членов библиотечного комитета не может быть в финансовом комитете.
3. Никто не может быть членом общего и библиотечного комитетов одновременно, если он не входит также и в финансовый комитет.
Венн вводит переменные: x — члены финансового, y — библиотечного и z — общего комитетов, и записывает правила в виде:

(1) Все x суть z.
(2) Все yz суть x.
(3) Никакие x не суть y.

В символической форме он получает:

$$(1)\ x\bar{z} = 0,$$
$$(2)\ \bar{x}yz = 0,$$
$$(3)\ xy = 0.$$

Далее он замечает,— уже в применении к любым посылкам,— что, если какая-нибудь посылка непосредственно противоречит какой-нибудь другой или же повторяет ее, то противоречивость или избыточность такой системы нетрудно заметить. Но бывает так, что нужно еще проанализировать посылки, чтобы это обнаружить. Этот анализ, несколько подробнее разобранный нами, Венн осуществляет так:
После того как посылки приведены к нулевой форме, он «умножает» каждое слагаемое этой формы на все выражения вида $(a + \bar{a})$, где a — переменная, не входящая в это слагаемое, и затем «открывает скобки». Если при этом оказывается, что в посылках имеются все вообще возможные члены универсума, т. е., что «складывая» посылки, мы получим 1=0, то это свидетельствует об их противоречивости. Наоборот, если хотя бы один член отсутствует, то этого уже достаточно для того, чтобы сказать, что система непротиворечива (так как, сделав отсутствующий член равным единице, мы сделаем все остальные равными нулю). Если в разных посылках обнаруживаются одинаковые члены, то мы приходим к заключению об избыточности посылок.
В нашей задаче такой анализ посылок даст нам:

$$xy\bar{z} + x\bar{y}\bar{z} = 0,$$
$$\bar{x}yz = 0,$$
$$xyz + xy\bar{z} = 0.$$

И мы видим, что из восьми различных возможных членов в посылках имеются только четыре, т. е. система непротиворечива. В то же время в первой и третьей посылках есть общий член $xy\bar{z}$, т. е. система избыточна. Этот член, например, из последней посылки, можно

устранить, т. к. в силу первой он равен нулю, и мы получим эквивалентную первоначальной систему посылок:

$$xyz + x\bar{y}\bar{z} = 0,$$
$$\bar{x}yz = 0.$$
$$xyz = 0.$$

Объединяя две последние и используя то, что

$$\bar{x}yz + xyz = yz,$$

Венн получает (возвращаясь к первоначальной форме первой посылки) для правил устава упрощенную форму

$$x\bar{z} = 0,$$
$$yz = 0.$$

Иными словами, ответ Венна гласит: Правила устава непротиворечивы, но их можно упростить,— вместо трех написать два:

1. Финансовый комитет должен избираться из состава общего комитета.
2. Никто из членов библиотечного комитета не может быть в общем комитете.

Метод анализа посылок, использованный здесь Венном, очень удобен и для решения задачи с помощью его диаграмм. На соответствующей диаграмме (рис. 39) видно, что заштрихованная область образует класс $x\bar{z} + yz$, т. е. в символической форме: $x\bar{z} + yz = 0$ или $x\bar{z} = 0$, $yz = 0$, что совпадает с полученным выше ответом.

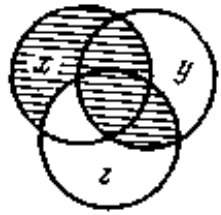


Рис. 39

Эту задачу,— особенно, если сформулировать ее как требование упростить систему посылок,— можно решить, конечно, и проще. Так, Шредер приводит следующее решение: он записывает посылки в виде одного уравнения

$$x\bar{z} + xyz + xy = 0$$

и, применив законы выявления и поглощения, получает сначала

$$x\bar{z} + yz + xy = 0,$$

а потом

$$x\bar{z} + yz = 0$$

(член xy получается с помощью «выявления»). Однако вопрос о критерии избыточности посылок при этом остается неуточненным.

5. Задача о провалившихся на экзамене. Это — задача Венна.

Условия задачи гласят:

Множество провалившихся на некотором экзамене оказалось в точности состоящим из мальчиков, сдававших греческий язык, и девочек, сдававших латынь. Требуется дать полное описание всех сдававших мальчиков в остальных терминах, употребляемых в условии.

Венн обозначает: x — класс сдававших латынь, z — класс сдававших греческий, w — класс мальчиков, $\bar{w}$ — класс девочек (за универсум принят класс всех сдававших), y — класс провалившихся.

Условие задачи состоит при этом в том, что $y = x\bar{w} + zw$, и требуется определить из него w. В нулевой форме это условие дает

$$x\bar{y}\bar{w} + \bar{y}zw + \bar{x}y\bar{w} + y\bar{z}w = 0,$$

или

$$(\bar{y}z + y\bar{z})w + (x\bar{y} + \bar{x}y)\bar{w} = 0.$$

Решение этой задачи по Шредеру дает в качестве необходимого и достаточного условия ее разрешимости

$$(\bar{y}z + y\bar{z})(x\bar{y} + \bar{x}y) = 0,$$

или

$$x\bar{y}z + \bar{x}y\bar{z} = 0,$$

а в качестве решения

$$w = (x\bar{y} + \bar{x}y)\bar{u} + (yz + \bar{y}\bar{z})u,$$

или в другой форме

$$w = x\bar{y}\bar{z} + \bar{x}yz + u(xyz + \bar{x}\bar{y}\bar{z}).$$

Диаграмма же Венна для этого условия имеет вид, показанный на рис. 40, т. е. дает

$$w = x\bar{y}\bar{z} + \bar{x}yz + uxyz + u\bar{x}\bar{y}\bar{z}.$$

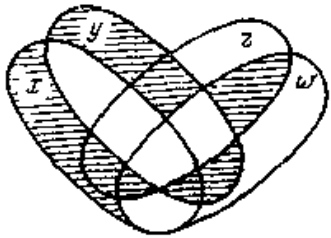


Рис. 40

В связи с этим Венн ставит вопрос о том, как следует интерпретировать на диаграмме неопределенный знак *и* (у самого Венна $\frac{0}{0}$), фигурирующий в решении задачи. В ответ на этот вопрос Венн пишет: «Уделим некоторое внимание диаграмматическому истолкованию символа $\frac{0}{0}$. Когда мы смотрим на четыре сохраняющихся класса, образующих фигуру *w,* то можем подумать, что все они находятся на одном и том же уровне достоверности и что в них нет ничего соответствующего строгому различению, представленному символически через

$$w = x\bar{y}\bar{z} + \bar{x}yz + \frac{0}{0}xyz + \frac{0}{0}\bar{x}\bar{y}\bar{z}.$$

Но размышления скоро выявляют это различие. Два первых класса могут быть полностью описаны в терминах *x, y* и z. Два последних не могут быть так описаны без употребления *w.* Но как так *w* и есть как раз тот самый термин, который требуется определить, то включение его в определение эквивалентно отсутствию определения вообще. Мы можем сделать вид, что даем определение, написав

$$w = x\bar{y}\bar{z} + \bar{x}yz + wxyz + w\bar{x}\bar{y}\bar{z},$$

но такой способ избежать неопределенности во всяком случае обманчив. Этот класс *w* поэтому содержит целые классы $x\bar{y}\bar{z}$ и $\bar{x}yz$ и «часть» классов xyz и $\bar{x}\bar{y}\bar{z}$. Окажется ли эта «часть» или «некоторые» на самом деле целым или только частью, или даже пустой, будет зависеть от обстоятельств».

Мы видим, таким образом, что Венн дает диаграмматическое истолкование и тому аналитическому способу решения логических уравнений, который изложен у него и Шредера.

6. Задача о девицах. Остановимся на диаграммамическом решении задачи Порецкого.

«Относительно девиц, бывших на данном бале, даны следующие 14 посылок:

1) каждая из девиц была или благовоспитанна, или весела, или молода, или красива;

2) когда начались танцы, то оказалось, что все нетанцующие девицы были некрасивы, и что каждая из танцующих была или молода, или весела, или благовоспитанна;

3) в другой момент, когда все пожилые девицы образовали отдельный кружок, о каждой из прочих девиц можно было сказать, что она или красива, или весела, или благовоспитанна;

4) если выделить всех девиц немолодых и некрасивых, то останутся только благовоспитанные и веселые девицы;

5) если же выделить всех девиц невеселых, то останутся благовоспитанные, молодые и красивые;

6) таких девиц, которые, обладая молодостью и веселостью, не обладали бы в то же время ни красотой, ни благовоспитанностью, на балу не было вовсе;

7) между молодыми девицами не было таких, которые, обладая красотой и веселостью, были бы не благовоспитанны;

8) каждая благовоспитанная девица была или молода, или весела, или красива;

9) все девицы, соединявшие красоту с благовоспитанностью, были одни веселы, другие молоды;

10) каждой невеселой девице не доставало или молодости, или красоты, или благовоспитанности;

11) все те веселые девицы, которые, не отличаясь молодостью, обладали благовоспитанностью, были красивы;

12) немолодые девицы были одни не благовоспитанны, другие не веселы, третьи не красивы;

13) между некрасивыми девицами не было таких, которые с благовоспитанностью соединяли бы молодость и веселость;

14) когда уехали все неблаговоспитанные, все немолодые, все невеселые и все некрасивые девицы, никаких девиц на балу более не осталось.

Узнать, прежде всего, возможна ли подобная сложная задача? Нет ли между ее посылками противоречий? Если окажется, что задача возможна, то описать точным образом весь мир девиц бала и определить отношение между различными категориями этих девиц...

Пусть a — благовоспитанные, b — веселые, c—молодые, d—красивые девицы бала. Посылки суть: **1)** $1 = a + b + c + d$; **2)** $1 = a + b + c + \bar{d}$; **3)** $1 = a + b + \bar{c} + d$; **4)** $1 = a + b + \bar{c} + \bar{d}$; **5)** $1 = a + \bar{b} + c + d$; **6)** $0 = \bar{a}bc\bar{d}$, **или** $1 = a + \bar{b} + \bar{c} + d$; **7)** $0 = \bar{a}bcd$, **или** $1 = a + \bar{b} + \bar{c} + \bar{d}$; **8)** $a = a(b + c + d)$, **или** $1 = \bar{a} + b + c + d$; **9)** $ad = ad(b + c)$, **или** $1 = \bar{a} + b + c + \bar{d}$; **10)** $\bar{b} = \bar{b}(\bar{a}+\bar{c}+\bar{d})$, **или** $1 = \bar{a} + b + \bar{c} + \bar{d}$; **11)** $ab\bar{c} = ab\bar{c}d$, или $1 = \bar{a} + \bar{b} + c + d$; **12)** $\bar{c} = \bar{c}(\bar{a} + \bar{b} + \bar{d})$, или $1 = \bar{a} + \bar{b} + c + \bar{d}$; **13)** $abc\bar{d} = 0$, или $1 = \bar{a} + \bar{b} + \bar{c} + d$; **14)** $1 = \bar{a} + \bar{b} + \bar{c} + \bar{d}$».

Задачу Порецкий решает аналитически. Система посылок избыточна. Перемножая все посылки, записанные в виде $1 = A_i (i = 1, \ldots, 14)$, он получает $1 = A$; справа от знака

равенства стоит выражение *A,* отличное от тождественного нуля; в этом и только в этом случае система посылок $1 = A_i\ (i = 1, \ldots, m;\ A_i$ — сумма (дизъюнкция) произведений (конъюнкций) некоторых из данных классов $a_1 \ldots, a_n$ и их дополнений) непротиворечива. Порецкому пришлось перемножать выражения вида $1 = B$, $1 = C$ друг на друга не менее 27 раз.

Более просто вопрос об упрощении системы посылок, а, следовательно, и вопрос об их непротиворечивости, может быть решен с помощью диаграммы Венна, построенной на рис. 41.

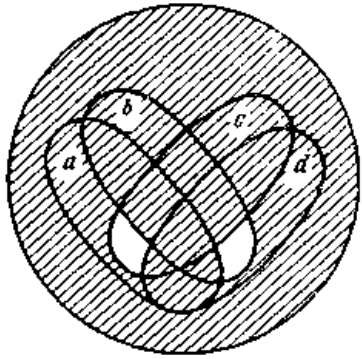


Рис. 41

Например, в восьмой посылке говорится, что непустыми могут быть только ячейки, принадлежащие хотя бы одному из следующих классов: $\bar{a}$, b, c, или d; следовательно, ячейка, не принадлежащая ни одному из этих классов (т. е. ячейка $a\bar{b}\bar{c}\bar{d}$) пуста (эту же ячейку можно вычеркнуть, не преобразуя предварительно $a = a\,(b + c + d)$ к виду $1 = \bar{a} + b + c + d$). На диаграмме не заштрихованы только две ячейки: $a\bar{b}c\bar{d}$ и $\bar{a}b\bar{c}d$. Поэтому система из 14 посылок эквивалентна выражению

$$1 = a\bar{b}c\bar{d} + \bar{a}b\bar{c}d$$

(что совпадает с полученным Порецким ответом).

Таким образом, поставленная задача вполне возможна, а всех девиц бала можно разбить на две группы:

1) девицы, которые, будучи благовоспитанны и молоды, не были ни веселы, ни красивы;

2) девицы, которые, будучи веселы и красивы, не были ни благовоспитанны, ни молоды.

К этой задаче мы вернемся в приложении, где определим отношения между различными категориями девиц, присутствующих на балу.

7. Теорема Гаубера. Нами уже рассматривалась теорема Гаубера. Докажем ее графически.
Дано: $\alpha\bar{\beta} = 0,\ y\bar{\beta} = 0,\ x\bar{\alpha} = 0,\ \alpha + \beta = x + y.$
Тогда $\alpha\bar{x} = 0,\ \beta\bar{y} = 0,\ xy = 0.$

В силу условия классы $\alpha\bar{\beta},\ y\bar{\beta},\ x\bar{\alpha}$ на диаграмме четырех переменных $\alpha,\ \beta,\ x,\ y$ (рис. 42) пусты; классы $\alpha + \beta$ и $x + y$ графически совпадают, т. е. часть класса $x + y$, не принадлежащая α+β, и часть класса α+β, не принадлежащая $x + y$, пусты (рис. 42).

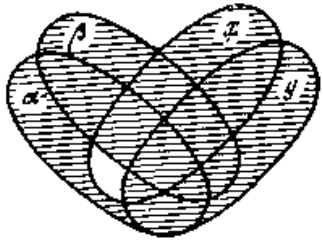


Рис. 42

На диаграмме видно, что классы $\alpha\bar{x},\ \beta\bar{y},\ xy$ пусты (заштрихованы), что и требовалось показать.

8. О выводе логических следствий. Примеры вывода следствий из посылок уже встречались выше. Разберем еще два: один из работы Венна, другой — Порецкого.

Пример 1. Посылки:
(а) Все *x* суть или *y* и *z*, или не-*y*.
(б) Если какие-нибудь *xy* суть *z*, то оии суть *w*.
(в) Никакие *wx* не суть *yz*.
Следствие: Никакие *x* не суть *y*.
В посылках — четыре различных класса: *x, y,* z, *w*. Поэтому плоскость делится не менее, чем на 2^4 ячеек, (на рис. 43 плоскость разбита на 16 ячеек с помощью четырех эллипсов).

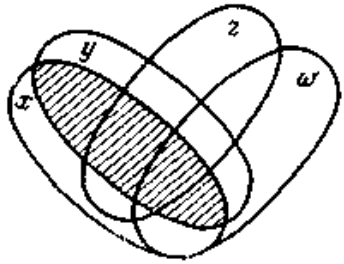


Рис. 43

Из первой посылки следует, что все элементы класса *x* заключены или в классе *yz,* или находятся вне класса *y,* т. е. часть класса *x,* принадлежащая *y* и не принадлежащая *z* (не зависимо от того, принадлежит она *w* или нет), пуста. Надиаграмме (рис. 43) заштрихованы ячейки

$$xy\bar{z}w$$

и

$$xy\bar{z}\bar{w}.$$

Во второй посылке говорится о том, что все элементы класса *xy,* принадлежащие классу z, заключены в классе *w,* т. е. класс $xyz\bar{w}$ пуст (на рис. 43 соответствующая ячейка заштрихована).

В третьей посылке утверждается, что класс *xyzw* пуст. На диаграмме (рис. 43) ячейка *xyzw* заштрихована.

В результате получаем, что на рис. 43 ячейка *xy* полностью заштрихована, т. е., что «Никакие *x* не суть *y*».

Следует подчеркнуть, что этот пример решается нами в отличие от Венна графически (даже посылки не переписываем предварительно в аналитической форме).

У Венна диаграммы используются только для геометрической иллюстрации результата аналитических преобразований. Мы показываем, что все преобразования можно проводить на диаграммах (без использования аналитического аппарата). Иногда, разумеется, удобно использовать комбинированный метод — сочетание аналитического и графического.

Этот комбинированный метод опишем подробнее на примере решения следующей задачи Порецкого (при этом рассматривается только часть логики классов, описываемая полностью на языке формул исчисления высказываний).

Пример 2.

Дана система равенств:

$$\left.\begin{aligned} a &= ab, \\ a &= a + c. \end{aligned}\right\} \qquad (1.20)$$

Требуется доказать:

1) Система (1.20) эквивалентна (в терминологии Порецкого — «тождественна») уравнению

$$a = ab + \bar{a}c \qquad (1.21)$$

(дополнение Порецкий в разных работах обозначает по-разному, так, как и Шредер, он пишет a_1, а не $\bar{a}$).

2) Из (1.20) следует уравнение

$$a = b(a + c), \qquad (1.22)$$

но не обратно. Порецкий в этом случае говорит, что система (1.20) «равнозначна с равенством» (1.22).

Формулу (1.21) Порецкий называет «полным», а формулу (1.22) «точным» определением *а.*

Первый шаг комбинированного метода состоит в представлении информации, содержащейся в условиях задачи, в виде равенств типа $A = 0$.

Так как равенство $A = B$ эквивалентно равенству

$$A\bar{B} + \bar{A}B = 0,$$

то (1.20), (1.21) и (1.22) перепишем в виде:

$$\left.\begin{aligned} a(\overline{ab}) + \bar{a}ab = 0, \\ a\overline{(a+c)} + \bar{a}(a+c) = 0, \end{aligned}\right\} \qquad (1.23)$$

$$a\overline{(ab+\bar{a}c)} + \bar{a}(ab+\bar{a}c) = 0, \qquad (1.24)$$

$$a\overline{(b(a+c))} + \bar{a}b(a+c) = 0. \qquad (1.25)$$

На втором шаге равенства типа $A = 0$ преобразуются в равенства типа $C = 0$ (эквивалентным образом), где C есть сумма произведений, членами которых являются буквы, обозначающие графически различные классы (заданные в условии задачи), и буквы с чертой, обозначающие дополнения соответствующих классов (при этом предполагается, что в каждое слагаемое каждая из букв может входить только один раз, с чертой или без черты), т. е. C есть аналог дизъюнктивной нормальной формы в исчислении высказываний. Преобразование A в C можно проводить по известным правилам логики классов, а также на диаграммах, предварительно определив соответствующие операции над ними. На последнем мы не будем останавливаться, т. к. в следующей главе изучаются операции над диаграммами Венна в исчислении высказываний.

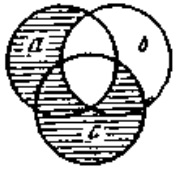


Рис. 44

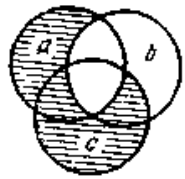


Рис. 45

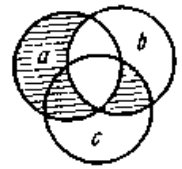


Рис. 46

После преобразований получаем диаграммы: диаграмма на рис. 44 соответствует объединению (сумме) уравнений системы (1.23), на рис. 45 — уравнению (1.24), на рис. 46 — уравнению (1.25).

Будем говорить, что диаграмма $\mathfrak{A}$ является следствием диаграммы $\mathfrak{B}$ тогда и только тогда, когда все незаштрихованные ячейки диаграммы $\mathfrak{B}$ не заштрихованы и на диаграмме $\mathfrak{A}$. Предполагаем при

этом, что число переменных на диаграмме $\mathfrak{A}$ совпадает с числом переменных на диаграмме $\mathfrak{B}$.

Диаграммы $\mathfrak{A}$ и $\mathfrak{B}$ называем равными тогда и только тогда, когда $\mathfrak{A}$ является следствием $\mathfrak{B}$, а $\mathfrak{B}$ является следствием $\mathfrak{A}$. (Для сравнения см. аналогичные понятия в следующих главах.)

Нетрудно показать, что система равенств *S* логически следует из системы равенств *T* тогда и только тогда, когда диаграмма, соответствующая системе *S*, является следствием диаграммы, соответствующей системе *T*.

Аналогично легко убедиться, что две системы равенств *S* и *T* эквивалентны тогда и только тогда, когда соответствующие им диаграммы равны.

В задаче Порецкого диаграммы на рис. 44 и 45 равны, а диаграмма рис. 46 является следствием диаграммы на рис. 44, но не обратно. Поэтому система (1.20) эквивалентна уравнению (1.21), а уравнение (1.22) логически следует из системы (1.20).

Вспомним, наконец, пример Венна, рассмотренный нами в п.1.4.1.4). Посмотрим диаграмму (рис. 47), соответствующую уравнению

$$w(1 - xz - y\bar{z})(1 - yz - x\bar{y}) = \bar{x}z,$$

которым символически выражается по Венну условие этой задачи. На диаграмме видно, что в классе *w* могут быть не пусты только ячейки $\bar{x}y\bar{z}$, x и $\bar{x}y\bar{z}$, через которые, как мы уже знаем, выражается *w* при условии $\bar{x}yz = 0$.

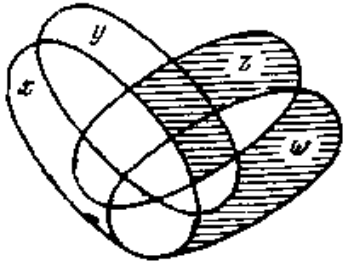


Рис. 47

9. Графический метод решения логических уравнений. Выше мы имели примеры соответствия между диаграммами Венна и логическими уравнениями: по заданному логическому уравнению строили соответствующую ему диаграмму Венна, и, наоборот, по известной диаграмме Венна находили соответствующее ей логическое уравнение.

Покажем, что на диаграмме, соответствующей логическому уравнению относительно неизвестной a_i, можно увидеть как условие разрешимости уравнения, так и общий вид его решения.

Пусть D — диаграмма Венна, соответствующая заданному логическому уравнению $\mathfrak{A}$. Будем предполагать, что на диаграмме D по крайней мере одна ячейка заштрихована — в противном условие задачи, записанное на языке логического уравнения $\mathfrak{A}$, содержит противоречие: из него можно вывести равенство 1=0. Пусть $a_1, \ldots, a_n$— все графически различные переменные диаграммы D; требуется разрешить уравнение $\mathfrak{A}$ относительно неизвестной a_i.

Каждой ячейке диаграммы D можно поставить в соответствие произведение (конъюнкцию), в которое не входит a_i (и $\bar{a}_i$):

$$\tilde{a}_1 \ldots \tilde{a}_{i-1}\tilde{a}_{i+1} \ldots \tilde{a}_n, \qquad (1.26)$$

$\tilde{a}_j$ есть a_j, если ячейка принадлежит классу a_j, и $\tilde{a}_j$ есть $\bar{a}_j$, если ячейка принадлежит дополнению класса a_j ($j = 1, \ldots, i-1, i+1, \ldots, n$); $\tilde{a}_1 \ldots \tilde{a}_{i-1}\tilde{a}_{i+1} \ldots \tilde{a}_n$ эквивалентно

$\tilde{a}_1 \ldots \tilde{a}_{i-1}a_i\tilde{a}_{i+1} \ldots \tilde{a}_n + \tilde{a}_1 \ldots \tilde{a}_{i-1}\bar{a}_i\tilde{a}_{i+1} \ldots \tilde{a}_n$.

На диаграмме D можно выделить ячейки как принадлежащие, так и не принадлежащие a_i. Уравнение $\mathfrak{A}$ эквивалентно строящемуся по диаграмме D уравнению:

$$Ha_i + G\bar{a}_i = 0, \qquad (1.27)$$

где H и G — суммы произведений вида (1.26), определяемых по всем заштрихованным ячейкам классов a_i и $\bar{a}_i$ соответственно.

Например, диаграмме рис. 40 соответствует уравнение

$$(\bar{x}\bar{y}z + \bar{x}y\bar{z} + x\bar{y}z + xy\bar{z})w + (\bar{x}y\bar{z} + \bar{x}yz + x\bar{y}\bar{z} + x\bar{y}z)\bar{w} = 0 \qquad (1.28)$$

(в классе w заштрихованы ячейки с номерами 0011, 0101, 1011 и 1101, а в классе $\bar{w}$ — 0100, 0110, 1000 и 1010;

переменные диаграммы рассматриваем в порядке x, y, z, w).

Уравнение (1.27) разрешимо тогда и только тогда, когда $HG = 0$ (см. п. 1.4.1.2, 1.4.1.4).

Условие $HG=0$ эквивалентно равенству $R=0$, где R — сумма произведений вида (1.26), которые пишутся по всем тем заштрихованным ячейкам фигуры a_i, соседние которых относительно a_i также заштрихованы.

Ячейки, соответствующие $\tilde{a}_1 \ldots \tilde{a}_{i-1}a_i\tilde{a}_{i+1} \ldots \tilde{a}_n$ и $\tilde{a}_1 \ldots \tilde{a}_{i-1}\bar{a}_i\tilde{a}_{i+1} \ldots \tilde{a}_n$, будем называть соседними относительно фигуры (класса) a_i. Двоичные n-значные номера двух соседних между собою относительно a_i ячеек отличаются только тем, в одном на i-м месте находится единица, в другом — нуль.

Например, $\bar{x}y\bar{z} + x\bar{y}z = 0$ — необходимое и достаточное условие разрешимости уравнения относительно w, которому соответствует

диаграмма на рис. 40 (на диаграмме пары ячеек с номерами 0101, 0100 и 1011, 1010 заштрихованы).

Действительно, пусть $HG = 0$. Предположим, что неверно $R = 0$. Тогда существует $j\ (j = 1, \ldots, k)$ такое, что неверно $R_j = 0$ ($R_1, \ldots, R_k$ — произведения вида (1.26), написанные по всем заштрихованным ячейкам фигуры a_i, соседние которых относительно a_i также заштрихованы, т. е. $R_1 + \ldots + R_k$ есть R). Следовательно, так как R_j входит в H и в G, неверно, что $H = 0$ и что $G = 0$, а это противоречит тому, что $HG = 0$. Таким образом, если $HG = 0$, то $R = 0$.

Наоборот, если $R = 0$, то $H = 0$ или $G = 0$ (т. к. все общие члены сумм H и G, и только они, являются слагаемыми R), т. е. $HG = 0$.

Как показано в п.1.4.1.2 и 1.4.1.4, решение уравнения (1.27) можно представить в виде

$$a_i = \bar{H}G + u\bar{H}\bar{G} \tag{1.29}$$

при условии $HG = 0$, где u — произвольный класс.

На диаграмме D в класс $\bar{H}G$ входят:

(1) незаштрихованные ячейки класса a_i, соседние которых относительно a_i заштрихованы, и

(2) заштрихованные ячейки из $\bar{a}_i$, соседние которых относительно a_i не заштрихованы, т. е. $\bar{H}G$ эквивалентно Q, где Q — сумма произведений вида (1.26), написанных по всем незаштрихованным ячейкам класса a_i, соседние которых относительно a_i заштрихованы.

Аналогично, $\bar{H}\bar{G}$ эквивалентно S, где S — сумма произведений вида (1.26), написанных по всем незаштрихованным ячейкам класса a_i соседние которых относительно a_i не заштрихованы.

Следовательно, равенство (1.29) эквивалентно равенству

$$Q_i = Q + uS, \tag{1.30}$$

где u — произвольный класс, Q — сумма произведений вида (1.26), написанных по всем незаштрихованным ячейкам из a_i, соседние которых относительно a_i заштрихованы, S — сумма произведений вида (1.26), написанных по всем незаштрихованным ячейкам из a_i, соседние которых относительно a_i не заштрихованы.

Таким образом, решение уравнения (1.27) относительно a_i можно представить в виде (1.30) при условии $R=0$, где R — сумма произведений вида (1.26), написанных по всем заштрихованным ячейкам из a_i соседние которых относительно a_i заштрихованы. Результат подстановки $Q+uS$ вместо a_i в равенство (1.27) эквивалентен, при $R = 0$, тождеству $0 = 0$.

Следует обратить внимание на то, что равенство

$$a_i = Q + uS$$

не эквивалентно уравнению (1.27) ни при каких значениях *u*. В самом деле, равенство $a_i = Q + uS$ эквивалентно

$$\bar{Q}(\bar{S} + \bar{u})\, a_i + (Q + Su)\, \bar{a}_i = 0, \qquad (1.31)$$

где класс $\bar{Q}\bar{S}\ a_i$ охватывает все заштрихованные ячейки уравнения (1.27), принадлежащие a_i, в класс $Q\bar{a}_i$ входят все заштрихованные ячейки из $\bar{a}_i$, определяемые (1.27), и соседние относительно a_i ячейки которых не пусты. В зависимости от класса *u* равенство (1.31) может определять пустые ячейки, пустота которых не следует из (1.27).

Существуют такие классы, например, класс a_i, после подстановки которых вместо *u* из (1.31) получаем равенство, объединение которого с $R = 0$ эквивалентно (1.27).

Вернемся к диаграмме рис. 40. В классе *w* не заштрихованы ячейки с номерами 0001, 1111, 0111, 1001, соседние относительно *w* последних двух ячеек заштрихованы. Поэтому выражение класса *w* через классы *x, y, z* — при условии $\bar{x}yz + x\bar{y}z = 0$ — имеет вид:

$$w = \bar{x}yz + x\bar{y}\bar{z} + u(\bar{x}\bar{y}\bar{z} + xyz), \qquad (1.32)$$

где *u* — произвольный класс (что совпадает с ответом задачи о провалившихся на экзамене).

Сделаем проверку. В (1.28) вместо *w* подставим правую часть равенства (1.32):

$$(\bar{x}\bar{y}z + \bar{x}y\bar{z} + x\bar{y}z + xy\bar{z})(\bar{x}yz + x\bar{y}\bar{z} +$$

$$+ u\,(\bar{x}\bar{y}\bar{z} + xyz)) + (\bar{x}y\bar{z} + \bar{x}yz + x\bar{y}\bar{z} + x\bar{y}z) \times$$

$$\times(\bar{x}\bar{y}z + \bar{x}y\bar{z} + x\bar{y}z + xy\bar{z} + \bar{u}(\bar{x}\bar{y}\bar{z} + xyz)) = 0,$$

после раскрытия скобок получаем

$$\bar{x}y\bar{z} + x\bar{y}z = 0,$$

т. е. условие разрешимости уравнения (1.28). Приведем (1.32) к нулевой форме

$$(\bar{x}yz + x\bar{y}\bar{z} + u(\bar{x}\bar{y}\bar{z} + xyz))\,\bar{w} +$$

$$+ (\bar{x}\bar{y}z + \bar{x}y\bar{z} + x\bar{y}z + xy\bar{z} + \bar{u}(\bar{x}\bar{y}\bar{z} + xyz))\,w = 0. \qquad (1.33)$$

Полученное равенство не эквивалентно (1.28) и не является его следствием. Вместо *u* в (1.33) подставим *w*, получим равенство

$$(\bar{x}\bar{y}z + \bar{x}y\bar{z} + x\bar{y}z + xy\bar{z})\bar{w} + (\bar{x}yz + x\bar{y}\bar{z})w = 0,$$

которое вместе с $\bar{x}y\bar{z} + x\bar{y}z = 0$, эквивалентно (1.28).

10. Исключение неизвестных. С примерами исключения неизвестных с помощью диаграмм Венна мы уже встречались в задаче Буля. Остановимся теперь на общем методе такого исключения. На диаграммах Венна получаем ответ на оба вопроса:

1) возможно ли исключение,

2) если возможно, то каков его результат.

Для описания графического метода исключения переменных удобнее воспользоваться табличным представлением диаграмм Венна.

На одной стороне таблицы (для определенности возьмем левую сторону) расположим только имена классов (и их дополнений), которые требуется исключить. Тогда, если в построенной диаграмме будут полностью заштрихованные столбцы, то на первый вопрос можно дать положительный ответ: указанные в условии задачи классы могут быть исключены. Если же на диаграмме нет полностью заштрихованных столбцов, то ответ на первый вопрос будет отрицательным: указанные в условии задачи классы не исключаются.

В первом случае (когда ответ на первый вопрос положителен) результат исключения пишется по всем заштрихованным столбцам, в имена которых не включаются классы, обозначенные на левой стороне таблицы Венна. Этот метод продемонстрируем на примере.

Дано:

$$x\bar{z}\,(\bar{u} + wy + \bar{w}\bar{y}) = 0,$$
$$uxw\,(y\bar{z} + \bar{y}z) = 0,$$
$$x\,(u + y)\,(zw + \bar{z}\bar{w}) = 0,$$
$$(\bar{x} + \bar{u}\bar{y})\,(z\bar{w} + \bar{z}w) = 0.$$

Требуется исключить *u* и *y,* а затем определить *x*. Задачу Венн решает графически: строится диаграмма, описывающая заданную систему уравнений (рис. 48); выделяются все пустые классы, не содержащие *u* и *y;* такими пустыми классами оказываются xzw, $\bar{x}z\bar{w}$ и $x\bar{z}w$; следовательно, результат исключения *u* и *y* записывается в виде уравнения

$$xzw + xz\bar{w} + x\bar{z}w = 0.$$

Разрешая полученное уравнение относительно *x,* Венн получает

$$x = z\bar{w} + \bar{z}w + v\bar{z}\bar{w}$$

(на диаграмме видно, что в классе *x* могут быть непустыми ячейки $z\bar{w}$, $\bar{z}w$, $\bar{z}\bar{w}$).

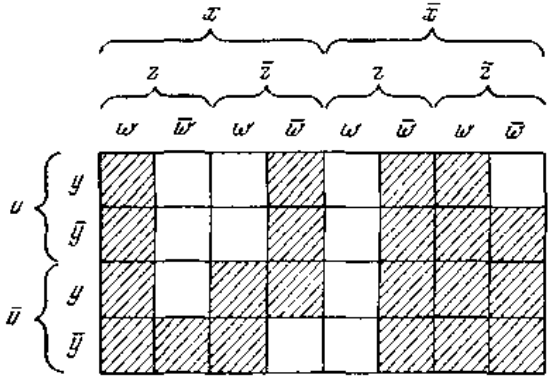


Рис. 48

Во втором издании книги «Символическая логика» Венн привел пример исключения неизвестных из системы, содержащей неравенства (соответствующие частным суждениям). Этот пример он решает с помощью диаграмм.

Дано:

$$\begin{cases} x = \bar{c}y + e\bar{y}, \\ z \neq c\bar{w} + \bar{e}w. \end{cases} \tag{1.34}$$

Требуется установить отношение между *x, y,* z, *w* после исключения *c* и *e*.

Преобразуя равенства вида $A = B$ в $A\bar{B} + \bar{A}B = 0$, а неравенства вида $A \neq B$ в $A\bar{B} + \bar{A}B > 0$, Венн из (1.34) получает

$$x(cy + \bar{e}\bar{y}) + (\bar{c}y + e\bar{y})\bar{x} = 0,$$
$$z(\bar{c}\bar{w} + ew) + (c\bar{w} + \bar{e}w)\bar{z} > 0.$$

В посылках (1.34) находятся одно универсальное и одно частное суждения. В этом случае в классах, которые могут быть непустыми, Венн ставит звездочки, предполагая, что, по крайней мере, один из отмеченных классов не пуст. Условия (1.34) Венн выражает диаграммой, которая показана на рис. 33. Из этой диаграммы следует:

1. Не существует ни одного полностью заштрихованного столбца, т. е. соотношение между *x, y,* z, *w* нельзя записать в виде равенства.
2. Существуют классы, в имена которых не входят буквы *c* и *e,* но объединение которых содержит все звездочки рассматриваемой диаграммы; такими классами являются xz, $x\bar{z}$, yw, $\bar{y}\bar{w}$. Следовательно, результатом исключения переменных *c* и *e* из (1.34) служит неравенство:

$$xz + \overline{xz} + yw + \overline{yw} > 0.$$

Аналогичный ответ можно получить по формулам (1.19), исключая переменные последовательно, например, сначала *c*, потом *e.*

В этом примере Венн ограничился системой, в которую входит только одно неравенство. Если число неравенств в системе увеличить, то между звездочками, соответствующими одному неравенству, должна быть установлена связь, с помощью которой они графически отличаются от звездочек, выражающих другие неравенства. Это следует из того, что для каждого неравенства по крайней мере один из отмеченных классов (соответствующих данному неравенству) не пуст.

Последователи Венна соединяют все звездочки, соответствующие одному частному суждению, прямыми, получая одну ломаную (иногда звездочки совсем опускают).

Например, истинность модуса disamis:

Некоторые *M* суть *P*.
Все *M* суть *S*
Некоторые *S* суть *P*.

С. Лущевска-Романова проверяет с помощью диаграммы, показанной на рис. 49.

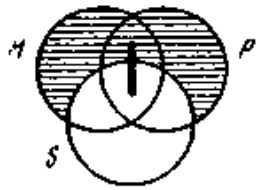


Рис. 49

Э. Беркли многозначность положения непустого класса, соответствующего предложению «Некоторые *a* суть *b*» *(ab≠0),* показывает на диаграмме трех переменных *a, b, c* (рис. 50).

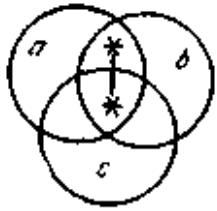


Рис. 50

В заключение главы подчеркнем, что аппарат диаграмм Венна оказывается применимым при решении всех вообще задач логики, входивших в проблематику математической логики в конце XIX века.

Пренебрежительная оценка этого аппарата некоторыми логиками на рубеже XIX и XX столетий,— например, в «Алгебре логики» Л. Кутюра,— являются поэтому неоправданной.

Графические методы Венна сохраняют интерес и сейчас. Их можно использовать при решении таких вопросов, как обзор логических следствий из заданной информации или минимизация формул исчислений. Диаграммы Венна можно применять в графических методах математической логики с неменьшим успехом, чем, скажем, карты Карнау (1953 г.) или диаграммы Вейча (1952 г.). В последующих разделах мы остановимся на решении некоторых задач логики с помощью диаграмм Венна. Прежде всего, в этой связи, рассмотрим диаграммы Венна с точки зрения классического исчисления высказываний. Хотя Венн и не изучал специально вопрос о применении диаграмм в исчислении высказываний, однако он останавливается на высказываниях как частном случае предложений логики классов.

Геометрическая наглядность (особенно, когда приходится иметь дело с небольшим числом пропозициональных переменных) диаграмм в исчислении высказываний побудила американского ученого Мак-Каллока считать их особенно пригодными при построении схем, отражающих работу нервных сетей мозга. При этом У. Мак-Каллок обращает основное внимание на непустые ячейки диаграммы, в них он ставит точки. Используя диаграммы Венна при описании работы нейронных схем (мы остановимся на этих вопросах дальше), Мак-Каллок не исследовал ряда естественно возникающих при этом вопросов. Это относится, в частности, к алгоритмам вывода логических следствий из посылок средствами исчисления высказываний, построенного графически с помощью диаграмм Венна.

Как известно, все предложения логики классов не могут быть формализованы в исчислении высказываний в силу отсутствия возможности представления частных суждений. Венн в 1894 г. дополнил книгу частными суждениями, используя для этого знак неравенства, а на диаграммах — звездочки. Однако графическими исследовал только один частный случай. Последователи графического метода Венна ввели для записи частных суждений ломаные (иногда содержащие в вершинах звездочки). Полного же разбора графических методов, относящихся к частным суждениям, в литературе нет. Поэтому дальше нами будет рассмотрен случай, когда посылки выразимы на языке исчисления одноместных предикатов, вводится понятие простого логического следствия из данных посылок и предлагается метод, позволяющий на диаграммах обозреть все простые логические следствия. Предварительно определяются диаграммы Венна в исчислении одноместных предикатов и описывается способ,

позволяющий с помощью только диаграмм Венна решить проблему разрешения.

Изучение диаграмм Венна подсказало А.С.Кузичеву некоторые изменения решения проблемы разрешения для исчисления одноместных предикатов. Венн не занимался проблемами разрешимости и разрешения. Но для того, чтобы обосновать его метод преобразования информации (в которую могут входить и частные предложения), А.С.Кузичевым решается проблема разрешения для формул исчисления одноместных предикатов с помощью только диаграмм Венна. Из ее решения получается общее правило, позволяющее обозревать логические следствия данных посылок.

4.2. Диаграммы Венна в классическом исчислении высказываний

4.2.1. Соответствие между диаграммами Венна и бинарными матрицами *n* переменных

Определение. Диаграммой Венна *n* переменных в классическом исчислении высказываний будем называть символ Венна *n* переменных, в некоторых ячейках которого может стоять по одной точке.

Ячейки символа Венна *n* переменных занумерованы числами от 0 до $2^n - 1$, поэтому диаграмму Венна можно представить в виде

$$(\beta_0 \beta_1 \ldots \beta_{2^n-1}), \tag{2.1}$$

где

$$\beta_i \doteq \begin{cases} 1, \text{ если в } i\text{-ой ячейке соответствующей диаграммы находится точка,} \\ 0, \text{ если } i\text{-ая ячейка рассматриваемой диаграммы пуста} \end{cases}$$

($\doteq$ — знак графического равенства, $i = 0, 1, \ldots, 2^n - 1$).

При больших *n* вместо (2.1) удобнее использовать матричную запись:

$$\begin{pmatrix} \beta_0 & \beta_1 & \cdots & \beta_{2^k-1} \\ \beta_{2^k} & \beta_{2^k+1} & \cdots & \beta_{2^{k+1}-1} \\ \cdots & \cdots & \cdots & \\ \beta_{2^k(2^m-1)} & \beta_{2^k(2^m-1)+1} & \cdots & \beta_{2^{k+m}-1} \end{pmatrix}, \tag{2.2}$$

числа *k* и *m* выбираются так, чтобы $k \geqslant 0$, $m \geqslant 0$, $k + m \neq 0$, $k + m = n$, β_i есть 0 или 1 ($i = 0, 1, \ldots, 2^n - 1$).

Ясно, что, если задана матрица вида (2.2), то легко строится соответствующая ей диаграмма Венна n переменных. (В этом разделе для краткости вместо слов «диаграмма Венна n переменных в классическом исчислении высказываний» часто будем употреблять слова «диаграмма Венна n переменных».)

Матрицу (2.2), где β_i есть 0 или1 $(i = 0, 1, \ldots, 2^n - 1)$, будем называть бинарной.

Например, при $m = 3$, $\kappa = 2$ имеем бинарную матрицу:

$$\begin{pmatrix} \beta_0 & \beta_1 & \beta_2 & \beta_3 \\ \beta_4 & \beta_5 & \beta_6 & \beta_7 \\ \beta_8 & \beta_9 & \beta_{10} & \beta_{11} \\ \beta_{12} & \beta_{13} & \beta_{14} & \beta_{15} \\ \beta_{16} & \beta_{17} & \beta_{18} & \beta_{19} \\ \beta_{20} & \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{24} & \beta_{25} & \beta_{26} & \beta_{27} \\ \beta_{28} & \beta_{29} & \beta_{30} & \beta_{31} \end{pmatrix} \qquad (2.3)$$

Числа $0, \ldots, 2^n - 1$, записанные в n-значной двоичной записи, можно рассматривать как элементарные последовательности n переменных.

Все различные элементарные последовательности n переменных будем объединять в таблицы (матрицы) вида (2.2), в которых β_i есть двоичная запись с n знаками числа i. Так, матрице (2.3) соответствует таблица элементарных последовательностей пяти переменных:

$$\begin{pmatrix} 0\,0\,0\,0\,0 & 0\,0\,0\,0\,1 & 0\,0\,0\,1\,0 & 0\,0\,0\,1\,1 \\ 0\,0\,1\,0\,0 & 0\,0\,1\,0\,1 & 0\,0\,1\,1\,0 & 0\,0\,1\,1\,1 \\ 0\,1\,0\,0\,0 & 0\,1\,0\,0\,1 & 0\,1\,0\,1\,0 & 0\,1\,0\,1\,1 \\ 0\,1\,1\,0\,0 & 0\,1\,1\,0\,1 & 0\,1\,1\,1\,0 & 0\,1\,1\,1\,1 \\ 1\,0\,0\,0\,0 & 1\,0\,0\,0\,1 & 1\,0\,0\,1\,0 & 1\,0\,0\,1\,1 \\ 1\,0\,1\,0\,0 & 1\,0\,1\,0\,1 & 1\,0\,1\,1\,0 & 1\,0\,1\,1\,1 \\ 1\,1\,0\,0\,0 & 1\,1\,0\,0\,1 & 1\,1\,0\,1\,0 & 1\,1\,0\,1\,1 \\ 1\,1\,1\,0\,0 & 1\,1\,1\,0\,1 & 1\,1\,1\,1\,0 & 1\,1\,1\,1\,1 \end{pmatrix} \qquad (2.4)$$

В связи с возможностью представления диаграмм Венна n переменных бинарными матрицами, будем употреблять слова «бинарная матрица n переменных» как имеющие тот же смысл, что и слова «диаграмма Венна n переменных в исчислении высказываний».

Между диаграммами Венна в исчислении высказываний и диаграммами Венна в логике классов, не содержащими звездочек, можно установить взаимнооднозначное соответствие. Например, по диаграмме, построенной при решении задачи Буля (рис. 36), определяется бинарная матрица:

$$\begin{pmatrix} 0&0&0&0&0&0&1&1 \\ 0&1&0&0&0&0&1&1 \\ 1&0&0&1&0&1&0&0 \\ 0&0&0&1&1&1&0&0 \end{pmatrix}, \qquad (2.5)$$

где нули соответствуют заштрихованным ячейкам, а единицы — незаштрихованным, $k = 3$, $m = 2$.

4.2.2. Операции над диаграммами Венна

1. Увеличение числа переменных (переход от n переменных к $n + k$, $k > 0$, переменным). Пусть дана диаграмма Венна n переменных $D(n)$. Пусть $\alpha_1, \ldots, \alpha_h$ — все ячейки, в которых на диаграмме $D(n)$ находятся точки. Построим символ Венна $n+k$ переменных ($k>0$). Для этого каждую ячейку символа Венна диаграммы $D(n)$ разделим на 2^k ячеек; во всех ячейках, на которые разбиваются ячейки $\alpha_1, \ldots, \alpha_h$, и только в них, поставим по одной точке. В результате получится $D(n + k)$ — диаграмма Венна $n + k$ переменных.

В дальнейшем при рассмотрении операций над несколькими диаграммами мы будем предполагать, что над ними проведена операция увеличения числа переменных, так что они содержат одинаковое число переменных.

2. Отрицание (дополнение). Пусть дана диаграмма $D(n)$, где n — число переменных. Построим новую диаграмму Венна n переменных следующим образом. Начертим символ Венна n переменных. Во всех его ячейках, в которых на $D(n)$ нет точек, поставим по одной точке; ячейки, в которых на $D(n)$ есть точки, остаются пустыми. Полученную диаграмму будем называть отрицанием диаграммы $D(n)$ и обозначать $\neg D(n)$.

3. Конъюнкция (пересечение). Пусть даны диаграммы $D_1(n), D_2(n)$. Построим символ Венна n переменных. В тех и только тех его ячейках, в которых на D_1 и D_2 одновременно находятся точки, поставим по одной точке; полученную диаграмму будем называть конъюнкцией диаграммы D_1 и D_2 и обозначать $(D_1 \,\&\, D_2)$.

4. Дизъюнкция (объединение). Пусть даны диаграммы $D_1(n)$ и $D_2(n)$. Построим новый символ Венна n переменных. Во всех его ячейках, в которых на D_1 или на D_2 находятся точки (и только в них), поставим по одной точке; полученную диаграмму будем называть дизъюнкцией диаграмм D_1 и D_2 и обозначать $(D_1 \vee D_2)$.

Все указанные операции могут быть сформулированы и на языке бинарных матриц. Для примера остановимся на операции увеличения числа переменных.

Две строки (столбца) матрицы будем называть соседними относительно переменной a_i, если соответствующие им элементарные последовательности попарно отличны друг от друга только тем, что в одной из них на i-ом месте находится нуль, а в другой на i-м месте — единица, в остальном же эти элементарные последовательности совпадают.

Например, первый и второй, третий и четвертый столбцы матрицы (2.5) являются соседними относительно a_5 (предполагается, что переменные обозначены буквами

$$a_1, a_2, a_3, a_4, a_5).$$

В результате выполнения операции увеличения числа переменных (при добавлении переменной a_i) получим бинарную матрицу, во всех соседних относительно a_j строках (столбцах) которой стоят попарно одинаковые числа. Например, после добавления к числу переменных матрицы (2.5) новой переменной b (которая ставится между переменными a_2 и a_3) получим бинарную матрицу шести переменных $a_1, a_2, b, a_3, a_4, a_5$:

$$\begin{pmatrix} 0&0&0&0&0&0&1&1 \\ 0&0&0&0&0&0&1&1 \\ 0&1&0&0&0&0&1&1 \\ 0&1&0&0&0&0&1&1 \\ 1&0&0&1&0&1&0&0 \\ 1&0&0&1&0&1&0&0 \\ 0&0&0&1&1&1&0&0 \\ 0&0&0&1&1&1&0&0 \end{pmatrix} \qquad (2.6)$$

Между диаграммами Венна в исчислении высказываний и формулами этого исчисления установим соответствие. При этом мы не будем подробно излагать исчисление высказываний, отсылая читателя к соответствующей литературе, напомним лишь некоторые понятия. Рассматривается следующий алфавит: $a_1 \ldots a_n$ — пропорциональные переменные, $\neg \,\&\, \vee \supset$ — логические знаки ($\neg$ — знак отрицания, & — знак конъюнкции, $\vee$ — знак дизъюнкции, $\supset$ — знак импликации), ( ) — скобки.

Формулы исчисления высказываний определяются следующими порождающими правилами:

1. Если a_i — пропозициональная переменная, то a_i считается формулой.

2. Если Ф — формула, то $\neg\Phi$ считается формулой (иногда вместо $\neg\Phi$ будем писать $\bar{\Phi}$ или Ф').
3.Если Ф и Ψ— формулы, то $(\Phi \,\&\, \Psi)$, $(\Phi \vee \Psi)$, $(\Phi \supset \Psi)$ считаются формулами (иногда вместо $(\Phi \,\&\, \Psi)$ будем писать $(\Phi\,\Psi)$, а вместо $(\Phi \vee \Psi)$ — $(\Phi + \Psi)$.

4.2.3. Построение диаграмм Венна по данным формулам

Пусть дана формула Ф исчисления высказываний, составленная из переменных $a_1, \ldots, a_n$. (О формуле Ф говорят, что она составлена из пропозициональных переменных $a_1, \ldots, a_n$, если никакая отличная от $a_1, \ldots, a_n$ пропозициональная переменная не входит в Ф.) Требуется построить диаграмму Венна, соответствующую формуле Ф.
Диаграмму (бинарную матрицу), соответствующую данной формуле Ф, будем обозначать через [Ф], или $[\Phi]^n$, где n указывает на число графически различных переменных формулы Ф.
а) Каждой переменной a_i формулы Ф поставим в соответствие диаграмму Венна, получающуюся следующим образом. Построим символ Венна n переменных. Во всех ячейках i-ой фигуры, и только в них, поставим по одной точке. Получим диаграмму, соответствующую переменной a_i (т. е. диаграмму $[a_i]^n$).
б) Каждому логическому знаку формулы Ф поставим в соответствие операции над диаграммами Венна следующим образом:
1. Предположим, что $\neg F$ — подформула формулы Ф и что диаграмма $[F]$ построена. Построим диаграмму $[\neg F]$.
Диаграмму $[\neg F]$ определим как результат операции отрицания диаграммы $[F]$

$$[\neg F] \rightleftarrows \neg [F],$$

где $\rightleftarrows$ обозначает равенство по определению.
2. Предположим, что $(E \,\&\, F)$, $(E \vee F)$ и $(E \supset F)$ — подформулы формулы Ф и что диаграммы $[E]$ и $[F]$ построены. Построим диаграммы $[E \,\&\, F]$, $[E \vee F]$, $[E \supset F]$:

$$[E \,\&\, F] \rightleftarrows ([E] \,\&\, [F]),$$
$$[E \vee F] \rightleftarrows ([E] \vee [F]),$$
$$[E \supset F] \rightleftarrows (\neg [E] \vee [F]).$$

(В аналогичных выражениях для краткости внешние круглые скобки формул будем опускать.)
Описанный способ определения диаграммы (матрицы), соответствующей данной формуле Ф, имеет вид индуктивного

построения. На первом шаге (пункт а) строятся диаграммы (матрицы), соответствующие всем графически различным переменным формулы Ф. Индукция ведется по шагам построения формулы Ф (пункт б).

Определения. Пусть D_1, D_2 — бинарные матрицы одинакового числа переменных. Будем говорить, что D_1 входит в D_2 (записываем $D_1 \subset D_2$), если все единицы матрицы D_1 являются единицами матрицы D_2 (иными словами, $D_1 \subset D_2$, если каждой единице с номером α матрицы D_1 соответствует единица с тем же номером α на матрице D_2).

Матрицы D_1 и D_2 будем называть равными (записываем $D_1 = D_2$), если $D_1 \subset D_2$ и $D_2 \subset D_1$.

Условимся обозначать: $D_1 \neq D_2$, когда матрицы D_1 и D_2 не являются равными.

Две формулы Ф и Ψ, каждая из которых составлена из переменных $a_1, \ldots, a_n$, будем называть *эквивалентными,* если $[\Phi]^n = [\Psi]^n$.

Две формулы Ф и Ψ будем называть *различными,* если неверно, что $[\Phi] \subset [\Psi]$, и неверно, что $[\Psi] \subset [\Phi]$.

Формулу Ф назовем *тождественно истинной* (универсально общезначимой), если ее бинарная матрица (бинарная матрица, построенная по формуле Ф), не содержит нулей. Например, формула $((a_1 \vee a_3 \vee \neg a_4) \vee ((a_2 \vee a_5) \,\&\, \neg (a_2 \,\&\, a_5)))$ не является универсально общезначимой, т. к. ее бинарная матрица

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

содержит нули.

Формулу Ф называем тождественно ложной, если ее бинарная матрица не содержит единиц.

Формула Ф называется дизъюнктивной (конъюнктивной) нормальной формулой [сокращенно — ДНФ (КНФ)], если Ф есть дизъюнкция (конъюнкция) конъюнкций (дизъюнкций), членами которых являются пропозициональные переменные и их отрицания, при этом все дизъюнктивные (конъюнктивные) члены формулы Ф будем считать попарно неэквивалентными между собой.

ДНФ (КНФ) Ф называется совершенной дизъюнктивной (конъюнктивной) нормальной формулой [сокращенно — СДНФ (СКНФ)], если в каждый дизъюнктивный (конъюнктивный) ее член входят все графически различные пропозициональные переменные формулы Ф, при этом каждая переменная может входить (с

отрицанием или без) в дизъюнктивный (конъюнктивный) член толькв один раз.

4.2.4. Построение формул по диаграммам Веина

Пусть дана диаграмма $D(a_1, \ldots, a_n)$ (в скобках указаны все графически различные переменные $a_1, \ldots, a_n$ диаграммы Венна). Пусть $\alpha_1, \ldots, \alpha_h$—все различные ячейки, в которых на диаграмме D находятся точки. Ячейка α_i есть $\tilde{a}_1^i \ldots \tilde{a}_n^i$, где

$$\tilde{a}_j^i = \begin{cases} a_j, & \text{если } \alpha_i \text{ принадлежит фигуре } a_j, \\ \bar{a}_j, & \text{если } \alpha_i \text{ принадлежит дополнению фигуры } a_j \end{cases}$$
$$j = 1, \ldots, n; \quad i = 1, \ldots, h.$$

Ячейке α_i поставим в соответствие формулу

$$\Phi_i = (\varepsilon_1^i a_1 \& \ldots \& \varepsilon_n^i a_n),$$

где

$$\varepsilon_j^i = \begin{cases} \bigwedge, & \text{если } \tilde{a}_j^i \text{ есть } a_j, \\ \neg, & \text{если } a_j^i \text{ есть } \bar{a}_j, \end{cases}$$

$\bigwedge$ — обозначает пустое слово. (Некоторые скобки в формулах опускаем в силу ассоциативности как конъюнкции, так и дизъюнкции.) Диаграмме $D(a_1, \ldots, a_n)$ поставим в соответствие СДНФ

$$(\Phi_1 \vee \ldots \vee \Phi_h).$$

Формулу, соответствующую диаграмме Венна D, будем обозначать $[D]$. В СДНФ часто будем опускать все скобки, кроме внешних, и знак &, предполагая, что конъюнкция связывает сильнее дизъюнкции. Например, бинарной матрице (2.5) соответствует формула

$$((\neg a \& \neg b \& c \& d \& \neg e) \vee (\neg a \& \neg b \& c \& d \& e) \vee$$
$$\vee (\neg a \& b \& \neg c \& \neg d \& \neg e) \vee (\neg a \& b \& c \& d \& \neg e) \vee$$
$$\vee (\neg a \& b \& c \& d \& e) \vee (a \& \neg b \& \neg c \& \neg d \& \neg e) \vee (a \&$$
$$\& \neg b \& \neg c \& d \& e) \vee (a \& \neg b \& c \& \neg d \& e) \vee (a \& b \& \neg c \&$$
$$\& d \& e) \vee (a \& b \& c \& \neg d \& \neg e) \vee (a \& b \& c \& \neg d \& e)).$$

Если диаграмма $D(a_1, \ldots, a_n)$ не содержит точек, то в соответствие ей ставим тождественно ложную формулу, например, $(a_1 \& \neg a_1)$.

Если на диаграмме число ячеек, занятых точками, больше, чем число пустых ячеек, то для уменьшения логической длины построение удобнее начинать с пустых ячеек диаграммы и диаграмме $D(a_1, \ldots, a_n)$ ставить в соответствие формулу $\neg(\Phi_{h+1} \vee \ldots \vee \Phi_2^n)$, не являющуюся ДНФ. При этом используется эквивалентность:

$$(\Phi_1 \vee \ldots \vee \Phi_h) \equiv \neg(\Phi_{h+1} \vee \ldots \vee \Phi_{2^n}).$$

Формулу Ψ называют совершенной дизъюнктивной (конъюнктивной) нормальной формой формулы Ф, если Ψ есть СДНФ (СКНФ) и $\Phi \equiv \Psi$.

Нетрудно доказать, что для любой формулы Ф, не являющейся тождественно ложной, формула [[Ф]] есть совершенная дизъюнктивная нормальная форма формулы Ф (предполагая, что формулы пишутся по ячейкам, занятым точками).

В силу установленного соответствия между диаграммами Венна и формулами исчисления высказываний можно говорить о графическом (диаграммном) построении исчисления высказываний.

4.2.5. Вывод логических следствий с помощью диаграмм Венна

Формулу Ψ называют логическим следствием посылок $\Phi_1, \ldots, \Phi_m$ тогда и только тогда, когда

$$((\Phi_1 \& \ldots \& \Phi_m) \supset \Psi)$$

есть тождественно истинная формула.

Покажем, что формула Ψ будет логическим следствием посылок $\Phi_1, \ldots, \Phi_m$ тогда и только тогда, когда

$$[\Phi_1 \& \ldots \& \Phi_m]^n \subset [\Psi]^n,$$

где n — число графически неравных переменных, из которых составлена формула $(\Phi_1 \& \ldots \& \Phi_m)$.

1) Пусть Ψ есть логическое следствие формул $\Phi_1, \ldots, \Phi_m$. Предположим, что в ячейках $\alpha_1, \ldots, \alpha_k$ на диаграмме $[\Phi_1 \& \ldots \& \Phi_m]$ есть точки, а на [Ψ] — нет. Тогда на диаграмме

$$[(\Phi_1 \& \ldots \& \Phi_m) \supset \Psi]^n$$

ячейки $\alpha_1, \ldots, \alpha_k$ пусты, но $((\Phi_1 \& \ldots \& \Phi_m) \supset \Psi)$ есть тождественно истинная формула, т. е. формула, которой соответствует диаграмма, на которой нет пустых (не содержащих точки) ячеек. Мы пришли, таким образом, к противоречию.

Следовательно, если Ψ — логическое следствие посылок $\Phi_1, \ldots, \Phi_m$, то $[\Phi_1 \& \ldots \& \Phi_m] \subset [\Psi]$.

2) Пусть $[\Phi_1 \& \ldots \& \Phi_m] \subset [\Psi]$. Тогда $[\neg[\Phi_1 \& \ldots \& \Phi_m] \vee [\Psi]]$ есть тождественно истинная формула. Следовательно, поскольку формула $[\neg[\Phi_1 \& \ldots \& \Phi_m] \vee \Psi]]$ равносильна формуле

$((\Phi_1 \ \& \ldots \ \& \Phi_m) \supset \Psi)$, формула Ψ есть логическое следствие из посылок $\Phi_1, \ldots, \Phi_m$.

Таким образом, для получения логического следствия из посылок $\Phi_1, \ldots, \Phi_m$, составленного из δ переменных, где $\delta \geqslant n$, а n — число графически неравных переменных формулы $(\Phi_1 \ \& \ldots \ \& \Phi_m)$ надо построить диаграмму δ переменных $[\Phi_1 \ \& \ldots \ \& \Phi_m]^{\delta}$; далее, в некоторых из пустых ячеек диаграммы $[\Phi_1 \& \ldots \& \Phi_m]$ поставить по одной точке, потом написать формулу, соответствующую полученной диаграмме; эта формула и будет логическим следствием посылок $\Phi_1, \ldots, \Phi_m$.

Проставляя точки в различных пустых ячейках диаграммы

$$[\Phi_1 \& \ldots \& \Phi_m],$$

мы будем получать различные логические следствия из посылок $\Phi_1, \ldots, \Phi_m$.

Перебрав все возможные комбинации пустых ячеек диаграммы $[\Phi_1 \ \& \ldots \ \& \Phi_m]$ и проставив в них по одной точке, мы получим все возможные попарно неэквивалентные следствия из посылок $\Phi_1, \ldots, \Phi_m$, составленные из δ переменных.

4.2.6. Простые логические следствия

Логическое следствие Ψ из посылок $\Phi_1, \ldots, \Phi_m$ называют простым следствием из этих посылок, если Ψ есть дизъюнкция переменных (быть может, с отрицанием), не поглощаемая никаким другим логическим следствием того же вида из посылок $\Phi_1, \ldots, \Phi_m$.

Понятие поглощения некоторой формулы (другой) формулой становится ясным из следующих законов поглощения в исчислении высказываний:

$$(\Phi \& (\Phi \vee \Psi)) \equiv \Phi,$$
$$(\Phi \vee (\Phi \& \Psi)) \equiv \Phi,$$

где Φ, Ψ — любые формулы; о формуле Φ говорят, в первом случае, что она поглощает формулу $(\Phi \vee \Psi)$, во втором — что она поглощает формулу $(\Phi \ \& \ \Psi)$. Например, формула $(\neg a \vee b)$ поглощает формулу $(\neg a \vee b \vee c)$, но не поглощает формулу $(\neg a \ \vee c \vee d)$.

Конъюнкцию всех попарно различных простых логических следствий посылок $\Phi_1, \ldots, \Phi_m$ называют *силлогистическим многочленом* формулы $(\Phi_1 \& \ldots \& \Phi_m)$.

Один из способов получения силлогистического многочлена основан на применении законов поглощения и выявления.

Законы выявления в исчислении высказываний: $((\Phi \vee \vee \chi) \& (\Psi \vee \neg \chi)) \equiv ((\Phi \vee \chi) \& (\Psi \vee \neg \chi) \& (\Phi \vee \vee \Psi))$, $((\Phi \& \chi) \vee (\Psi \& \neg \chi)) \equiv ((\Phi \& \chi) \vee (\Psi \& \neg \chi) \vee \vee (\Phi \& \Psi))$, где Φ, Ψ, χ — любые формулы.

Силлогистический многочлен получается следующим образом:

1. Берется произвольная конъюнктивная нормальная форма χ формулы $(\Phi_1 \& \ldots \& \Phi_m)$.

2. В формуле χ производятся все возможные выявления и поглощения; формула, которая получится в результате выполнения описанной процедуры, и будет силлогистическим многочленом формулы $(\Phi_1 \& \ldots \ldots \& \Phi_m)$.

Можно показать, что независимо от формулы χ и порядка применения законов выявления и поглощения силлогистический многочлен формулы $(\Phi_1 \& \ldots \& \Phi_m)$ строится единственным образом (с точностью до перестановки членов).

Перейдем теперь к обзору всех простых логических следствий из посылок $\Phi_1, \ldots, \Phi_m$ с помощью диаграмм Венна. Предварительно разберем несколько предложений, из которых и будет вытекать общий метод нахождения простых логических следствий.

Пусть $a_1, \ldots, a_n$ — все графически неравные переменные, из которых составлена формула $(\Phi_1 \& \ldots \& \Phi_m)$.

1. Если на бинарной матрице $[\Phi_1 \& \ldots \& \Phi_m]^n$ находится только одна единица, то простыми логическими следствиями посылок $\Phi_1, \ldots, \Phi_m$ будут $\varepsilon_1 a_1, \ldots, \varepsilon_n a_n$, где ε_i есть пустое слово, если в элементарной последовательности γ, являющейся номером единицы, на i-ом месте стоит 1, и ε_i есть $\neg$, если на i-ом месте в последовательности находится 0 $(i = 1, \ldots, n)$; формула $(\varepsilon_1 a_1 \& \ldots \& \varepsilon_n a_n)$ будет силлогистическим многочленом формулы $(\Phi_1 \& \ldots \& \Phi_m)$.

Например, на бинарной матрице

$$\begin{pmatrix} 0\,0\,0\,0\,0\,0\,0\,1 \\ 0\,0\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0\,0\,0\,0 \\ 0\,0\,0\,0\,0\,0\,0\,0 \end{pmatrix} \qquad (2.7)$$

находится только одна единица, которая имеет номер 00111 (семь в двоичной системе), поэтому силлогистический многочлен формулы $(\Phi_1 \& \ldots \& \Phi_m)$, бинарной матрицей которой является (2.7), есть $(\neg a_1 \& \neg a_2 \& a_3 \& a_4 \& a_5)$.

2. Пусть бинарная матрица формулы $(\Phi_1 \& \ldots \& \Phi_m)$ содержит больше одной единицы. Пусть ψ — логическое следствие из посылок $\Phi_1, \ldots, \Phi_m$.

Если единицы расположены на бинарной матрице формулы Ψ так, что им соответствуют все элементарные последовательности, в которых на местах $i_1, \ldots, i_k$ находятся, соответственно, числа $\varepsilon_1, \ldots, \varepsilon_k$, то формула Ψ эквивалентна формуле

$$F = (\gamma_{i_1} a_{i_1} \vee \ldots \vee \gamma_{i_k} a_{i_k}),$$

где γ_{i_j} есть пустое слово, если $\varepsilon_j = 1$, и γ_{i_j} есть $\neg$, если $\varepsilon_j = 0$; в этом случае формула F, зависящая от k переменных, есть логическое следствие из посылок $\Phi_1, \ldots, \Phi_m$.

Например, на бинарной матрице

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 \end{pmatrix} \tag{2.8}$$

формулы Ψ двадцать четыре единицы, им соответствуют все элементарные последовательности пяти переменных, в которых на третьем месте находится нуль, а на пятом единица, поэтому $\Psi \equiv (\neg a_3 \vee a_5)$.

3. Пусть Φ и Ψ — формулы. Если $[\Phi] \subset [\Psi]$, то формула Ψ поглощается формулой Φ.

Например, формула, бинарной матрицей которой является (2.8), поглощается формулой, бинарной матрицей которой является (2.7).

4. Пусть для некоторого s конъюнкция диаграмм всех простых логических следствий $\Psi_1, \ldots, \Psi_r$, зависящих не более, чем от s переменных, равна диаграмме конъюнкции посылок $\Phi_1, \ldots, \Phi_m$

$$([\Psi_1] \& \ldots \& [\Psi_r]) = [\Phi_1 \& \ldots \& \Phi_m],$$

и ни для какого k (где $k = 1, \ldots, s-1$) не имеет места равенство

$$([\Psi_{j_1}] \& \ldots \& [\Psi_{j_k}]) = [\Phi_1 \& \ldots \& \Phi_m],$$

где $\Psi_{j_1}, \ldots, \Psi_{j_k}$ — все простые логические следствия посылок $\Phi_1, \ldots, \Phi_m$, зависящие не более, чем от k переменных. Тогда $(\Psi_1 \& \ldots \& \Psi_r)$ — силлогистический многочлен формулы $(\Phi_1 \& \ldots \& \Phi_m)$.

Предложения 1—4 легко могут быть доказаны.

Для примера рассмотрим предложение 3.

По условию: $[\Phi] \subset [\Psi]$. Поэтому

$$\Psi \equiv ([[\Phi]] \vee \chi),$$

где χ — формула, построенная по диаграмме, полученной из диаграммы $[\Psi]$ после стирания на ней точек, принадлежащихтакже и диаграмме $[\Phi]$, т. е. χ — есть формула

$[[\Psi] \& \neg [\Phi]]$.

$$(\Phi \& \Psi) \equiv ([[\Phi]] \& ([[\Phi]] \vee \chi)) \equiv [[\Phi]] \equiv \Phi.$$

Следовательно, формула Ф поглощает формулу Ψ, что и требовалось показать.

В качестве следствия предложения 2 можно получить условие исключения некоторых пропозициональных переменных из данной формулы.

Говорят, что пропозициональная переменная a_i исключается из формулы Ф, если существует такая формула Ψ, в которую переменная a_i не входит, и $\Phi \equiv \Psi$.

Пропозициональная переменная a_i может быть исключена (вместе с отрицанием) из формулы Ψ тогда и только тогда, когда во всех соседних относительно a_i строках (столбцах) бинарной матрицы $[\Psi]$ находятся попарно одинаковые числа. (Для сравнения см. операцию 1 увеличения числа переменных на диаграммах Венна в п. 4.2.2 и метод исключения неизвестных по Венну в первой главе.)

Например, из формулы Ψ, бинарной матрицей которой является (2.6), переменную b можно исключить; в результате получится формула Ф, равносильная формуле Ψ; матрица $[\Phi]$ есть (2.5).

Перейдем к алгорифму получения всех простых логических следствий из посылок $\Phi_1, \ldots, \Phi_m$.

Пусть $\Psi_1, \ldots, \Psi_{l_1}$ — все возможные различные логические следствия из $\Phi_1, \ldots, \Phi_m$, состоящие (каждое) из одной переменной. Если $([\Psi_1] \& \ldots \& [\Psi_{l_1}]) = [\Phi_1 \& \ldots \& \Phi_m]$, то $(\Psi_1 \& \ldots \& \Psi_{l_1})$— силлогистический многочлен формулы $(\Phi_1 \& \ldots \& \Phi_m)$. Если $([\Psi_1] \& \ldots \& [\Psi_{l_1}]) \neq [\Phi_1 \& \ldots \ldots \Phi_m]$, то возьмем $\Psi_{l_1+1}, \ldots, \Psi_{l_2}$— все возможные различные логические следствия из $\Phi_1, \ldots, \Phi_m$, состоящие из двух переменных, такие, что для любого i (где $i = 1, \ldots, l_1$) и для любого j (где $j = l_1 + 1, \ldots l_2$) неверно, что $[\Psi_i] \subset [\Psi_j]$. Если при этом $([\Psi_1] \& \ldots \& [\Psi_{l_2}]) = [\Phi_1 \& \ldots \& \Phi_m]$, то $(\Psi_1 \& \ldots \& \Psi_{l_2})$ — силлогистический многочлен формулы $(\Phi_1 \& \ldots \& \Phi_m)$. Если же $([\Psi_1] \& \ldots \& [\Psi_{l_2}]) \neq \Phi_1 \& \ldots \& \Phi_m]$, то берутся $\Psi_{l_2+1}, \ldots, \Psi_{l_3}$— все различные логические следствия из $\Phi_1, \ldots, \Phi_m$, состоящие из трех переменных, такие, что для любого i (где $i = 1, \ldots, l_2$) и для любого j (где $j = l_2 + 1, \ldots, l_3$) неверно, что $[\Psi_i] \subset [\Psi_j]$. Эта процедура продолжается, пока мы не найдем такое k, что $([\Psi_1] \& \ldots \& [\Psi_k]) = [\Phi_1 \& \ldots \& \Phi_m]$; тогда

$(\Psi_1 \& \ldots \& \Psi_k)$ — силлогистический многочлен формулы $(\Phi_1 \& \ldots \& \Phi_m)$.

Пример. Пусть диаграмма Венна D формулы $(\Phi_1 \& \ldots \ldots \& \Phi_m)$ имеет вид, показанный на рис. 51. Построим соответствующий силлогистический многочлен.

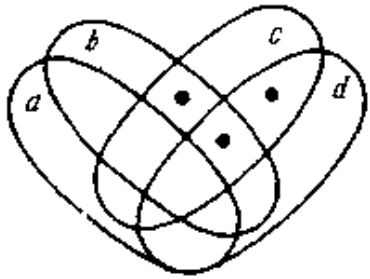


Рис. 51

Все точки диаграммы D принадлежат классам c и $\bar{a}$, каждый из которых состоит из одной переменной. Нетрудно проверить, что, кроме c и $\bar{a}$, на рассматриваемой диаграмме нет классов, составленных из одной переменной и содержащих все ячейки, в которых на диаграмме расположены точки.

Поэтому $\neg a$ и c суть все различные логические следствия посылок $\Phi_1, \ldots, \Phi_m$, состоящие из одной переменной.

Так как $([\neg a] \& [c]) \neq D$, то на диаграмме D выделим все такие различные классы, состоящие из двух переменных, которые содержат все ячейки диаграммы D с точками и в которые не входят ни класс $\bar{a}$, ни класс c. На диаграмме D видно, что такой класс один, он представляет собой объединение классов b и d.

$$([\neg a] \& [c] \& [b \vee d]) = D.$$

Следовательно, формула $(\neg a \& c \& (b \vee d))$ является требуемым силлогистическим многочленом.

В изложенном методе на каждом шаге находятся все простые логические следствия, состоящие из одного и того же числа переменных. На первом шаге находятся только простые логические следствия, являющиеся пропозициональными переменными или их отрицаниями, на втором — простые логические следствия, представляющие собой дизъюнкции переменных или их отрицаний, и т. д.

Одновременно решается и задача получения всех простых логических следствий, не содержащих некоторых пропозициональных переменных. Так, на первом шаге могут быть получены все простые логические следствия посылок $\Phi_1, \ldots, \Phi_m$, не содержащие *(n* — 1) пропозициональных переменных, где *n* — количество графически различных пропозициональных букв формулы $(\Phi_1 \& \ldots \& \Phi_m)$.
При нахождении простых логических следствий из данных посылок мы использовали силлогистический многочлен, являющийся некоторой конъюнктивной нормальной формой конъюнкции посылок. Задача решалась геометрически, члены силлогистического многочлена определялись с помощью построения соответствующих диаграмм Венна.
Нетрудно обосновать следующий способ написания совершенных конъюнктивных нормальных форм формул по заданным диаграммам Венна.
Рассматриваем пустые ячейки данной диаграммы Венна. Классу *x* ставим в соответствие отрицание — $\neg x$, дополнению класса *x* — пропозициональную переменную *x;* пересечению классов — дизъюнкцию, объединению — конъюнкцию; следовательно, по диаграмме Венна строится совершенная конъюнктивная нормальная формула. Легко доказать, что найденная формула эквивалентна совершенной дизъюнктивной нормальной формуле, которая определяется по ячейкам, заполненным точками. (Напомним, что Венн при рассмотрении пустых ячеек ставил в соответствие классам — переменные, дополнению — отрицание, пересечению — конъюнкцию, объединению — дизъюнкцию, т. е. Венн фактически исследовал формулы в их отрицательной форме, обращая основное внимание на случаи, когда значением формул является 0.)
Очевидно, что диаграммы Венна позволяют среди двух совершенных нормальных форм (конъюнктивной и дизъюнктивной) формул выбрать ту, которая имеет меньшую логическую длину. Кроме того, описанный метод построения силлогистического многочлена формулы Ф может быть использован при определении минимальных по логической длине (или по числу вхождений букв) конъюнктивных или дизъюнктивных нормальных форм формулы Ф.

4.2.7. Диаграмма Венна как оператор

Формула Ф исчисления высказываний называется избыточной, если существует эквивалентная ей формула Ψ, имеющая меньшее число вхождений пропозициональных переменных или логических знаков.

В литературе, относящейся к техническим приложениям математической логики, много внимания уделяется проблеме минимизации формул. При рассмотрении вопросов, связанных, например, с проблемой создания надежных устройств из относительно ненадежных элементов, существенную роль играют, наоборот, избыточные формулы. Избыточные формулы, как мы увидим ниже, удобно строить с помощью сетей диаграмм Венна. При этом понятие «диаграмма Венна» расширяется — диаграммы играют роль операторов.

Прежде чем переходить к описанию строения и функционирования сетей, покажем, как дерево построения формулы можно преобразовать в сеть диаграмм.

Как известно, каждой формуле исчисления высказываний, составленной из пропозициональных переменных $a_1, \ldots, a_n$, можно поставить в соответствие дерево ее построения. Например, формуле

$$\Phi = (((a_1 \,\&\, (\neg a_2 \,\&\, \neg a_3)) \vee (a_1 \,\&\, \neg (a_3 \supset a_2))) \vee a_2)$$

взаимно однозначно соответствует дерево, показанное на рис. 52.

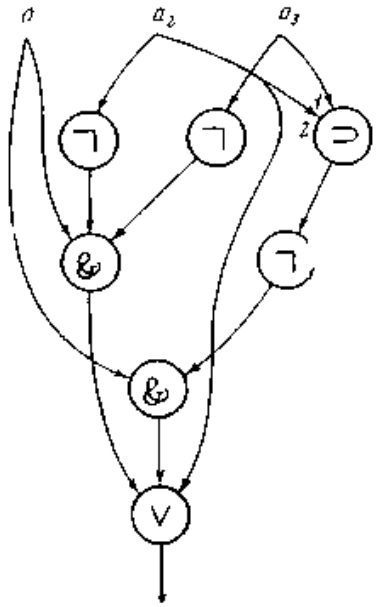


Рис. 52

На рисунке в вершинах расположены логические знаки формулы, кривые, оканчивающиеся в вершинах, нумеруются (когда порядок не существен, номера опускаются).

Дерево построения формулы можно преобразовать в сеть бинарных матриц (диаграмм Венна) следующим образом:

Первый ряд (ранг) образуем из матриц, соответствующих пропозициональным переменным, из которых составлена формула. (Построение таких матриц см. в п.1.4.2.3.)

Остальные ранги получаем заменой логических знаков в дереве построения формулы соответствующими бинарными матрицами; при этом число переменных матрицы совпадает с количеством кривых, оканчивающихся в рассматриваемой вершине.

а) Допустим, что сеть, соответствующая дереву формулы A, построена. Сеть, соответствующую дереву формулы $\neg A$, получим, заменяя в нем вершину  матрицей $(1\downarrow 0)$.

б) Допустим, что сети, соответствующие деревьям формул $A_1, \ldots, A_m$, построены. Сеть, соответствующую дереву формулы $(A_1\&, \ldots, \& A_m)$, получим, заменяя в нем вершину, например, при $m = 3$,  матрицей

$$\begin{pmatrix} \downarrow & \downarrow & \downarrow & \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

При построении сети, соответствующей дереву формулы $(A_1\vee, \ldots, \vee A_m)$, заменяем, например, при $m = 3$, вершину  матрицей

$$\begin{pmatrix} \downarrow & \downarrow & \downarrow & \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

При построении сети, соответствующей дереву формулы $(A_1 \supset A_2)$, заменяем вершину  матрицей

$$\begin{pmatrix} 1\downarrow & 2\downarrow & \\ 1 & 1 & 0 & 1 \end{pmatrix}.$$

Например, по дереву формулы Ф (рис. 52) строим сеть:

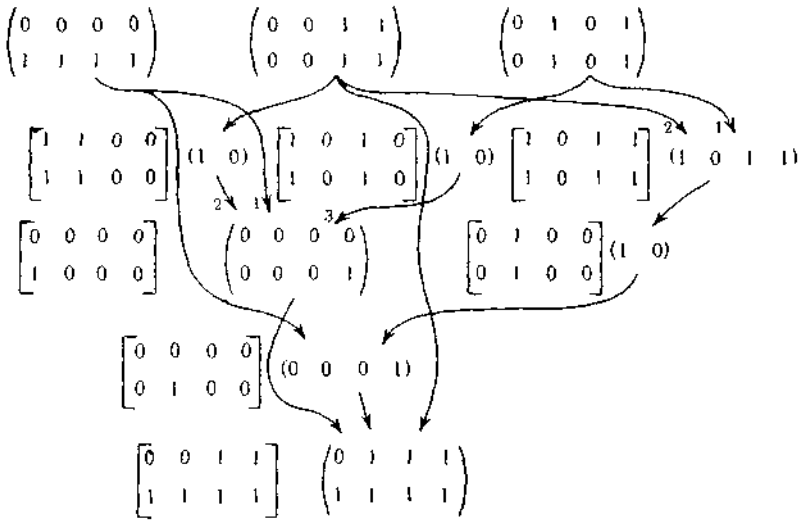


Рис. 53-57

Только что построенную сеть читатель может рассматривать в качестве примера; почему вершинам дерева построения формулы мы сопоставили такие матрицы, а не другие, и каков смысл матриц, стоящих в квадратных скобках, станет ясно после того, как мы опишем строение и функционирование сетей произвольного вида.

Описание строения и функционирования сетей диаграмм. Сеть состоит из нескольких рядов (рангов) диаграмм. Диаграммы первого ранга являются входными диаграммами, диаграммы последнего ранга — выходными (на рисунках ранги сети диаграмм нумеруются сверху вниз).

Каждая диаграмма сети диаграмм, начиная со второго ранга, связана кривыми с одной или несколькими диаграммами сети, на которых эти кривые начинаются. Кривые оканчиваются стрелками на диаграммах. Все кривые данной диаграммы нумеруются числами, которые ставятся слева от стрелки. Число кривых, оканчивающихся на данной диаграмме, равно числу различных переменных этой диаграммы. Каждую диаграмму сети будем рассматривать как некоторый оператор, функционирование которого ведет к построению диаграммы независимых (входных) переменных. Последнюю будем называть результирующей диаграммой данного оператора.

Обозначения: $D^{\tau}_{r,i}$ — диаграмма Венна, расположенная в r-ом ранге на i-ом месте (диаграммы в ранге нумеруются слева направо); $\lfloor D^{\tau}_{r,i} \rfloor$ — результирующая диаграмма оператора $D^{\tau}_{r,i}$.

Результирующие диаграммы операторов последнего ранга сети будем называть результирующими диаграммами сети. Их число совпадает с числом выходных диаграмм сети.

Одноранговые сети. Каждая диаграмма n переменных является одноранговой сетью с одним выходом. Результирующей диаграммой такой сети считаем данную диаграмму или результат операции увеличения числа переменных над данной диаграммой. В последнем случае результирующую диаграмму помещаем в квадратных скобках рядом с оператором.

Двухранговые сети с одним выходом. Первый ранг образуют k-диаграмм соответственно $n_1, \ldots, n_k$ переменных. Во втором ранге находится одна диаграмма n_{k+1} переменных, на которой оканчиваются стрелками n_{k+1} кривых, начинающихся на диаграммах первого ранга. Результирующие диаграммы первого ранга строятся аналогично уже рассмотренному случаю одноранговых сетей; пусть все они имеют одинаковое число переменных n,

$$n \geqslant \max\{n_1, \ldots, n_k\}.$$

Пусть $\beta_1, \ldots, \beta_h$ — все ячейки, в которых на $D^{\tau}_{2,1}$ находятся точки. Образуем диаграммы $D_1, \ldots, D_m$ где индексы $1, \ldots, m$ совпадают с номерами кривых, оканчивающихся на $D^{\tau}_{2,1}$ (здесь $m = n_{k+1}$): если i-ая кривая начинается на $D^{\tau}_{1,j}$ (где $j = 1, \ldots, k$), то положим $D_i \rightleftharpoons D^{\tau}_{1,j}$ (здесь $i = 1, \ldots, m$).

Пусть для определенности β_1 имеет номер (в двоичной системе) $1 \ldots 10 \ldots 0$, количество единиц которого равно i. Построим символ Венна n переменных. В ячейках, совпадающих по нумерации с ячейками, в которых на $D_1, \ldots, D_i$ находятся точки, а на $D_{i+1}, \ldots, D_m$ нет точек (и только таких ячейках), поставим по одной точке.

Аналогично поступаем с остальными ячейками β_l на $D^{\tau}_{2,1}$ $1 < l \leqslant h$; при этом занятые точками ячейки результирующей диаграммы (для $\beta_1, \ldots, \beta_{l-1}$) не рассматриваются, Получим результирующую диаграмму $\lfloor D_{2,1} \rfloor$, которую на рисунках помещаем в квадратных скобках около оператора.

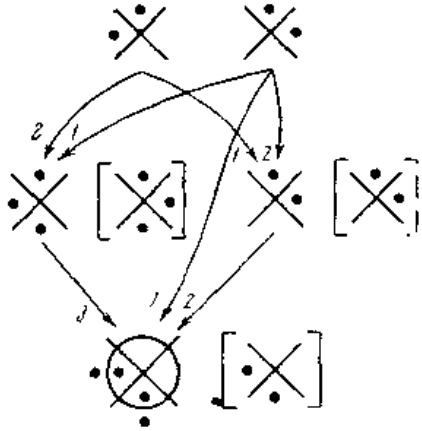


Рис. 58

Построение результирующих диаграмм многоранговых сетей ведется последовательно: на *i*-м шаге конструируются результирующие диаграммы *i*-го ранга; при этом диаграммы *i*-го ранга играют роль операторов над связанными с ними результирующими диаграммами в вышестоящих рангах. Результирующие диаграммы последнего ранга являются результирующими диаграммами сети.

Например, на рис. 53—57 в квадратных скобках расположены результирующие диаграммы соответствующих операторов.

Регулярные сети диаграмм Венна. Сеть диаграмм будем называть регулярной, если

1) диаграммы *i*-го ранга содержат одинаковое число переменных;

2) на каждой диаграмме *i*-го ранга может начинаться только одна кривая, оканчивающаяся на данной диаграмме (*i* + 1)-го ранга;

3) диаграммы (*i*+1)-го ранга соединяются кривыми только с диаграммами *i*-го ранга;

4) нумерация кривых, оканчивающихся на диаграмме (*i* + 1)-го ранга, совпадает с нумерацией диаграмм в *i*-м ранге (слева направо); *i*>1.

Ясно, что в случае регулярных сетей кривые на рисунках можно опускать.

Две сети будем называть эквивалентными, если их результирующие диаграммы совпадают.

Для любой сети диаграмм (без обратных связей) можно построить эквивалентную ей регулярную сеть. Построение осуществляется с помощью следующих двух операций:

1. *Изменение нумерации кривых оператора.* Пусть дана двухранговая сеть с одним выходом; $\beta_1, \ldots, \beta_n$— все ячейки, в

которых на $D^{\tau}_{2,1}$ находятся точки; 1, ..., m — номера кривых сети. Построим новую двухранговую сеть с одним выходом (операторы этой сети обозначим $\mathcal{D}_{ij}$), обладающую следующими свойствами: диаграммы первого ранга совпадают с результирующими диаграммами первого ранга исходной сети и имеют одинаковое число переменных; кривые сети по расположению также совпадают с кривыми исходной сети, но их номера переставляются — число l заменяется на l', где $1 \leqslant l' \leqslant m$; каждая ячейка β на $D^{\tau}_{2,1}$ имеет номер $\alpha^{\beta}_1, \ldots, \alpha^{\beta}_m$, где α^{β}_i есть 1 или 0 $(i = 1, \ldots, m)$, ячейке β на $D^{\tau}_{2,1}$ соответствует на $D^{\tau}_{2,1}$ ячейка δ с номером $\gamma^{\delta}_1, \ldots, \gamma^{\delta}_m$, где $\gamma^{\delta}_{j'}$ совпадает с α^{β}_j, $j' = 1, \ldots, m$; на $\mathcal{D}_{2,1}$ точки стоят только в ячейках $\delta_1, \ldots, \delta_k$, соответствующих ячейкам на $\beta_1, \ldots, \beta_k$ $D^{\tau}_{2,1}$.

Нетрудно убедиться, что полученная сеть эквивалентна исходной.

2. *Увеличение числа кривых оператора.* Пусть дана двухранговая сеть с одним выходом; $\beta_1, \ldots, \beta_k$ — все ячейками, в которых на $D^{\tau}_{2,1}$ находятся точки; 1, ..., m — номера кривых сети. Построим новую одновыходную двухранговую сеть, операторы этой сети обозначим $\overline{D}_{ij}$.

Диаграммами первого ранга будем считать результирующие диаграммы первого ранга исходной сети и какую-то произвольную диаграмму; это значит, что число диаграмм первого ранга новой сети равно k+1, где k — число диаграмм в первом ранге исходной сети. Новая диаграмма может быть расположена в любом месте первого ранга. Будем предполагать, что все диаграммы первого ранга сети имеют одинаковое число переменных $n, n = \max\{n_1, \ldots, n_k\}$, где n_i — число переменных на $D^{\tau}_{1,i}$, $i = 1, \ldots, k$. Сеть имеет m +1 кривую: m кривых по расположению и нумерации совпадают с кривыми исходной сети, (m+1)-я кривая начинается на новой (k + 1)-й диаграмме первого ранга. Оператором $\overline{D}_{2,1}$ будем считать диаграмму, являющуюся результатом увеличения числа переменных диаграммы $D^{\tau}_{2,1}$ на единицу; новая переменная соответствует (m + 1)-й кривой.

Нетрудно проверить, что в результате получаем сеть, эквивалентную исходной.

Например, используя операции 1 и 2, преобразуем сеть на рис. 58 в эквивалентную ей регулярную сеть (рис. 59).

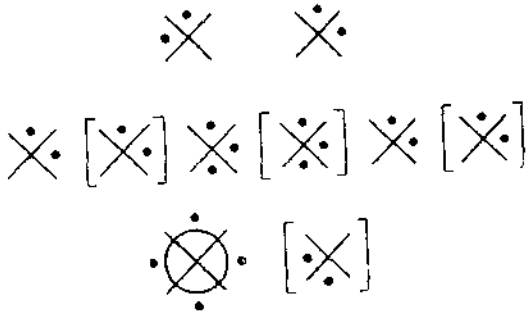


Рис. 59

Обратные связи в сетях диаграмм В е н н а. Рассмотрим сети, на каждой диаграмме которых могут оканчиваться кривые, связывающие эту диаграмму с любыми диаграммами сети (в том числе саму с собой). Иначе говоря, на каждой из диаграмм n_i переменных, начиная со второго ранга, $i > k_1$, стрелками оканчивается n_i кривых, соединяющих эту диаграмму с некоторыми диаграммами сети, и на каждой из диаграмм первого ранга n_j переменных, $1 \leqslant j \leqslant k_1$, могут также оканчиваться стрелками n_j кривых.

В таких сетях понятие ранга не имеет значения (оно наиболее существенно для регулярных сетей), но для удобства будем им пользоваться; будем называть, как и выше, диаграммы первого ранга входными, а диаграммы последнего ранга — выходными.

Будем говорить, что сеть является сетью с обратной связью, если в ней имеется по крайней мере одна диаграмма $D^{\tau}_{r,j}$, на которой находятся стрелки кривой, связывающейэту диаграмму с одной из диаграмм $(r+i)$-го ранга, $i \geq 0$.

Рассматривать работу сетей будем в равноотстоящие моменты времени. Такой временной подход необходим прежде всего для сетей с обратной связью; сети без обратных связей можно исследовать независимо от времени.

Построение результирующих диаграмм в равноотстоящие моменты времени происходит следующим образом.

Будем предполагать, что в момент времени t (начальный момент) строятся только результирующие диаграммы всех входных операторов (правила построения те же, что и для диаграмм первого ранга двухранговой сети); для остальных диаграмм их результирующие диаграммы пусты, т. е. не содержат точек. В $(t + i)$-й момент времени $(i>0)$ строятся только результирующие диграмм, на которых оканчиваются стрелки кривых; правила построения те же, что и для

диаграмм второго ранга двухранговой сети без обратной связи; для остальных диаграмм их результирующие пусты.

Пример. На рис. 60 — трехранговая сеть с обратной связью; вторая кривая диаграммы $D^{\tau}_{2,1}$ и первая кривая диаграммы $D^{\tau}_{2,2}$ начинаются на диаграмме $D^{\tau}_{3,1}$. На рис. 61—67 в квадратных скобках расположены результирующие диаграммы операторов сети в различные моменты времени.

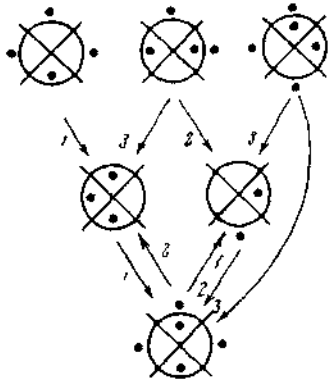


Рис. 60

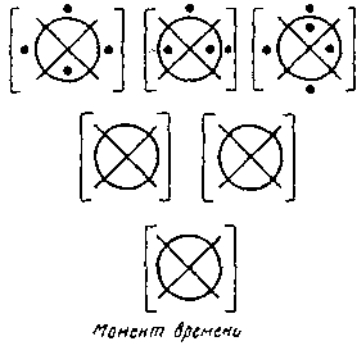


Рис 61

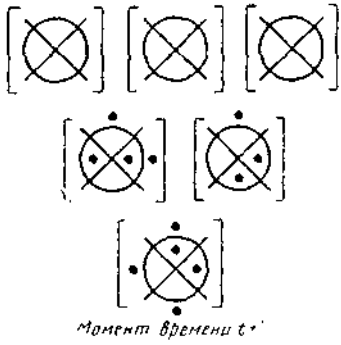


Рис. 62

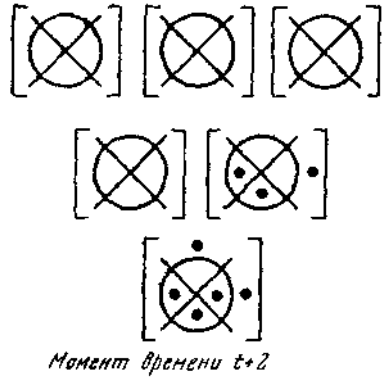


Рис. 63

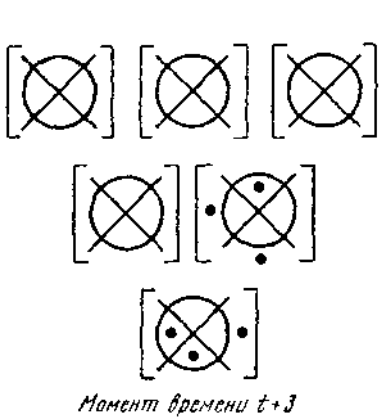


Рис. 64

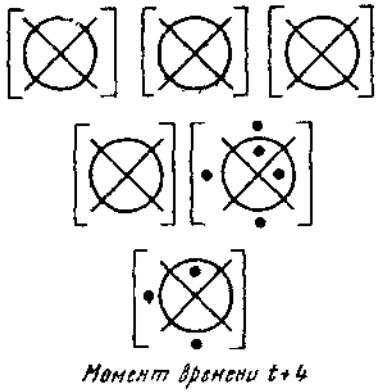
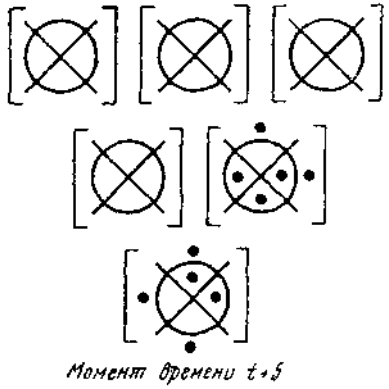


Рис. 65



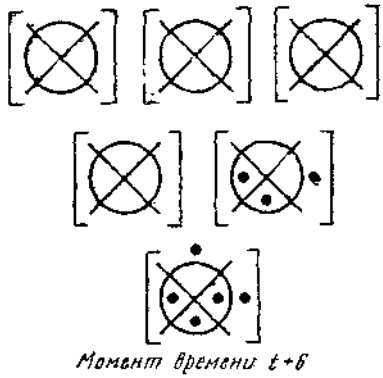


Рис. 66

Из чертежа видно, что результирующая диаграмма сети, т. е. результат работы $D^{\tau}_{3,1}$, не пуста, начиная с $(t + 1)$-го момента времени, и что результаты функционирования всех операторов сети периодически повторяются, начиная с $(t + 2)$-го момента времени,— результирующие диаграммы в момент $t + 2$ графически совпадают с результирующими диаграммами в момент $t + 6$.

Определение. Будем говорить, что сеть образует временной цикл длины μ, если результирующие диаграммы всех операторов сети периодически (с периодом μ) повторяются, начиная с некоторого момента времени; этот момент времени будем называть началом цикла.

В нашем примере момент времени t +2 является началом цикла, длина которого μ равна 4.

4.2.8. Вероятностные диаграммы

На сетях диаграмм Венна точки в некоторых ячейках операторов не оказывают влияния на соответствующие результирующие диаграммы. Аналогично можно заметить, что в некоторых других ячейках (пустых) операторов можно поставить точки, не изменяя результата работы (что, конечно, ведет к увеличению избыточности сети).

Каждой ячейке диаграммы Венна n переменных взаимно однозначно соответствует определенная последовательность из n единиц и нулей. Среди всех последовательностей, образуемых переменными оператора, можно выделить те, которые не влияют на результат работы этого оператора. Для выделения на операторе ячеек, точки в которых не изменяют результата функционирования оператора или только входных диаграмм сети, можно использовать так называемые вероятностные диаграммы.

Определение. Матрицу иа 2^k столбцов и 2^m строк

$$\begin{pmatrix} b_0 & b_1 & \dots & b_{2^k-1} \\ b_{2^k} & b_{2^k+1} & \dots & b_{2^{k+1}-1} \\ \dots & \dots & \dots & \\ b_{2^k(2^m-1)} & b_{2(2^m-1)+1} & \dots & b_{2^k+m-1} \end{pmatrix},$$

где $k \geqslant 0,\ m \geqslant 0,\ k + m \neq 0,$ будем называть вероятностной матрицей n переменных *(n=k+m),* если элементами матрицы являются буквы 0, *p,* 1, т. е. если $b_i \in \{0, p, 1\},\ i = 0, \dots, 2^n - 1.$

На языке диаграмм получаем, что вероятностная диаграмма n переменных есть символ Венна n переменных, во всех ячейках которого расположена одна из трех букв 1, *p,* 0.

В настоящем параграфе рассматривается случай, когда каждая из буквр вероятностной матрицы может принимать значение 1 или 0, независимо от остальных букв *p* на этой матрице.

Между вероятностными и бинарными матрицами n переменных можно установить соответствие: по вероятностной матрице n переменных A можно получить 2^l бинарных матриц, где l — количество букв p на матрице $A,$ заменяя каждую из букв p на 0 и 1, независимо от остальных букв *p.*

Например, вероятностной матрице

$$\begin{pmatrix} p & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & p & 0 & 1 \end{pmatrix}$$

соответствуют четыре бинарные матрицы

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

Ясно, что строки (столбцы) бинарных матриц можно воспринимать как числа, записанные в двоичной системе счисления. Тогда (для сокращения записи), переводя числа в десятичную систему счисления, любую бинарную матрицу можно представить в виде колонки (строки) десятичных чисел.

Например, последнюю из приведенных выше матриц можно записать в виде:

$$\begin{pmatrix} 10 \\ 9 \\ 12 \\ 5 \end{pmatrix}$$

или (14, 3, 8, 5).

Сеть вероятностных диаграмм является бинарной, если диаграммы сети не содержат букв *p*.

Бинарную диаграмму будем называть пустой, если на ней нет единиц.

Ясно, что каждой сети бинарных диаграмм взаимно однозначно соответствует сеть диаграмм Венна.

Из вероятностных диаграмм можно конструировать сети. Построение осуществляется аналогично случаю сетей диаграмм Венна. Единственное отличие от последних состоит в предположении, что все входные диаграммы, содержащие, по крайней мере, по одной букве *p,* имеют одинаковое число переменных — обозначим его через *n,*— а все остальные диаграммы первого ранга имеют число переменных, не превосходящее *n.* Для простоты будем предполагать далее, что все диаграммы первого ранга имеют одинаковое число переменных.

Как известно, основой надежности является избыточность. Ниже избыточность сетей вероятностных диаграмм будет использована при конструировании надежных устройств из мало надежных элементов.

Как обычно, решению задачи синтеза будет предшествовать анализ сетей вероятностных диаграмм — построение их результирующих диаграмм.

Способ 1: Перебор всех бинарных сетей, соответствующих данной сети вероятностных диаграмм. Пусть дана сеть *G* вероятностных диаграмм с одним выходом. Пусть *I*— число всех букв *p* на диаграммах сети *G*. Заменяя все буквы *p* единицами и нулями всеми возможными способами, получаем бинарные сети. Число всех бинарных сетей, соответствующих данной сети *G*, равно 2^I. Для каждой из 2^I сетей бинарных диаграмм можно построить соответствующую результирующую диаграмму *n* переменных, *n* — число различных переменных в первом ранге.

Предположим, что все 2^I результирующие диаграммы построены; обозначим их через $D_1, \ldots, D2^I$. Тогда результирующая вероятностная диаграмма исходной сети такова: возьмем символ Венна *n* переменных; если в ячейках $s_1, \ldots, s_k$ на каждой из диаграмм есть $D_1, \ldots, D2^I$ единицы, то в соответствующих ячейках $s_1, \ldots, s_k$ символа Венна поставим единицы; если ячейки $i_1, \ldots, i_r$ на каждой из диаграмм $D_1, \ldots, D2^I$ содержат нули, то в соответствующих ячейках символа Венна поставим нули; в каждой из остальных ячеек символа Венна поставим буквы *p*. В полученной результирующей диаграмме сети *G* буквы *p* расположены в тех и только тех ячейках, в которых по крайней мере на одной из диаграмм $D_1, \ldots, D_2^i$ находятся нули и в которых на некоторых из этих диаграмм обязательно есть единицы.

Если дана сеть *G* вероятностных диаграмм с несколькими выходами, то результирующие диаграммы для каждого выхода можно строить независимо от остальных выходов. Остановимся на функционировании сетей вероятностных диаграмм во времени, обращая особое внимание на сети с обратными связями.

Каждой сети вероятностных диаграмм *G* соответствует 2^I сетей диаграмм Венна, где *I*— число всех букв *p* на диаграммах сети *G*. Если *G* — сеть с несколькими выходами, то и каждая из 2^I сетей диаграмм Венна имеет столько же выходов.

В каждый данный момент времени $t + i$ (здесь $i > 0$) для каждой из 2^I сетей диаграмм Венна можно построить свою результирующую диаграмму (см. п.1.4.2.7). Перебирая все 2^I сетей диаграмм Венна в момент времени $t+ i$, можно построить результирующие вероятностные диаграммы исходной сети *G* в момент времени $t + i$ аналогично случаю сетей вероятностных диаграмм без обратных связей (не зависящих от времени).

Для каждой ячейки первого ранга буква в ячейке с тем же номером на результирующей вероятностной диаграмме определяется независимо от букв в остальных ячейках первого ранга, а каждая буква любого оператора, начиная со второго ранга, на бинарной сети функционирует независимо от остальных букв. Поэтому при построении результирующей вероятностной диаграммы можно ограничиться рассмотрением только бинарных сетей, которые получаются, когда все буквы *p* на каждой вероятностной диаграмме исходной сети принимают одинаковые значения.

Способ 2: Перебор элементарных последовательностей.

Пусть *G* — произвольная сеть вероятностных диаграмм, *n* — число различных переменных на каждой из диаграмм первого ранга.

Если *G* — сеть с обратной связью, то ее поведение исследуется в фиксированные моменты времени $t + i$, $i \geqslant 0$. Если *G* — сеть без обратных связей, то ее функционирование можно изучать как независимо от времени, так и в зависимости от момента времени,— сеть работает аналогично сетям с обратными связями.

При анализе сетей, зависящих от времени, основную роль играют связи между диаграммами; на чертежах эти связи выражаются с помощью кривых (иногда кривые можно опускать, тогда они подразумеваются). В момент времени $t + i$, $i \geqslant 0$ функционируют только те диаграммы сети, на которых оканчиваются стрелки кривых. Все остальные диаграммы сети будем рассматривать как пустые,— результатом их работы также являются пустые диаграммы. В момент времени *t* (начальный момент) функционируют только диаграммы первого ранга — входные диаграммы, а все остальные диаграммы считаются пустыми.

Таким образом, в каждый данный момент времени вместо сети *G* рассматривается другая сеть, отличающаяся от *G* только тем, что некоторые диаграммы сети *G* воспринимаются как пустые. Поэтому при изложении второго способа построения результирующих диаграмм мы не будем специально выделять случай, когда работа сети зависит от времени.

Каждой ячейке на диаграммах первого ранга соответствует определенная последовательность S_l из букв 1, *p,* 0, где *l* — номер ячейки $(l = 0, \ldots, 2^n - 1)$, $S_l = \delta_{1,l} \ldots \delta_{k,l}$, где *k*— число диаграмм в первом ранге, $\delta_{i,l} = 1$, 0 или *p* в зависимости от того, какая из этих букв (1, 0, *p)* стоит в *l*-й ячейке на *i*-й диаграмме первого ранга, $i = 1, \ldots, k$. Если среди $\delta_{1,l}, \ldots, \delta_{k,l}$ нет букв *p,* то S_l есть элементарная последовательность. Если среди $\delta_{1,l}, \ldots, \delta_{k,l}$ встречается *r* букв *p,* то придавая *p* значение 1 или 0, получим 2^r

различных элементарных последовательностей. Для каждой последовательности S_I в случае регулярных сетей можно проследить, переходя от диаграмм j-го ранга к диаграммам $(j+1)$-го ранга $(1 \leqslant j \leqslant g-1$, где g — число рангов в сети), какое значение 1,0, или p принимает данный выход сети G. Для сетей более сложных, чем регулярные, существенную роль играют связи между диаграммами.

Способ 3: Уменьшение числа рангов в регулярной сети вероятностных диаграмм. Первые два способа построения результирующих диаграмм сети связаны с перебором. Для регулярных сетей можно предложить еще один метод нахождения результирующих диаграмм. Он состоит в последовательном уменьшении числа рангов, для чего два последних ранга сети заменяются вероятностными диаграммами их работы. При этом каждую выходную диаграмму сети можно рассаматривать как оператор с определенными правилами функционирования. Уменьшая число рангов, мы получаем новую сеть, эквивалентную исходной.

Две сети G_1 и G_2 вероятностных диаграмм назовем эквивалентными, если все результирующие диаграммы сети G_1 и только они являются результирующими диаграммами сети G_2.

Построим вероятностные диаграммы работы сети, состоящей из двух рангов — последнего и предпоследнего — исходной сети G. Заменим найденными диаграммами два последних ранга сети G. Получим сеть G_1 вероятностных диаграмм, состоящую из $g-1$ рангов, где g — число рангов в сети G. Сеть G_1 эквивалентна сети G. Действуя таким образом, последовательно уменьшаем число рангов в сети G. В конце концов мы получим одноранговую сеть, диаграммы которой являются результирующими диаграммами сети G.

На каждом шаге описанной процедуры строятся результирующие диаграммы двухранговых сетей. Выходные диаграммы двухранговых сетей можно считать операторами, работающими над диаграммами первого ранга. Введение операторов позволяет полностью исключить перебор. Так как результирующие диаграммы для каждого оператора строятся независимо друг от друга, то мы ограничимся случаем одновыходных сетей. Будем также предполагать, что на выходной диаграмме содержится, по крайней мере, одна из букв 1, p (в противном случае результирующая диаграмма имеет, как и оператор, только нули, и тогда выходную диаграмму мы называем пустым оператором).

Правила работы выходного оператора двухранговой сети.

$П_0$: оператор N-переноса. Пусть выходная диаграмма $D^p_{2,1}$ сети G имеет только две ячейки, $n_{k+1} = 1$, $k = 1$, т. е. в первом ранге находится только одна диаграмма $D^p_{1,1}(n)$. Пусть на диаграмме

$D^p_{2,1}(1)$ в нулевой ячейке находится 0, в первой — 1; на диаграмме $D^p_{1,1}$ в ячейках $\beta_1, \ldots, \beta_l$ находятся единицы, в ячейках $\beta_{l+1}, \ldots, \beta_r$ — буквы *p,* в ячейках $\beta_{r+1}, \ldots, \beta_{2^n}$ — нули.

Построим символ Венна *n* переменных; в ячейки поставим буквы, графически равные буквам, лежащим в соответствующих ячейках диаграммы $D^p_{1,1}$. Получим вероятностную диаграмму $\lfloor D^p_{2,1} \rfloor$ — результат работы данной сети *G.* В этом случае диаграмму $D^p_{2,1}$ называем оператором *N*-переноса. В силу построения $\lfloor D^p_{2,1} \rfloor \doteq D^p_{1,1}(n)$.

Π_0^p: оператор N_p-переноса. Пусть предположения правила Π_0 сохраняются, за исключением того, что в первой ячейке на $D^p_{2,1}$ стоит не единица, а *p.* Диаграмму $D^p_{2,1}$ называем оператором N_p-переноса. В ячейках $\beta_1, \ldots, \beta_r$ нового символа Венна поставим букву *p,* в ячейках $\beta_{r+1}, \ldots, \beta_{2^n}$ — нули. Получим вероятностную диаграмму $\lfloor D^p_{2,1} \rfloor$.

Π_1: оператор *N*-отрицания. Пусть предположения правила Π_0 сохраняются, за исключением того, что на диаграмме $D^p_{2,1}$ в нулевой ячейке находится 1, а в первой — 0. Такую диаграмму $D^p_{2,1}$ будем называть оператором *N*-отрицания. Построим символ Венна *n* переменных, в ячейках $\beta_{r+1}, \ldots, \beta_{2^n}$ этого символа поставим 1, в ячейках $\beta_{l+1}, \ldots, \ldots, \beta_r$ — поставим *p,* в ячейках $\beta_1, \ldots \beta_l$, поставим 0. Получим вероятностную диаграмму $\lfloor D^p_{2,1} \rfloor$.

Π_1^p: оператор N_p-отрицания. Пусть предположения правила Π_1 сохраняются, только в нулевой ячейке на диаграмме $D^p_{2,1}$ стоит не единица, а буква *p.* В этом случае диаграмму $D^p_{2,1}$ называем оператором N_p-отрицания. В ячейках $\beta_{l+1}, \ldots, \beta_{2^n}$ символа Венна поставим *p,* в ячейках $\beta_1, \ldots, \ldots, \beta_l$ — нули. Получим вероятностную диаграмму $\lfloor D^p_{2,1} \rfloor$.

Π_2: оператор *N*-конъюнкции. Пусть рассматривается сеть *G,* выходная диаграмма которой имеет четыре ячейки $n_{k+1} = 2$, $k = 2$, т. е. в первом ранге сети *G* расположены две диаграммы: $D^p_{1,1}(n)$ и $D^p_{1,2}(n)$. Пусть на $D^p_{2,1}(2)$ в третьей ячейке стоит 1, а в остальных ячейках — 0; в ячейках $\beta_1, \ldots, \beta_l$ диаграмм $D^p_{1,1}$ и $D^p_{1,2}$ находятся единицы; в ячейках $\beta_{l+1}, \ldots, \beta_s$ диаграмм $D^p_{1,1}$ и $D^p_{1,2}$ находится *p,* в ячейках $\beta_{s+1}, \ldots, \beta_m$ — 1 на диаграмме $D^p_{1,1}$ и буквы *p* на диаграмме $D^p_{1,2}$, в ячейках $\beta_{m+1}, \ldots, \beta_q$ — p на $D^p_{1,1}$ и 1 на $D^p_{1,2}$, во всех остальных ячейках стоит нуль, по крайней мере, на одной из диаграмм $D^p_{1,1}$ или $D^p_{1,2}$. Диаграмму $D^p_{2,1}$ будем называть оператором *N*-конъюнкции.

Построим символ Венна *n* переменных, в ячейках $\beta_1, \ldots, \beta_l$ которого поставим 1, в ячейках $\beta_{l+1}, \ldots, \beta_q$ поставим *p,* а во всех остальных ячейках — 0. Получим вероятностную диаграмму $\lfloor D^p_{2,1} \rfloor$ — результат работы оператора *N*-конъюнкции.

Π_2^p: оператор N_p-конъюнкции. Пусть предположения правила Π_2 сохраняются, за тем исключением, что в третьей ячейке диаграммы $D^p_{2,1}$ стоит не единица, а *p.* В ячейках $\beta_1, \ldots, \beta_q$ построенного символа Венна поставим *p,* во всех остальных — 0. Получим результирующую вероятностную диаграмму $\lfloor D^p_{2,1} \rfloor$. В этом случае $D^p_{2,1}$ называем оператором N_p-конъюнкции.

Π_3. Рассмотрим случай, когда на выходной диаграмме $D^p_{2,1}$ находится только одна отличная от нуля буква — единица. Предположим, что диаграмма $D^p_{2,1}$ отличается от выходных диаграмм в правилах Π_0, Π_1, Π_2. Пусть эта единственная единица на $D^p_{2,1}(m)$ лежит в ячейке $\Delta \overline{\cdot} z_1 \ldots z_m$, где $d_1, \ldots, d_m$ — все графически различные переменные диаграммы $D^p_{2,1}$, $z_j \overline{\cdot} \tilde{d}_j$,

$$z_j \overline{\cdot} \begin{cases} d_j, & \text{если } \Delta \text{ принадлежит фигуре } d_j, \\ \bar{d}_j, & \text{если } \Delta \text{ не принадлежит фигуре } d_j, \end{cases} j = 1, \ldots, m.$$

Для определенности пусть $\Delta \overline{\cdot} \bar{d}_1 \ldots \bar{d}_i d_{i+1} \ldots d_m$ $(0 \leqslant i \leqslant m,\ i = 0$ означает, что $\Delta \overline{\cdot} d_1 \ldots d_m$, $i = m$ означает, что $\Delta \overline{\cdot} \bar{d}_1 \ldots \bar{d}_m)$.

Построим *N*-отрицания диаграмм $D^p_{1,1}(n), \ldots, D^p_{1,i}(n)$ (по правилу Π_1). К полученным диаграммам $\lfloor D^p_{1,1} \rfloor(n), \ldots, \lfloor D^p_{1,i} \rfloor(n)$ и к диаграммам $D^p_{1,i+1}, \ldots, D^p_{1,m}$ применим последовательно правило Π_2: от диаграмм $\lfloor D^p_{1,1} \rfloor$ и $\lfloor D^p_{1,2} \rfloor$ перейдем к диаграмме D^p_1, от диаграмм D^p_1 и $\lfloor D^p_{1,3} \rfloor$ — к диаграмме D^p_2 и так далее; от диаграмм D^p_{i-1} и $\lfloor D^p_{1,i+1} \rfloor$ — к диаграмме D^p_i, и так далее; наконец, от D^p_{m-2} и $D^p_{1,m}$ — к диаграмме D^p_{m-1}, являющейся результирующей диаграммой работы исходной сети.

Правило Π_3 можно сформулировать также следующие образом:

Перенесем на символ Венна результирующей диаграммы, во-первых, те и только те единицы, для которых соответствующие ячейки на всех диаграммах $D^p_{1,1}, \ldots, D^p_{1,i}$ пусты, а ячейки на всех диаграммах $D^p_{1,i+1}, \ldots, D^p_{1,m}$ заполнены единицами, во-вторых, те и только те нули, для которых в соответствующих ячейках по крайней мере на одной из диаграмм $D^p_{1,1}, \ldots, D^p_{1,i}$ стоит единица, или по крайней мере на одной из диаграмм $D^p_{1,i+1}, \ldots, D^p_{1,m}$ — нуль. Во всех остальных ячейках результирующей диаграммы (не заполненных нулями и единицами) расположим по одной букве *p.*

Правило Π_3 является обобщением правил Π_0, Π_1 и Π_2. Поэтому в дальнейшем предполагаем, что Π_0, Π_1 и Π_2 — частные случаи правила Π_3.

Π_3^p. Правило Π_3^p отличается от Π_3 только тем, что на диаграмме $D^p_{2,1}$ единица заменена на p, а вместо правила Π_2 применяется правило Π_2^p. Правило формулируется так:

Перенести на символ Венна результирующей диаграммы те и только те нули, для которых в соответствующих ячейках, по крайней мере на одной из диаграмм $D^p_{1,1}, \ldots, D^p_{1,i}$, находится единица или на одной из диаграмм $D^p_{1,i+1}, \ldots, D^p_{1,m}$ — нуль; во всех остальных ячейках поставить буквы p. Аналогично правилу Π_3, правило Π_3^p является обобщением правил Π_0^p, Π_1^p и Π_2^p.

Π_4: оператор N-дизъюнкции. Пусть на $D^p_{2,1}(m)$ находится s, $s>1$, отличных от нуля букв $1, p$: $\Delta_1, \ldots, \Delta_r, \ldots, \Delta_s$, где $\Delta_i = 1$, $i = 1, \ldots, r$ (где $r > 0$); $\Delta_j = p$, $j = r+1, \ldots, s$ (где $s \geq r$).

По правилу Π_3 для Δ_1, если $\Delta_1 = 1$, можно построить вероятностную диаграмму; не обращая внимания на буквы $\Delta_2, \ldots, \Delta_s$, сотрем на ней все нули; получим диаграмму C_1. По правилу Π_3 для Δ_2, при $\Delta_2 = 1$, построим вероятностную диаграмму; не обращая внимания на буквы $\Delta_1, \Delta_3, \ldots, \Delta_s$, перенесем с нее все знаки 1, p в соответствующие ячейки диаграммы C_1; получим диаграмму C_2, в некоторых ячейках которой могут содержаться две буквы p. И так далее.

По правилу Π_3 для Δ_r, при $\Delta_r = 1$, построим вероятностную диаграмму, не обращая внимания на буквы $\Delta_1, \ldots, \Delta_{r-1}, \Delta_{r+1}, \ldots, \Delta_s$; перенесем с нее все знаки 1, p в соответствующие ячейки диаграммы C_{r-1}; получим диаграмму C_r.

По диаграмме C_r построим вероятностную диаграмму $D^p_r(n)$. В ячейках β_j этой диаграммы, соответствующих пустым ячейкам диаграммы C_r, поставим 0; в ячейке β_i поставим 1, если 1 находится в ячейке β_i диаграммы C_r, или, если в m-членной последовательности из знаков 1, 0, p, соответствующей ячейке β_i вероятностных диаграмм первого ранга, находится только t букв p, а в ячейке β_i диаграммы C_r находится 2^t букв p. Пусть $\beta_1, \ldots, \beta_u$ — все ячейки, в которых поставлена одна из букв 1, 0. Во всех остальных ячейках поставим по одной букве p. Получим диаграмму $D^p_r(n)$. Если $r = 0$, то будем считать, что на $D^p_r(n)$ находятся только нули.

По правилу Π_3^p для Δ_{r+1}, при $\Delta_{r+1} = p$, построим вероятностную диаграмму, не обращая внимания на буквы $\Delta_1, \ldots, \Delta_r, \Delta_{r+2}, \ldots, \Delta_s$; перенесем с нее знаки p в соответствующие ячейки диаграммы $D^p_r(n)$,

не занятые 1, *p* (заменяя 0 на *p*); получим диаграмму $D^p_{r+1}(n)$. И так далее.

По правилу $П_3^p$ для Δ_s, при $\Delta_{s-1} \neq p$, построим вероятностную диаграмму, не обращая внимания на буквы $\Delta_1, \ldots, \Delta_{s-1}$; перенесем с нее знаки *p* в соответствующие ячейки диаграммы $D^p_{s-1}(n)$, не занятые 1, *p* (заменяя 0 на *p);* получим диаграмму $D^p_s(n)$ — результат работы исходной сети. В этом случае диаграмму $D^p_{2,1}(m)$ называем оператором *N*-дизъюнкции, а диаграмму $D^p_s(n)$ — результатом функционирования оператора *N*-дизъюнкции.

При построении вероятностной диаграммы работы регулярной сети третий способ предпочтительнее, потому что выходная диаграмма двухранговых сетей есть оператор с указанными правилами работы. У многоранговых регулярных сетей последовательное уменьшение рангов можно вести снизу вверх, заменяя два последних ранга сети вероятностной диаграммой их работы, но нельзя вести сверху вниз,— нельзя заменять первые два ранга сети результатами работы соответствующих двухранговых сетей.

Рассмотренные правила работы выходных операторов двухранговых сетей связаны с операциями исчисления высказываний не только по названию. Если ограничиться бинарными сетями, то оператору *N*-отрицания соответствует операция отрицания, оператору *N*-конъюнкции — операция конъюнкции, а результату работы оператора *N*-дизъюнкции соответствует s-кратное применение операции дизъюнкции (*s* — количество единиц оператора *N*-дизъюнкции).

Определения. Сеть вероятностных диаграмм называется надежной, если результирующие диаграммы этой сети не содержат букв *p*.

Сеть называется не вполне надежной, если среди ее результирующих диаграмм имеется по крайней мере одна, содержащая буквы *p*.

4.2.9. Надежные сети вероятностных диаграмм

На языке вероятностных диаграмм задача построения надежных сетей из не вполне надежных элементов ставится следующим образом.

Даны бинарные диаграммы *n* переменных

$$D^p_1(n), \ldots, D^p_k(n), \quad k \geqslant 1.$$

Требуется построить надежную сеть *G* вероятностных диаграмм, состоящую из *g* рангов, $g > 1$, и имеющую в *r*-м ранге n_r вероятностных диаграмм, содержащих максимально возможное число букв $p, r = 1, \ldots, g$, так, чтобы результирующими диаграммами сети *G* являлись только данные бинарные диаграммы

$$D_1^p(n), \ldots, D_h^p(n).$$

Предложим способ синтеза двухранговых надежных сетей вероятностных диаграмм с одним выходом. Сети будем строить так, чтобы диаграммы первого ранга имели в среднем одинаковое число букв *p*. Для простоты ограничимся рассмотрением регулярных сетей.

Пусть $\alpha_1, \ldots, \alpha_l$ — номера ячеек данной бинарной диаграммы, в которых расположены единицы.

1. Начертим *m* символов Венна *n* переменных: $M_1, \ldots\ldots, M_m$. В ячейках с номерами $\alpha_1, \ldots, \alpha_l$ каждого символа поставим единицы. Расположим $2^n - l$ нулей так, чтобы для любого k, $k \in \{0, \ldots, 2^n - 1\}$, $k \neq \alpha_1, \ldots, \alpha_l$, существовал символ M_i, $i = 1, \ldots, m$ в k-й ячейке которого находится нуль, и чтобы на каждом M_i осталось в среднем одинаковое число пустых ячеек, в которых поставим по одной букве *p*. Построенные вероятностные диаграммы обозначим $\mathfrak{A}_{1,1}, \ldots, \mathfrak{A}_{1,m}$. Образуем TV-отрицания диаграмм

$\mathfrak{A}_{1,i}$, $i = 1, \ldots, m$: $\neg\mathfrak{A}_{1,i}$, где $\neg$ — знак N-отрицания.

2. Из каждой пары $\mathfrak{A}_{1,i}$ и $\neg\mathfrak{A}_{1,i}$ выделим в первый рангсети по одной диаграмме, которую обозначим $\mathfrak{B}_{1,i}$, $i = 1, \ldots, m$.

3. Начертим символ Венна M_{m+1} m переменных. Каждой ячейке $\beta_1 \ldots \beta_m$, где β_i есть 0 или 1, символа M_{m+1} поставим в соответствие диаграмму

$$\mathfrak{C}_{\beta_1 \ldots \beta_m} \overline{\circ} (\Delta_1 \mathfrak{B}_{1,1} \& \ldots \& \Delta_m \mathfrak{B}_{1,m}),$$

где $\Delta_j \overline{\cdot} A$, если $\beta_j = 1$, $\Delta_j \overline{\cdot} \neg$, если $\beta_j = 0$, $\&$ — знак N-конъюнкции, $\beta_1 \ldots \beta_m$ — номер ячейки на M_{m+1} в двоичной системе.

В ячейке $\beta_1^0 \ldots \beta_m^0$, где $\beta_j^0 = 1$, если $\mathfrak{B}_{1,j} \overline{\cdot} \mathfrak{A}_{1,j}$ и $\beta_j^0 = 0$, если $\mathfrak{B}_{1,j} \overline{\cdot} \neg\mathfrak{A}_{1,j}$, поставим единицу, предполагая, что $l > 1$. Тогда $\mathfrak{C}_{\beta_1^0 \ldots \beta_m^0} \overline{\cdot} \mathfrak{C}$, где $\mathfrak{C}$ — данная бинарная диаграмма, и для любой ячейки с номером $\beta_1, \ldots, \beta_m$, отличном от $\beta_1^0 \ldots \beta_m^0$, диаграмма $\mathfrak{C}_{\beta_1 \ldots \beta_m}$ не совпадает с $\mathfrak{C}$.

Если $l = 0$ (в этом случае данная бинарная диаграмма $\mathfrak{C}$ соответствует тождественно ложной формуле исчисленияв высказываний), то в ячейке $\beta_1^0 \ldots \beta_m^0$ на M_{m+1} поставим

$$p, \mathfrak{C}_{\beta_1^0 \ldots \beta_m^0} \equiv 0.$$

4. Среди ячеек на M_{m+1}, отличных от $\beta_1^0 \ldots \beta_m^0$, выделим такие, для которых $\mathfrak{C}_{\beta_1 \ldots \beta_m} \equiv 0$. В этих ячейках на M_{m+1} поставим по одной букве *p*. Во всех остальных ячейках M_{m+1} поставим нули, получим сеть, каждая вероятностная диаграмма первого ранга которой имеет μ букв *p,*

$$\mu = 2^n - \frac{2^n - l}{m} - l, \text{ если } \frac{2^n - l}{m} \text{ — целое, } \mu = 2^n - \left[\frac{2^n - l}{m}\right] - l$$

или $\mu = 2^n - \left[\frac{2^n - l}{m}\right] - l - 1$, если $\frac{2^n - l}{m}$ не является целым $\left(\left[\frac{2^n - l}{m}\right]\right.$ — целая часть числа $\left.\frac{2^n - l}{m}\right)$; при этом в любой другой двухранговой сети (с одним выходом и с одной единицей на выходном операторе), результирующей диаграммой которой является диаграмма $\mathfrak{C}$, существует диаграмма первого ранга, количество букв *p* на которой не превосходит $2^n - \left[\frac{2^n - l}{m}\right] - l$.

5. Каждой ячейке диаграмм первого ранга построенной сети соответствует определенная *m*-членная последовательность из букв 1, 0, *p*. Такую последовательность будем называть N_i-последовательностью, где *i* — номер соответствующей ячейки. В каждой N_i -последовательности имеется не более *m* — 1 букв *p*. Все N_i -последовательности делятся на группы, состоящие из одинаковых последовательностей. Всего таких групп не более *m* +1.

Если в ячейке с номером $\beta_1^{00} \ldots \beta_m^{00}$ (в двоичной системе) символа M_{m+1} заменим 0 на *p*, то для того, чтобы диаграмма $\mathfrak{C}_{\beta_1^{00} \ldots \beta_m^{00}}$ была пустой (т. е. не влияла на результат работы сети), в каждой N_i-последовательности по крайней мере на одном из мест $j_1, \ldots, j_r$ $(\beta_{j_1}^{00} = \ldots = \beta_{j_r}^{00} = 1)$ или на одном из мест $s_1, \ldots, s_u$ $(\beta_{s_1}^{00} = \ldots = \beta_{s_u}^{00} = 0)$ поставим соответственно нуль или единицу. Таким образом, при увеличении числа букв *p* на M_{m+1} количество букв *p* на диаграммах первого ранга может убывать. Наибольшее число букв *p* на M_{m+1} не превосходит 2^m — 1, число 2^m — 1 может быть получено, например, когда во всех ячейках данной бинарной диаграммы $\mathfrak{C}$ находятся единицы.

Например, двухранговые надежные сети из не вполне надежных элементов (рис. 68—70), построенные этим способом, реализуют, соответственно, формулы

$$(\bar{a}_1\,\bar{a}_2\,a_3 \vee \bar{a}_1\,a_2\,\bar{a}_3),$$
$$(a_1\,\bar{a}_2\,a_3 \vee a_1\,a_2\,\bar{a}_3),$$
$$(\bar{a}_1\,\bar{a}_2\,\bar{a}_3 \vee \bar{a}_1\,a_2\,a_3 \vee a_1\bar{a}_2\,a_3).$$

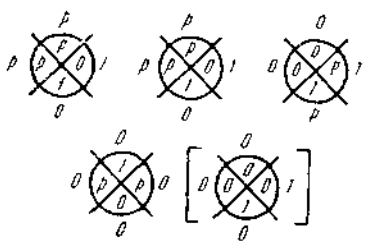


Рис. 68

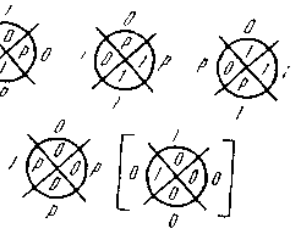


Рис. 69

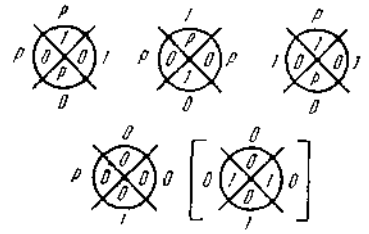


Рис. 70

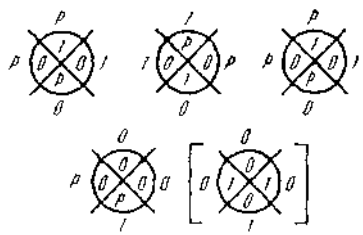


Рис. 71

Количество букв на входном операторе рис. 70 можно увеличить, не изменяя результата работы сети (рис. 71). Нетрудно показать, что невозможно составить надежную сеть, у которой (1) выходная диаграмма имеет две переменные, (2) все операторы содержат не менее одной буквы *p* и (3) формула результирующей диаграммы не является ни тождественно ложной, ни тождественно истинной.

Описанный способ является обобщением способа, предложенного Мак-Каллоком. Использование свойств оператора *N*-дизъюнкции позволяет увеличить число единиц на результирующей диаграмме. Так, на рис. 72 синтезирована надежная сеть из не вполне надежных элементов, реализующая СДНФ

$$(\bar{a}_1\bar{a}_2a_3 \vee \bar{a}_1a_2\bar{a}_3 \vee \bar{a}_1a_2a_3 \vee a_1\bar{a}_2\bar{a}_3 \vee a_1\bar{a}_2a_3),$$

или эквивалентную ей СКНФ

$$((a_1 \vee a_2 \vee a_3) \,\&\, (\neg a_1 \vee \neg a_2 \vee a_3) \,\&\, (\neg a_1 \vee \neg a_2 \vee \neg a_3).$$

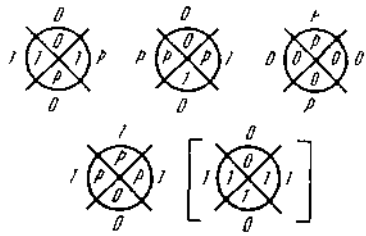


Рис. 72

Кроме того, можно также конструировать надежные сети из не вполне надежных элементов, имеющие боле двух рангов (рис. 73) или несколько выходов (рис. 74).

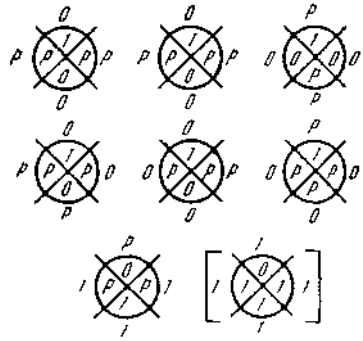


Рис. 73

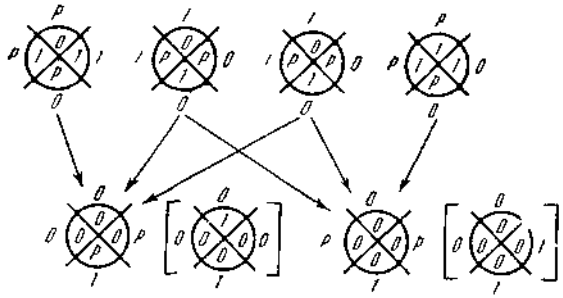


Рис. 74

4.2.10. Вероятностные диаграммы (продолжение)

Определяя выше (п.1.4.2.8) вероятностную диаграмму (матрицу) и изучая сети таких диаграмм (матриц), мы нигде по существу не пользовались понятием вероятности. Буква p играла роль неинтерпретируемого символа, о котором говорилось только, что он может принимать значения 0 и 1. Между тем название для этого класса диаграмм выбрано не случайно, и диаграмме действительно можно придать вероятностный смысл. Первые шаги в этом направлении были сделаны Мак-Каллоком. Оказывается, что при этом более глубоко проясняются идеи, лежащие в основе алгоритмов построения надежных сетей.

Изложение начнем с разбора соответствующего аналитического аппарата.

Случайной булевой функцией F n аргументов назовем случайную функцию, определенную для всевозможных наборов значений аргументов, принимающих два значения 0 и 1, значениями которой являются случайные величины, могущие принимать тоже только два значения 0 и 1.

Как и выше, будем рассматривать наборы n нулей и единиц как двоичные записи чисел от 0 до 2^n — 1. Тогда на наборе, которому соответствует число $i, 0 \leqslant i \leqslant 2^n - 1$, значением функции F окажется случайная величина A_i, принимающая значение 1 с вероятностью p_i, и значение 0 с вероятностью 1 — p_i.

Для случайной булевой функции F рассмотрим функцию Ф n переменных, определенную на тех же наборах, что и F, и равную p_i на наборе, которому соответствует число i. Такую функцию Ф назовем вероятностной функцией случайной функции F.

Очевидно, что задание вероятностной функции вполне определяет случайную булеву функцию. Поэтому в дальнейшем будем иметь дело исключительно с вероятностными функциями. Независимое от понятия случайной булевой функции определение вероятностной функции Ф состоит в том, что это — функция n переменных, определенная на наборах нулей и единиц, значения которой удовлетворяют условию $0 \leqslant \Phi(x_1, \ldots, x_n) \leqslant 1$. В частности, все булевы функции входят в класс вероятностных функций.

Наглядный смысл вероятностных функций состоит в следующем. Пусть имеется некоторое техническое устройство с n входами и с одним выходом, долженствующее вычислять некоторую определенную булеву функцию n переменных (булеву функцию n переменных можно представлять в виде диаграммы Венна n переменных). Идеально функционирующее устройство в ответ на любую комбинацию входных

воздействий выдает вполне определенный выходной сигнал, обозначаемый 0 или 1. Предположим, однако, что устройство функционирует не идеально. В таком случае для любой комбинации входных воздействий определена лишь вероятность того, что на выходе будет 1 (в частности, она может превращаться в уверенность: $p = 1$ или $p — 0$).

Руководствуясь этой наглядной картиной и некоторыми элементарными теоретико-вероятностными соображениями, определим над вероятностными функциями ряд операций.

Во-первых, предположим, что выход устройства пропускается через инвертор, который, со своей стороны, работает вполне надежно. Очевидно, что, если на выходе устройства с вероятностью p была единица, то на выходе инвертора с такой же вероятностью будет нуль, а единица на выходе инвертора будет с вероятностью $1 — p$.

Это дает основание для следующего определения операции отрицания.

Пусть F — вероятностная функция. Операция отрицания дает новую вероятностную функцию, которую обозначим $\bar{F}$, причем $p_i^{\bar{F}} \rightleftharpoons 1 — p_i^F$ [выражение «p_i^{Φ}» обозначает значение функций Ф на наборе с номером i, предполагаем (как и выше), что наборы занумерованы теми числами i, двоичными изображениями которых эти наборы являются].

Во-вторых, пусть выходы двух устройств подведены к прибору, осуществляющему конъюнкцию. Если вероятность появления 1 на выходе первого устройства есть p, а вероятность появления 1 на выходе второго устройства есть q, то вероятность получить 1 на выходе прибора равна pq (в предположении, естественно, независимости событий).

Если же на прибор, реализующий конъюнкцию, дважды подается выход одного и того же устройства, то вероятность получить на выходе прибора 1 будет, очевидно, p (а не p^2). Введем теперь операцию конъюнкции.

Операция конъюнкции, примененная к различным вероятностным функциям F и Ф, дает функцию $(F \,\&\, \Phi)$ такую, что

$$p_i^{(F\&\Phi)} \rightleftharpoons p_i^F \cdot p_i^{\Phi}. \quad p_i^{(\Phi\&\Phi)} \rightleftharpoons p_i^{\Phi}.$$

Таким образом, при применении операции конъюнкции к двум функциям их значения на одном и том же наборе подвергаются операции квазиумножения, отличающегося от обычного умножения действительных чисел тем, что $p_i^{\Phi} \cdot p_i^{\Phi} = p_i^{\Phi}$. Поэтому мы помечаем

вероятности p_i индексом вверху, указывающим соответствующую функцию.

В-третьих, пусть выходы двух устройств подведены к прибору, осуществляющему дизъюнкцию. Вероятность того, что на выходе прибора появится 1, равна, очевидно,

$$pq + p(1 - q) + (1 - p)q = pq + p - pq + \\ + q - pq = p + q - pq.$$

Если мы имеем дело с одним устройством, то, как и для конъюнкции, вероятность появления единицы будет просто p. Легко видеть, что мы получим это автоматически, если будем понимать умножение как квазиумножение:

$$p + p - pp = p + p - p = p.$$

Определение. Операция дизъюнкции, примененная к вероятностным функциям F и Φ, дает вероятностную функцию $(F \vee \Phi)$ такую, что

$$p_i^{(F \vee \Phi)} \rightleftharpoons p_i^F + p_i^\Phi - p_i^F p_i^\Phi,$$

где умножение понимается как квазиумножение.

Наконец, дадим еще определение операции умножения вероятностной функции на число.

Операция умножения вероятностной функции F на число q, $0 \leqslant q \leqslant 1$, дает $\Phi \rightleftharpoons qF$, $p_i^\Phi \rightleftharpoons qp_i^F$.

Эта операция, очевидно, не выводит за пределы класса вероятностных функций. Ей тоже можно дать теоретико-вероятностное истолкование: p_i^Φ есть вероятность появления единицы при условии, что произошло некоторое другое событие, вероятность которого есть q, и которое не зависит от первого (например, при условии, что применяются преобразователи, о которых речь шла выше).

Обозначим через 1 функцию, которая на всех наборах равна 1, а через 0 — функцию, которая на всех наборах равна 0. Определим импликацию и эквивалентность вероятностных функций F и Φ: $(F \supset \Phi) \rightleftharpoons (\overline{F} \vee \Phi)$,

$$(F \equiv \Phi) \rightleftharpoons ((F \supset \Phi) \& (\Phi \supset F)).$$

Из определения операции непосредственно следует:

$$(F \& 1) \equiv F, \quad (1 \& F) \equiv F, \quad (F \& 0) \equiv 0, \quad (0 \& F) \equiv 0,$$
$$(F \vee 1) \equiv 1, \quad (1 \vee F) \equiv 1, \quad (F \vee 0) \equiv F, \quad (0 \vee F) \equiv F$$

для любой функции F.

Для конъюнкции и дизъюнкции непосредственно следует коммутативность этих операций:

$$(F \& \Phi) \equiv (\Phi \& F), \quad (F \vee \Phi) \equiv (\Phi \vee F)$$

и их ассоциативность:

$$(F \,\&\, (\Phi \,\&\, \Psi)) \equiv ((F \,\&\, \Phi) \,\&\, \Psi),$$
$$(F \vee (\Phi \vee \Psi)) \equiv ((F \vee \Phi) \vee \Psi).$$

Докажем ассоциативность дизъюнкции.

$((F \vee \Phi) \vee \Psi)$ означает: $p_i^F + p_i^\Phi - p_i^F p_i^\Phi + p_i^\Psi - (p_i^F + p_i^\Phi - p_i^F p_i^\Phi)\, p_i^\Psi = p_i^F + p_i^\Phi + p_i^\Psi - p_i^F p_i^\Phi - p_i^F p_i^\Psi - p_i^\Phi p_i^\Psi + p_i^F p_i^\Phi p_i^\Psi$.

с другой стороны, $(F \vee (\Phi \vee \Psi))$ дает: $p_i^F + p_i^\Phi + p_i^\Psi - p_i^\Phi p_i^\Psi - p_i^F (p_i^\Phi + p_i^\Psi - p_i^\Phi p_i^\Psi) = p_i^F + p_i^\Phi + p_i^\Psi - p_i^F p_i^\Phi - p_i^F p_i^\Psi - p_i^\Phi p_i^\Psi + p_i^F p_i^\Phi p_i^\Psi$,

таким образом, $((F \vee \Phi) \vee \Psi) \equiv (F \vee (\Phi \vee \Psi))$.

Докажем дистрибутивность дизъюнкции по отношению к конъюнкции.

$((F \vee \Phi) \,\&\, \Psi)$ дает: $(p_i^F + p_i^\Phi - p_i^F p_i^\Phi)\, p_i^\Psi = p_i^F p_i^\Psi + p_i^\Phi p_i^\Psi - p_i^F p_i^\Phi p_i^\Psi$,

с другой стороны, $((F \,\&\, \Psi) \vee (\Phi \,\&\, \Psi))$ означает $p_i^F p_i^\Psi + p_i^\Phi p_i^\Psi - p_i^F p_i^\Psi p_i^\Phi p_i^\Psi = p_i^F p_i^\Psi + p_i^\Phi p_i^\Psi - p_i^F p_i^\Phi p_i^\Psi$.

Таким образом, $((F \vee \Phi) \,\&\, \Psi) \equiv ((F \,\&\, \Psi) \vee (\Phi \,\&\, \Psi))$.

Докажем дистрибутивность конъюнкции по отношению к дизъюнкции.

$((F \,\&\, \Phi) \vee \Psi))$ дает: $p^F p^\Phi + p^\Psi - p^F p^\Phi p^\Psi$ (для сокращения вместо p_i пишем p); $((F \vee \Psi) \,\&\, (\Phi \vee \Psi))$ дает:

$(p^F + p^\Psi - p^F p^\Psi)\,(p^\Phi + p^\Psi - p^\Phi p^\Psi) = p^F p^\Phi + p^\Psi p^\Phi - p^F p^\Psi p^\Phi + p^F p^\Psi + p^\Psi p^\Psi - p^F p^\Psi p^\Psi - p^F p^\Phi p^\Psi - p^\Psi p^\Phi p^\Psi + p^F p^\Psi p^\Phi p^\Psi = p^F p^\Phi + p^\Phi p^\Psi - p^F p^\Phi p^\Psi + p^F p^\Psi + p^\Psi - p^F p^\Psi - p^F p^\Phi p^\Psi - p^\Phi p^\Psi + p^F p^\Phi p^\Psi = p^\Psi + p^F p^\Phi - p^F p^\Phi p^\Psi$,

таким образом, $((F \,\&\, \Phi) \vee \Psi)) \equiv ((F \vee \Psi) \,\&\, (\Phi \vee \Psi))$.

Законы де-Моргана тоже имеют место:

$\overline{(F \,\&\, \Phi)}$ дает: $1 - p^F p^\Psi$,

$(\bar{F} \vee \bar{\Phi})$ дает: $(1 - p^F) + (1 - p^\Phi) - (1 - p^F)(1 - p^\Phi) =$
$= 1 - p^F + 1 - p^\Phi - 1 + p^F + p^\Phi - p^F p^\Phi = 1 - p^F p^\Phi$,

т. е. $\overline{(F \,\&\, \Phi)} \equiv (\bar{F} \vee \bar{\Phi})$.

$\overline{(F \vee \Phi)}$ дает: $1 - p^F - p^\Phi + p^F p^\Psi$;

$(\bar{F} \,\&\, \bar{\Phi})$ дает: $(1 - p^F)(1 - p^\Phi) = 1 - p^F - p^\Phi + p^F p^\Phi$

и, таким образом, $\overline{(F \vee \Phi)} \equiv (\bar{F} \,\&\, \bar{\Phi})$.

Наконец, очевидным образом

$(F \,\&\, \bar{F}) \equiv 0$ дает: $p^F(1 - p^F) = p^F - p^F p^F = p^F - p^F = 0$,

$(F \vee \bar{F}) \equiv 1$ дает: $p^F + 1 - p^F = 1$.

Все это показывает, что алгебра вероятностных функций является моделью для аксиом булевой алгебры.

Имеет место теорема о разложении, являющаяся точным аналогом теоремы о разложении булевой функции в совершенную дизъюнктивную нормальную форму:

Пусть x_j^0 есть $\bar{x}_j$, x_j^1 есть x_j, вместо $(x_i \,\&\, x_k)$ пишем просто $x_i x_k$ тогда

$$F(x_1, \ldots, x_n) \equiv \bigvee_{0 \leqslant i \leqslant 2^n - 1} p_i^F x_1^{i_1} \ldots x_n^{i_n},$$

где $i_1 \ldots i_n$ есть двоичная запись числа i.

Дадим аналитический аналог понятия сети вероятностных диаграмм. Начнем с двухранговой регулярной сети.

Пусть дана вероятностная функция n переменных $F(b_1, \ldots, b_n)$ и n вероятностных функций (тоже n переменных) $b_1(a_1, \ldots, a_n), \ldots, b_n(a_1, \ldots, a_n)$. Эту совокупность можно, очевидно, рассматривать как аналитический аналог двухранговой регулярной сети вероятностных диаграмм, результирующей функцией которой является $F(b_1(a_1, \ldots, a_n), \ldots, b_n(a_1, \ldots, a_n))$. Здесь используется понятие подстановки вероятностной функции в вероятностную функцию; его достаточно определить для элементарных функций $x_1, \ldots, x_n$ и операций $\bar{}$, $\&$, $\vee$, так как по теореме о разложении любая вероятностная функция выражается с помощью этих средств. Определение обычное:

$$F_{\Phi\lfloor x_k \rfloor}^{x_i} \rightleftharpoons \begin{cases} \Phi, & \text{если } i = k, \\ x_k, & \text{если } i \neq k; \end{cases} \qquad F_{\Phi\lfloor \bar{\Psi} \rfloor}^{x_i} \rightleftharpoons \overline{F_{\Phi\lfloor \Psi \rfloor}^{x_i}};$$

$$F_{\Phi\lfloor (\Psi \,\&\, \chi) \rfloor}^{x_i} \rightleftharpoons (F_{\Phi\lfloor \Psi \rfloor}^{x_i} \,\&\, F_{\Phi\lfloor \chi \rfloor}^{x_i}),$$

$$F_{\Phi\lfloor (\Psi \vee \chi) \rfloor}^{x_i} \rightleftharpoons (F_{\Phi\lfloor \Psi \rfloor}^{x_i} \vee F_{\Phi\lfloor \chi \rfloor}^{x_i}).$$

Теорема. Пусть дана двухранговая сеть $F\ (b_1, \ldots, b_n),\ b_1\ (a_1, \ldots, a_n), \ldots, b_n\ (a_1, \ldots, a_n)$. Тогда результирующая функция сети Ф определяется соотношением

$$p^{\Phi} = \sum_{0 \leqslant j \leqslant 2^n - 1} p_j^F\, (p^{b_1})^{j_1} (p^{b_2})^{j_2} \ldots (p^{b_n})^{j_n},$$

где $j_1 \ldots j_n$ есть двоичная запись числа f; а $(p^{b_k})^1 \rightleftharpoons p^{b_k}$, $(p^{b_k})^0 \rightleftharpoons 1 - p^{b_k}$.

$$\Phi\,(a_1, \ldots, a_n) \equiv F\,(b_1\,(a_1, \ldots, a_n), \ldots, b_n\ (a_1, \ldots, a_n)).$$

Доказательство. Возьмем СДНФ для F и подставим вместо $b_1, \ldots, b_n$ их СДНФ, заметив предварительно, что $\bar{b}_i\,(a_1, \ldots, a_n) \equiv \bigvee_k (p_k^{b_i})^0\, a_1^{k_1} \ldots a_n^{k_n}$ (следует непосредственно из определения операции отрицания). В силу определения конъюнкции $((a_1^{k_1} \ldots a_n^{k_n})\ \&\ (a_1^{k_1'} \ldots a_n^{k_n'})) \equiv 0$, если набор $k_1 \ldots k_n$ отличен от набора $k_1' \ldots k_n'$. Поэтому мы получим

$$\Phi\,(a_1, \ldots, a_n) \equiv \bigvee_i p_i^F \Big(\bigvee_j (p_j^{b_1})^{i_1} (p_j^{b_2})^{i_2} \ldots (p_j^{b_n})^{i_n}\, a_1^{j_1} \ldots a_n^{j_n}\Big)$$

($i_1 \ldots i_n$ — двоичная запись числа i; $j_1 \ldots j_n$ — двоичная запись числа j). Теперь заметим, что если $i_1 \ldots i_n$ не совпадает с $i_1' \ldots i_n'$, то $(p_j^{b_1})^{i_1} \ldots (p_j^{b_n})^{i_n} (p_j^{b_1})^{i_1'} \ldots (p_j^{b_n})^{i_n'} = 0$ (в силу свойств квазиумножения). Следовательно,

$$(p_i^F\, (p_j^{b_1})^{i_1} \ldots (p_j^{b_n})^{i_n}\, a_1^{j_1} \ldots a_n^{j_n} \vee p_k^F\, (p_j^{b_1})^{j_1} \ldots (p_j^{b_n})^{k_n}\, a_1^{j_1} \ldots a_n^{j_n}) \equiv$$
$$\equiv (p_i^F\, (p_j^{b_1})^{i_1} \ldots (p_j^{b_n})^{i_n} + p_k^F\, (p_j^{b_1})^{k_1} \ldots (p_j^{b_n})^{k_n})\, a_1^{j_1} \ldots a_n^{j_n},$$

и мы получаем искомое, сгруппировав $a_1^{j_1} \ldots a_n^{j_n}$.

Проиллюстрируем теорему на случае n=2. $\bar{p}$ означает p^0. Имеем $F\,(b_1, b_2) \equiv (p_0^F \bar{b}_1 \bar{b}_2 \vee p_1^F \bar{b}_1 b_2 \vee p_2^F b_1 \bar{b}_2 \vee$

$$\vee\ p_3^F b_1 b_2),\quad b_i \equiv (p_0^{b_i} \bar{a}_1 \bar{a}_2 \vee p_1^{b_i} \bar{a}_1 a_2 \vee p_2^{b_i} a_1 \bar{a}_2 \vee p_3^{b_i} a_1 a_2),\ i = 1, 2,$$

откуда $\bar{b}_i \equiv (\bar{p}_0^{b_i} \bar{a}_1 \bar{a}_2 \vee \bar{p}_1^{b_i} \bar{a}_1 a_2 \vee \bar{p}_2^{b_i} a_1 \bar{a}_2 \vee \bar{p}_3^{b_i} a_1 a_2)$. Подставим $b_i,\ \bar{b}_i$ в $F\,(b_1, b_2)$:

$$(p_0^F(\bar{p}_0^{b_1}\bar{a}_1\bar{a}_2 \vee \bar{p}_6^{b_1}\bar{a}_1 a_2 \vee \bar{p}_2^{b_1}a_1\bar{a}_2 \vee \bar{p}_3^{b_1}a_1a_2)(\bar{p}_0^{b_2}\bar{a}_1\bar{a}_2 \vee$$
$$\vee \bar{p}_1^{b}\bar{a}_1a_2 \vee \bar{p}_2^{b_2}a_1\bar{a}_2 \vee \bar{p}_3^{b_2}a_1a_2) \vee p_1^F(\bar{p}_0^{b}\bar{a}_1\bar{a}_2 \vee$$
$$\vee \bar{p}_1^{b_1}\bar{a}_1a_2 \vee \bar{p}_2^{b_1}a_1\bar{a}_2 \vee \bar{p}_3^{b_1}a_1a_2)(p_0^{b_2}\bar{a}_1\bar{a}_2 \vee p_1^{b_2}\bar{a}_1a_2 \vee$$
$$\vee p_2^{b_2}a_1\bar{a}_2 \vee p_3^{b_2}a_1a_2) \vee p_2^F(p_0^{b_1}\bar{a}_1\bar{a}_2 \vee p_1^{b_1}\bar{a}_1a_2 \vee$$
$$\vee p_2^{b_1}a_1\bar{a}_2 \vee p_3^{b_1}a_1a_2)(\bar{p}_0^{b_2}\bar{a}_1\bar{a}_2 \vee \bar{p}_1^{b_2}\bar{a}_1a_2 \vee \bar{p}_2^{b_2}a_1\bar{a}_2 \vee$$
$$\vee \bar{p}_3^{b_2}a_1a_2) \vee p_3^F(p_0^{b_1}\bar{a}_1\bar{a}_2 \vee p_1^{b_1}\bar{a}_1a_2 \vee p_2^{b_1}a_1\bar{a}_2 \vee p_3^{b_1}a_1a_2)$$
$$(p_0^{b_2}\bar{a}_1\bar{a}_2 \vee p_1^{b_2}\bar{a}_1a_2 \vee p_2^{b}a_1\bar{a}_2 \vee p_3^{b}a_1a_2)) \equiv$$

$$\equiv ((p_0^F\bar{p}_0^{b_1}\bar{p}_0^{b_2} + p_1^F\bar{p}_0^{b_1}p_0^{b_2} + p_2^Fp_0^{b_1}\bar{p}_0^{b_2} + p_3^Fp_0^{b_1}p_0^{b_2})\bar{a}_1\bar{a}_2 \vee$$
$$\vee (p_0^F\bar{p}_1^{b_1}\bar{p}_1^{b_2} + p_1^F\bar{p}_1^{b_1}p_1^{b_2} + p_2^Fp_1^{b_1}\bar{p}_1^{b_2} + p_3^Fp_1^{b_1}p_1^{b_2})\bar{a}_1a_2 \vee$$
$$\vee (p_0^F\bar{p}_2^{b_1}\bar{p}_2^{b_2} + p_1^F\bar{p}_2^{b_1}p_2^{b_2} + p_2^Fp_2^{b_1}\bar{p}_2^{b_2} + p_3^Fp_2^{b_1}p_2^{b_2})a_1\bar{a}_2 \vee$$
$$\vee (p_0^F\bar{p}_3^{b_1}\bar{p}_3^{b_2} + p_1^F\bar{p}_3^{b_1}p_3^{b_2} + p_2^Fp_3^{b_1}\bar{p}_3^{b_2} + p_3^Fp_3^{b_1}p_3^{b_2})a_1a_2).$$

Результат теоремы может быть получен и с помощью элементарных теоретико-вероятностных рассуждений. Воспользуемся тем же примером. Какова вероятность, что входная последовательность 00 дает на выходе 1? Последовательность 00 с вероятностью $\bar{p}_0^{b_1}\bar{p}_0^{b_2}$ дает на выходе первого ранга (значения функций b_1 и b_2 мы рассматриваем как вероятность получить 1 в первом ранге) 00, с вероятностью $\bar{p}_0^{b_1}p_0^{b_2}$ дает 01, с вероятностью $p_0^{b_1}\bar{p}_0^{b_2}$ дает 10 и с вероятностью $p_0^{b_1}p_0^{b_2}$ дает 11. Эти события не пересекаются. Поэтому вероятность получить 1 во втором ранге будет:

$$p_0^F\bar{p}_0^{b_1}\bar{p}_0^{b_2} + p_1^F\bar{p}_0^{b_1}p_0^{b_2} + p_2^Fp_0^{b_1}\bar{p}_0^{b_2} + p_3^Fp_0^{b_1}p_0^{b_2}.$$

Но это и есть найденное нами значение функции Ф (a_1, a_2) на наборе 00.

Нетрудно видеть, что сеть является надежной, если результирующая функция сети — булева. Задача состоит в том, чтобы по заданной булевой функции построить надежную сеть, для которой данная функция будет результирующей. Иными словами, надо подобрать $p_0^F, \ldots, p_{2^n-1}^F, p_0^{b_1}, \ldots, p_{2^n-1}^{b_n}$ так, чтобы

$$p^{\Phi} \overline{\circ} \sum_{0 \leqslant j \leqslant 2^n-1} p_j^F (p^{b_1})^{j_1} \ldots (p^{b_n})^{j_n},$$

где p^{Φ} — либо 0, либо 1.

Посмотрим, когда выражение справа будет равняться единице.

Так как рассмотрение для всех n совершенно единообразно, но запись условий — весьма громоздка, продемонстрируем это на примере $n = 3$. В этом случае

$$p_i^{\Phi} = p_0^F \bar{p}_i^{b_1} \bar{p}_i^{b_2} \bar{p}_i^{b_3} + p_1^F \bar{p}_i^{b_1} \bar{p}_i^{b_2} p_i^{b_3} + p_2^F \bar{p}_i^{b_1} p_i^{b_2} \bar{p}_i^{b_3} + p_3^F \bar{p}_i^{b_1} p_i^{b_2} p_i^{b_3} +$$
$$+ p_4^F p_i^{b_1} \bar{p}_i^{b_2} \bar{p}_i^{b_3} + p_5^F p_i^{b_1} \bar{p}_i^{b_2} p_i^{b_3} + p_6^F p_i^{b_1} p_i^{b_2} \bar{p}_i^{b_3} + p_7^F p_i^{b_1} p_i^{b_2} p_i^{b_3}.$$

Прежде всего, $p_i^{\Phi} = 1$, если набор $p_i^{b_1}, p_i^{b_2}, p_i^{b_3}$ состоит только из 0 и 1, а p_j^F, где j — номер этого набора, рассматриваемого как двоичное разложение, есть 1. Например, пусть $p_i^F = 1$, тогда при $p_i^{b_1} = 0, p_i^{b_2} = 0, p_i^{b_3} = 1, p_i^{\Phi} = 1 \cdot \bar{0} \cdot$

$\cdot \bar{0} \cdot 1 = 1 \cdot 1 \cdot 1 \cdot 1 = 1.$

Пусть теперь не все $p_i^{b_k}$ равны 0 или 1. Например, рассмотрим набор $\{0, 0, p\}$. Если мы его подставим в выражение для p_i^{Φ}, то $p_i^{\Phi} = p_0^F \bar{0}\bar{0}\bar{p} + p_1^F \bar{0}\bar{0}p = p_0^F \bar{p} + p_1^F p.$

Если $p_0^F = p_1^F = 1$, то $p_c^{\Phi} = \bar{p} + p = 1 - p + p = 1.$

Вообще, если наш набор будет содержать p_i, отличное от 0 и 1, и только одно, то для того, чтобы $p_i^{\Phi} = 1$ на этом наборе, придется два из p_j^F положить равными 1.

Наконец, если набор содержит два p_i, отличных от 0 и 1, то придется положить равными 1 четыре коэффициента p_j^F. Например, набор $\{1, p, q\}$ дает $p_i^{\Phi} = 1$, если $p_4^F = p_5^F = p_6^F = p_7^F = 1$, так как $\bar{p}\bar{q} + \bar{p}q + p\bar{q} + pq = 1.$

Набор $\{p, q, r\}$ дает 1 только тогда, когда все $p_i^F = 1$, т. е. $F(b_1, b_2, b_3)$ является тавтологией; практического интереса этот случай не представляет.

Для произвольного n обобщение условий очевидно.

Условия для $p_i^{\Phi} = 0$ получаются из условий $p_i^{\Phi} = 1$ заменой единицы нулем.

Используя условия, можно легко сформулировать алгоритмы построения надежных сетей. Сразу видно, однако, что сети строятся неоднозначно. С другой стороны, для построения надежной сети существенно лишь общее число единиц среди p_i^{Φ} и совершенно несущественны их номера,— в том смысле, что, если построена надежная сеть, реализующая булеву функцию, равную единице на k наборах, то для любой другой булевой функции, равной единицей на k — других, вообще говоря — наборах, надежная сеть получится из первой простым изменением нумерации p_i. Наконец, случай k единиц и l нулей эквивалентен случаю k нулей и l единиц (с заменой

$p_i^F = 1$ на $p_i^F = 0$ и наоборот). Поэтому для практически интересных случаев $n = 3$, $n = 4$ и т. п. (для не очень больших n) можно построить стандартные сети; например для $n = 3$, если не интересоваться тавтологией и противоречием, будет всего $1 + \frac{2^3 - 2}{2} = 4$ сети. Рассмотрим эти случаи в качестве примера.

1. Булева функция равна единице на четырех наборах. Пусть их номера i_1, i_2, i_3, i_4; на наборах с номерами j_1, j_2, j_3, j_4 функция равна нулю. Положим $p_0^F = p_1^F = p_7^F = p_5^F = p_3^F = 1$.

Получим $p_2^F \bar{p}_i^{b_1} p_i^{b_2} \bar{p}_i^{b_3} + \bar{p}_i^{b_1} p_i^{b_2} p_i^{b_3} + p_4^F p_i^{b_1} \bar{p}_i^{b_2} \bar{p}_i^{b_3} + p_i^{b_1} \bar{p}_i^{b_2} p_i^{b_3} +$
$+ p_6^F p_i^{b_1} p_i^{b_2} \bar{p}_i^{b_3} + p_i^{b_1} p_i^{b_2} p_i^{b_3}$.

Легко видеть, что на наборе $\{p_i^{b_1}, 1, 1\}$ это даст 1, на наборе $\{1, p_i^{b_2}, 1\}$ — то же, а на наборах $\{0, 0, 0\}$ и $\{0, 0, p_0^{b_3}\}$ — нуль. Поэтому достаточно считать, что, например,

$$p_{j_1}^{b_1} = p_{j_2}^{b_1} = p_{j_3}^{b_1} = p_{j_4}^{b_1} = 0, \quad p_{i_3}^{b_1} = p_{i_4}^{b_1} = 1.$$
$$p_{j_1}^{b_2} = p_{j_2}^{b_2} = p_{j_3}^{b_2} = p_{j_4}^{b_2} = 0; \quad p_{i_1}^{b_2} = p_{i_2}^{b_2} = 1.$$
$$p_{j_1}^{b_3} = 0, \quad p_{i_1}^{b_3} = p_{i_2}^{b_3} = p_{i_3}^{b_3} = p_{i_4}^{b_3} = 1.$$

2. Булева функция равна 1 на пяти наборах; p_i^F — те же номера наборов, дающих 1, суть $i_1, \ldots, i_5$; номера наборов, дающих 0, суть j_1, j_2, j_3.

$$p_{j_1}^{b_1} = p_{j_2}^{b_1} = p_{j_3}^{b_1} = 0; \quad p_{i_1}^{b_1} = p_{i_2}^{b_1} = 1.$$
$$p_{j_1}^{b_2} = p_{j_2}^{b_2} = p_{j_3}^{b_2} = 0; \quad p_{i_1}^{b_2} = p_{i_4}^{b_2} = p_{i_5}^{b_2} = 1.$$
$$p_{j_1}^{b_3} = 0; \quad p_{i_1}^{b_3} = p_{i_2}^{b_3} = p_{i_3}^{b_3} = p_{i_4}^{b_3} = p_{i_5}^{b_3} = 1.$$

3. Булева функция равна 1 на шести наборах $i_1, \ldots, i_6$, нулю — на двух наборах j_1, j_2.

$$p_0^F = 0, \quad p_7^F = p_6^F = p_5^F = p_3^F = 1.$$
$$b_1: \; p_{j_1} = p_{j_2} = 0; \quad p_{i_3} = p_{i_4} = p_{i_5} = p_{i_6} = 1.$$
$$b_2: \; p_{j_1} = p_{j_2} = 0; \quad p_{i_1} = p_{i_4} = p_{i_5} = p_{i_6} = 1.$$
$$b_3: \; p_{j_1} = p_{j_2} = 0; \quad p_{i_1} = p_{i_2} = p_{i_5} = p_{i_6} = 1.$$

4. Булева функция равна 1 на семи наборах $i_1, \ldots, i_7$ и нулю на наборе j_1; p_i^F — те же.

$$b_1: \; p_{j_1} = 0; \qquad p_{i_4} = p_{i_6} = p_{i_5} = p_{i_7} = 1.$$
$$b_2: \; p_{j_1} = 0; \quad p_{i_1} = p_{i_2} = p_{i_3} = p_{i_6} = p_{i_7} = 1$$
$$b_3: \; p_{j_1} = 0; \quad p_{i_1} = p_{i_2} = p_{i_3} = p_{i_4} = p_{i_5} = 1.$$

Пока мы рассматривали двухранговые сети. Совершенно аналогичным образом можно рассмотреть многоранговые регулярные сети. Например, регулярная трехранговая сеть требует указания $2n+1$ функций:

$$F(c_1, \ldots, c_n);$$
$$c_1(b_1, \ldots, b_n), \ldots, c_n(b_1, \ldots, b_n);$$
$$b_1(a_1, \ldots, a_n), \ldots, b_n(a_1, \ldots, a_n).$$

результирующая функция сети получается посредством подстановки b_j в c_j, а c_i в F. Надежность определяется так же, как в случае двухранговой регулярной сети, и методы остаются теми же.

Наконец, можно снять требование регулярности, т. е. использовать более или менее произвольные подстановки. Это делает анализ более громоздким, но не прибавляет существенных трудностей.

Вероятностные функции, нужные для синтеза надежных сетей, отличаются тем, что среди их значений обязательно есть единицы. Иными словами, они описывают работу таких устройств, служащих для реализации булевых функций, которые на определенные комбинации входных воздействий отвечают безошибочно. Можно сказать, что в реализуемых ими СДНФ часть членов присутствует обязательно, а часть — с некоторой вероятностью.

В качестве работающих именно таким образом элементов ниже разбираются формальные нейроны.

Описанный в настоящем параграфе аналитический аппарат предложен С. Л. Никогосовым 1966 г. На диаграммах все вычисления, как нетрудно видеть, упрощаются. Разберем два примера.

Пример 1. Возьмем сеть, отличающуюся от сети, изображенной на рис. 73, только тем, что все ее буквы p различны; для краткости вместо $p_{1,1,k}$, $p_{1,2,k}$, $p_{1,3,k}$, $p_{2,1,k}$, $p_{2,2,k}$, $p_{2,3,k}$, $p_{3,1,k}$ (индексы буквы $p_{i,j,k}$ обозначают: i — номер ранга, j — номер диаграммы в ранге, k — номер ячейки диаграммы) будем писать соответственно p_k, q_k, r_k, s_k, u_k, v_k, w_k:

$$\begin{pmatrix} 0 & 0 & p_2 & p_3 \\ p_4 & p_5 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & q_2 & q_3 \\ q_4 & q_5 & 0 & 1 \end{pmatrix} \begin{pmatrix} r_0 & r_1 & 0 & 0 \\ 0 & 0 & r_6 & 1 \end{pmatrix}$$

$$\begin{pmatrix} s_0 & 0 & 0 & s_3 \\ s_4 & s_5 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_5 & 0 & u_2 & u_3 \\ 0 & u_5 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & v_1 & 0 & v_3 \\ 0 & v_5 & v_6 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 1 & w_3 \\ 1 & w_5 & w_6 & 0 \end{pmatrix}.$$

Остановимся на функционировании первой диаграммы второго ранга. Обозначим через $[s_i]$ вероятность появления единицы в i-ой ячейке ее результирующей диаграммы. Получим

$$[s_0] = \bar{0}\bar{0}\bar{r}_0 s_0 + \bar{0}\bar{0}r_0 0 + \bar{0}0\bar{r}_0 0 + \bar{0}0r_0 s_3 + 0\bar{0}\bar{r}_0 s_4 + 0\bar{0}r_0 s_5 + 00\bar{r}_0 0 + 00r_0 1 = \bar{r}_0 s_0,$$

$$[s_1] = \bar{r}_1 s_0,$$

$$[s_2] = \bar{p}_2\bar{q}_2 s_0 + p_2\bar{q}_2 s_4,$$

$$[s_3] = \bar{p}_3\bar{q}_3 s_0 + p_3\bar{q}_3 s_4,$$

$$[s_4] = \bar{p}_4\bar{q}_4 s_0 + p_4\bar{q}_4 s_4,$$

$$[s_5] = \bar{p}_5\bar{q}_5 s_0 + p_5\bar{q}_5 s_4,$$

$$[s_6] = \bar{r}_6 s_0,$$

$$[s_7] = 1.$$

Аналогично для остальных диаграмм второго ранга:

$$[u_0] = 0,\ [u_1] = 0,\ [u_2] = \bar{p}_2 q_2 u_2,\ [u_3] = \bar{p}_3 q_3 u_2,$$
$$[u_4] = \bar{p}_4 q_4 u_2,\ [u_5] = \bar{p}_5 q_5 u_2,\ [u_6] = 0,\ [u_7] = 1;$$

$$[v_0] = r_0 v_1,\ [v_1] = r_1 v_1,\ [v_2] = p_2 q_2 v_6,\ [v_3] = p_3 q_3 v_6,$$
$$[v_4] = p_4 q_4 v_6,\ [v_5] = p_5 q_5 v_6,\ [v_6] = r_6 v_1,\ [v_7] = 1.$$

Итак результирующие диаграммы второго ранга имеют, соответственно, вид

$$\begin{bmatrix} [s_0] & [s_1] & [s_2] & [s_3] \\ [s_4] & [s_5] & [s_6] & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 & [u_2] & [u_3] \\ [u_4] & [u_5] & 0 & 1 \end{bmatrix}, \begin{bmatrix} [v_0] & [v_1] & [v_2] & [v_3] \\ [v_4] & [v_5] & [v_6] & 1 \end{bmatrix}.$$

Перейдем к построению результирующей диаграммы сети.

$$[w_0] = \overline{[s_0]}\,\overline{[v_0]} + \overline{[s_0]}\,[v_0] + [s_0]\,\overline{[v_0]} + [s_0]\,[v_0]\,w_5 =$$
$$= \overline{[s_0]\,[v_0]} + [s_0]\,[v_0]\,w_5 = \overline{\bar{r}_0 s_0 r_0 v_1} + \bar{r}_0 s_0 r_0 v_1 w_5 = 1,$$

$$[w_1] = \overline{[s_1]\,[v_1]} + [s_1]\,[v_1]\,w_5 = \overline{\bar{r}_1 s_0 r_1 v_1} + \bar{r}_1 s_0 r_1 v_1 w_5 = 1,$$

$$[w_2] = \overline{[s_2]}\,\overline{[u_2]}\,\overline{[v_2]} + \overline{[s_2]}\,\overline{[u_2]}\,[v_2] + \overline{[s_2]}\,[u_2]\,\overline{[v_2]} +$$
$$+ \overline{[s_2]}\,[u_2]\,[v_2]\,w_3 + [s_2]\,\overline{[u_2]}\,\overline{[v_2]} + [s_2]\,\overline{[u_2]}\,[v_2]\,w_5 +$$
$$+ [s_2]\,[u_2]\,\overline{[v_2]}\,w_6 = \overline{(\bar{p}_2\bar{q}_2 s_0 + p_2\bar{q}_2 s_4)}\,\bar{p}_2 q_2 u_2 +$$
$$+ \overline{(\bar{p}_2\bar{q}_2 s_0 + p_2\bar{q}_2 s_4)}\,\bar{p}_2 q_2 u_2 p_2 q_2 v_6 + (\bar{p}_2\bar{q}_2 s_0 + p_2\bar{q}_2 s_4)\cdot$$
$$\overline{\bar{p}_2 q_2 u_2}\cdot\overline{p_2 q_2 v_6} + \overline{(\bar{p}_2\bar{q}_2 s_0 + p_2\bar{q}_2 s_4)}\,\bar{p}_2 q_2 u_2 p_2 q_2 v_6 w_3 +$$
$$+ (\bar{p}_2\bar{q}_2 s_0 + p_2\bar{q}_2 s_4)\,\overline{\bar{p}_2 q_2 u_2}\,p_2 q_2 v_6 w_5 + (\bar{p}_2\bar{q}_2 s_0 + p_2\bar{q}_2 s_4)\cdot$$
$$\cdot\,\bar{p}_2 q_2 u_2\overline{p_2 q_2 v_6} w_6 = p_2 q_2 + p_2\bar{s}_4 + \bar{s}_0\bar{p}_2\bar{q}_2 + \bar{s}_0\bar{q}_2\bar{s}_4 +$$
$$+ \bar{s}_0\bar{p}_2\bar{u}_2 + q_2\bar{u}_2 + p_2\bar{u}_2\bar{s}_4 - \bar{s}_0\bar{u}_2\bar{s}_4 + \bar{p}_2 q_2 u_2 +$$
$$+ \bar{p}_2\bar{q}_2 s_0 + p_2\bar{q}_2 s_4 = p_2 + \bar{p}_2\bar{q}_2 + \bar{p}_2 q_2 = 1.$$

Аналогично получим, что $[w_3] = [w_4] = [w_5] = 1$, $[w_6] = 1$, $[w_7] = 0$.

Таким образом, полученная результирующая диаграмма совпадает с результатом работы надежной сети, изображенной на рис. 73.

4.3. Диаграммы Венна в классическом исчисления одноместных предикатов

Формулы исчисления одноместных предикатов определяются следующим образом.

Рассматривается алфавит, состоящий из предикатных переменных $P_1 \ldots P_s$, предметных переменных $x_1 \ldots x_h$, логических знаков $\neg\ \&\ \vee\ \supset\ \forall\exists$, скобок (), где s, h —целые положительные числа, $\forall$ — квантор общности, $\exists$ — квантор существования.

1. Если P_i — предикатная, x_j — предметная переменные, то слово $P_i(x_j)$ $(i = 1, \ldots, s,\ j = 1, 1 \ldots, h)$ считается атомарной формулой. Атомарная формула считается формулой.
2. Если Φ — формула, то $\neg\,\Phi$ считается формулой.
3. Если Φ и Ψ — формулы, то $(\Phi\ \&\ \Psi)$, $(\Phi \vee \Psi)$, $(\Phi \supset \Psi)$ считаются формулами.
4. Если Φ — формула, x — предметная переменная, то $\forall x\,\Phi$, $\exists x\Phi$ считаются формулами. О формуле Φ в этом случае

говорят, что она является областью действия соответствующего квантора.

Вхождение переменной x в формулу Ψ называется связанным, если оно является вхождением в область действия квантора $\forall x$ или $\exists x$; в противном случае вхождение называется свободным.

Если переменная x входит в качестве свободной (связанной) переменной в Ψ, то говорят, что x является свободной (связанной) переменной в Ψ.

Мы не будем описывать подробно исчисление одноместных предикатов, ограничимся введением лишь некоторых необходимых для изложения понятий.

4.3.1. Диаграммы Венна и формулы исчисления одноместных предикатов (определения, построение формул по диаграммам)

Квазибазисные высказывания формулы. Пусть Ф — формула, $Q_1 \ldots Q_m$—все графически различные предикатные переменные, из которых составлена формула Ф.

Формулы вида $(\alpha_1 Q_1(x) \,\&\, \ldots \,\&\, \alpha_m Q_m(x))$, где x — предметная переменная формулы Φ, α_i есть или Λ (пустое слово), или $\neg$, $i = 1, \ldots, m$, будем называть квазибазисными высказываниями формулы Ф. Например, $(\neg Q_1(x) \,\&\, Q_2(x) \,\&\, \neg Q_3(x))$ суть квазибазисное высказывание формулы

$$\forall y\,(Q_1(x) \supset Q_2(y)). \qquad (3.1)$$

Базисные высказывания формулы. Пусть Ф — формула, $\Psi(x)$ — квазибазисное высказывание формулы Ф. Формулу $\exists x \Psi(x)$ будем называть базисным высказыванием формулы Ф.

Например, $\exists x (\neg Q_1(x) \,\&\, Q_2(x) \,\&\, \neg Q_3(x))$ суть базисное высказывание формулы (3.1), а $\exists x (\neg Q_1(x) \,\&\, \neg Q_3(x))$ не является базисным высказываниемформулы (3.1).

Из построения базисных высказываний непосредственно следует, что всего графически различных базисных высказываний формулы Ф, содержащих предикатные переменные $Q_1, \ldots, Q_m$ и одну предметную переменную x, может быть 2^m.

Пусть X — произвольная предметная область (предполагаем, что предметная область не пуста — см. п.1.4.1.3; в качестве предметной области можно брать непустые множества «произвольной природы», однако для понимания материала достаточно ограничиться

множествами, состоящими из конечного числа элементов). Пусть $R_1, \ldots, R_m$— произвольные одноместные предикаты, определенные на *X*.

Каждому предикату $R_i\ (i = 1, \ldots, m)$ можно поставить в соответствие класс тех элементов *x* предметной области *X,* на которых предикат R_i выполняется, или, как говорят, $R_i\ (x)$ принимает значение 1 (единицей, обозначается «истина», нулем — «ложь»).

Всего графически различных последовательностей значений предикатов $R_1, \ldots, R_m$ будет 2^m, их можно собрать в следующую таблицу:

Таблица

	R_1	R_2	R_3	. . .	R_{m-1}	R_m
0)	0	0	0	. . .	0	0
1)	0	0	0	. . .	0	1
2)	0	0	0	. . .	1	0
3)	0	0	0	. . .	1	1
			. . .	. . .		. . .
2^m-3)	1	1	1	. . .	0	1
2^m-2)	1	1	1	. . .	1	0
2^m-1)	1	1	1	. . .	1	1

Нетрудно заметить, что последовательности значений предикатов $R_1, \ldots, R_m$ являются элементарными.

Для каждого элемента *x* рассматриваемой предметной области *X* каждая формула $R_i\ (x),\ i = 1, \ldots, m,$принимает значение 1 или 0. Поэтому каждому элементу *x* предметной области *X* отвечает одна и только одна *m*-членная элементарная последовательность, члены которой являются значениями $R_i(x)$. Объединяя все элементы, которым соответствует одна элементарная последовательность, разделим предметную область *X* на 2^m попарно непересекающихся классов. Так как предметная область *X* есть непустое множество, то по крайней мере один из 2^m классов не пуст.

Каждой m-членной элементарной последовательности β будет соответствовать только один из полученных 2^m классов, для элементов которого последовательность значений предикатов $R_1, \ldots, R_m$ совпадает с данной последовательностью β. Из последовательности β можно образовать, заменяя 0 на $\neg$, 1 на пустое слово Λ, последовательность α_β: $\alpha_1 \ldots \alpha_m$. Для элементов указанного класса последовательность значений предикатов $\alpha_1 R_1, \ldots \alpha_m R_m$ образует элементарную последовательность, имеющую номер $2^m - 1$ (эта последовательность состоит из m единиц).

Последовательности α_β соответствует одно базисное высказывание формулы Ф:

$$\exists x (\alpha_1 Q_1 (x) \& \ldots \& \alpha_m Q_m (x)).$$

Поскольку любую предметную область X для любых предикатов $R_1, \ldots, R_m$, определенных на X, можно разделить на 2^m попарно непересекающихся классов, по крайней мере, один из которых не пуст, то хотя бы одно базисное высказывание формулы Ф принимает значение 1.

Эквивалентные формулы. Каждая формула представляет собой определенное утверждение, истинное или ложное, когда оно относится к определенной предметной области X иногда все предикатные переменные и все свободные предметные переменные замещены, соответственно, индивидуальными предикатами, определенными на X, и индивидуальными предметами из X.

Говорят, что две формулы Ф и Ψ *эквивалентны на предметной области* X, если при всех таких замещениях они принимают одинаковое значение 1 или 0.

Если две формулы Ф и Ψ эквивалентны на любых предметных областях, то их называют просто эквивалентными, и записывают $\Phi \equiv \Psi$.

Общезначимые формулы. Если формула Ф принимает значение 1 для некоторой предметной области X, некоторых предикатов, определенных на X, и некоторых значений предметных переменных из X, то формулу Ф называют выполнимой.

Если Ф принимаег значение 1 для любых предикатов, определенных на некоторой предметной области X, и при любых значениях предметных переменных из X, то формулу Ф называют общезначимой на X.

Если формула Ф общезначима на любой предметной области, то формулу Ф называют универсально общезначимой или тождественно истинной.

Формулу называют тождественно ложной, или невыполнимой, если ни для какой предметной области ни при каких замещениях предикатов и предметных переменных она не принимает значения 1.

Формулы Ф и Ψ называются равносильными (относительно общезначимости), если Ф общезначима тогда и только тогда, когда общезначима Ψ.

Конъюнктивные приведенные формулы. Формулу Ф будем называть конъюнктивной приведенной формулой, если в нее входит только одна предметная переменная и если Ф есть конъюнкция дизъюнкций, членами которых являются базисные высказывания формулы Ф и отрицания базисных высказываний формулы Ф. (Предполагаем, что все рассматриваемые базисные высказывания формулы Ф составлены, как и всюду в дальнейшем, из одних и тех же предикатных переменных.)

Нетрудно доказать предложение 3.1:

Конъюнктивная приведенная формула Ф универсально общезначима тогда и только тогда, когда или каждый ее конъюнктивный член F_j есть дизъюнкция всех базисных высказываний формулы Ф, или в F_j одновременно входят, по крайней мере, одно базисное высказывание формулы Ф и его отрицание.

Конъюнктивные приведенные формы формул. Если Ф — формула, Ψ — конъюнктивная приведенная формула, и $\Phi \equiv \Psi$, то Ψ будем называть конъюнктивной приведенной *формой* формулы Ф.

Дизъюнктивные приведенные формулы. Формулу Ф будем называть дизъюнктивной приведенной формулой, если в нее входит только одна предметная переменная и если Ф есть дизъюнкция конъюнкций, членами которых являются базисные высказывания формулы Ф и отрицания базисных высказываний формулы Ф.

Нетрудно доказать следующее предложение 3.2:

Дизъюнктивная приведенная формула Ф тождественно ложна тогда и только тогда, когда или каждый ее дизъюнктивный член F_j есть конъюнкция всех отрицаний базисных высказываний формулы Ф, или в F_j одновременно входят, по крайней мере, одно базисное высказывание формулы Ф и его отрицание.

Дизъюнктивные приведенные формы формулы. Если Ф — формула, Ψ — дизъюнктивная приведенная формула и $\Phi \equiv \Psi$, то Ψ будем называть *дизъюнктивной приведенной формой* формулы Ф.

Диаграммы Венна. В формулу может входить более одной предметной переменной, поэтому остановимся на вопросе о расположении на плоскости k (где $k > 1$) попарно непересекающихся предметных областей.

Пусть $X_1, \ldots, X_k$ — любые предметные области. Разделим плоскость по методу Венна на 2^{m+1} ячеек (см. п.1.4.1.5), где m — число одноместных предикатов $Q_1, \ldots, Q_m$, определенных на областях $X_1, \ldots, X_k$.

Предметную область X_2 будем изображать внутренностью $(m+1)$-ой замкнутой линии Жордана на плоскости; предметную область X_3 будем изображать внутренностью $(m+2)$-ой замкнутой линии Жордана, проведенной по методу Венна так, что она не пересекает область X_2, а совпадает с ее границей от точки бывшего входа в область X_2 до точки бывшего выхода из X_2; при этом пересечение X_2 и X_3, то есть $X_2 \cap X_3$ — пустое множество; и так далее. Предметную область X_k будем изображать внутренностью $(m+k+1)$-ой замкнутой линии Жордана, проведенной по методу Венна так, что она не пересекает область $X_2 \cup \ldots \cup X_{k-1}$, а совпадает с границей области $X_2 \cup \ldots \cup X_{k-1}$ от точки бывшего входа в эту область до точки бывшего выхода из нее, при этом $X_2 \cap \ldots \cap X_k$ — пустое множество. Предметную область X_1 будем изображать дополнением к множеству $X_2 \cup \ldots \cup X_k$. При таком расположении $m+k-1$ фигур плоскость делится на $k \cdot 2^m$ ячеек, и каждая предметная область X_j — на 2^m ячеек с помощью фигур $Q_1, \ldots, Q_m$. Например, пусть $m = 2$, способы разбиения плоскости на $k \cdot 2^m$ ячеек приведены при $k = 1$, на рис. 75, при $k = 2$ — на рис. 76, при $k = 3$ — на рис. 77, при $k = 4$ — на рис. 78.

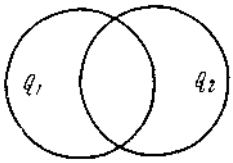


Рис. 75

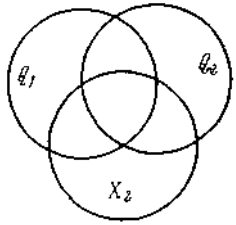


Рис. 76

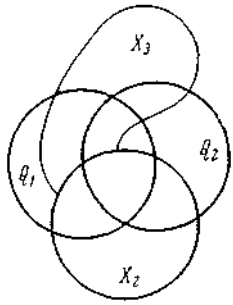


Рис. 77

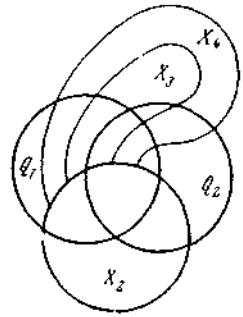


Рис. 78

Приведенный метод разбиения плоскости на $k \cdot 2^m$ ячеек с помощью $m + k$ — 1 фигур будем называть *обобщенным методом Венна.* Аналогично строятся обобщенные таблицы Венна и вводится обобщенный символ Венна. На рис. 79 изображена обобщенная таблица Венна при $k = 3$ и $m = 5$.

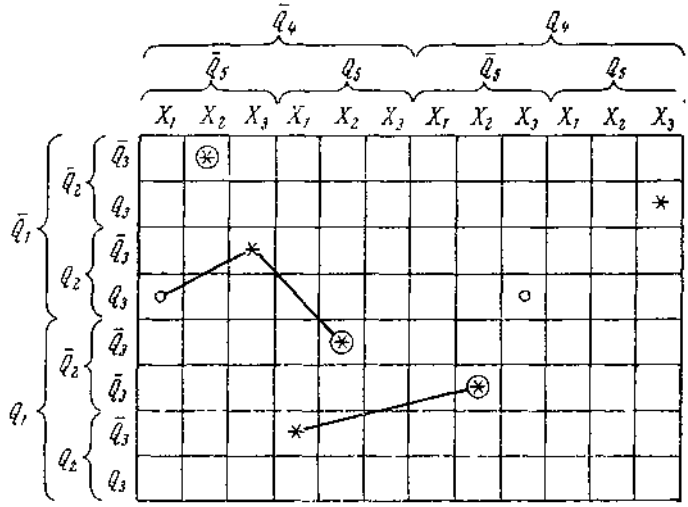


Рис. 79

Обобщенный символ Венна, в котором расположено несколько ∘ ∗ ⊛-ломаных, будем называть диаграммой Венна в исчислении одноместных предикатов. (Для краткости в настоящем разделе вместо слов «диаграмма Венна в исчислении одноместных предикатов» часто будем говорить «диаграмма Венна».)

Каждой диаграмме Венна можно поставить в конъюнктивное или дизъюнктивное соответствие некоторую формулу Ф следующим образом: Кружочку $\circ$ ставим в соответствие квазибазисное высказывание, звездочке $*$ — базисное высказывание, звездочке в кружке $\circledast$ — отрицание базисного высказывания некоторой формулы Ф. Каждой $\circ * \circledast$-ломаной дизъюнктивно (конъюнктивно) соответствует конъюнкция (дизъюнкция) формул, написанных по знакам $\circ$, $*$ и $\circledast$. Диаграмме Венна дизъюнктивно (конъюнктивно) поставим в соответствие дизъюнкцию (конъюнкцию) формул, соответствующих $\circ * \circledast$-ломаным. Смысл понятий дизъюнктивного (конъюнктивного) соответствия поясним на следующих примерах.

Диаграмме, изображенной на рис. 80, дизъюнктивно соответствует формула

$$(((P(x) \& \neg Q(x)) \& \exists x (P(x) \& \neg Q(x)) \& \neg \exists x (P(x) \& \& Q(x))) \vee (\exists y (P(y) \& \neg Q(y)) \& (\neg P(y) \& \neg Q(y)) \& \& \neg \exists x (\neg P(x) \& Q(x)))),$$

конъюнктивно — формула

$$(((P(x) \& \neg Q(x)) \vee \exists x (P(x) \& \neg Q(x)) \vee \neg \exists x (P(x) \& \& Q(x))) \& (\exists y (P(y) \& \neg Q(y)) \vee (\neg P(y) \& \neg Q(y)) \vee \neg \exists x (\neg P(x) \& Q(x))));$$

по диаграмме, изображенной на рис. 67, аналогично строятся формулы: дизъюнктивно —

$$(\neg \exists x_2 (\neg Q_1(x_2) \& \neg Q_2(x_2) \& \neg Q_3(x_2) \& \neg Q_4(x_2) \& \& \neg Q_5(x_2)) \vee ((\neg Q_1(x_1) \& Q_2(x_1) \& Q_3(x_1) \& \neg Q_4(x_1) \& \& \neg Q_5(x_1)) \& \exists x_3 (\neg Q_1(x_3) \& Q_2(x_3) \& \neg Q_3(x_3) \& \neg Q_4(x_3) \& \& \neg Q_5(x_3)) \& \neg \exists x_2 (Q_1(x_2) \& \neg Q_2(x_2) \& \neg Q_3(x_2) \& \neg Q_4(x_2) \& \& Q_5(x_2))) \vee (\exists x_1 (Q_1(x_1) \& Q_2(x_1) \& \neg Q_3(x_1) \& \neg Q_4(x_1) \& \& Q_5(x_1)) \& \neg \exists x_2 (Q_1(x_2) \& \neg Q_2(x_2) \& Q_3(x_2) \& Q_4(x_2) \& \& \neg Q_5(x_2)) \vee (\neg Q_1(x_3) \& Q_2(x_3) \& Q_3(x_3) \& Q_4(x_3) \& \& \neg Q_5(x_3)) \vee \exists x_3 (\neg Q_1(x_3) \& \neg Q_2(x_3) \& Q_3(x_3) \& Q_4(x_3) \& \& Q_5(x_3))),$$

конъюнктивно —

$$(\neg\exists x_2(\neg Q_1(x_2)\&\neg Q_2(x_2)\&\neg Q_3(x_2)\&\neg Q_4(x_2)\&\neg Q_5(x_2))\&$$
$$\&((\neg Q_1(x_1)\&Q_2(x_1)\&Q_3(x_1)\&\neg Q_4(x_1)\&\neg Q_5(x_1))\vee\exists x_3(\neg$$
$$Q_1(x_3)\&Q_2(x_3)\&\neg Q_3(x_3)\&\neg Q_4(x_3)\&\neg Q_5(x_3))\vee\neg\exists x_2(Q_1(x_2)\&$$
$$\&\neg Q_2(x_2)\&\neg Q_3(x_2)\&\neg Q_4(x_2)\&Q_5(x_2)))\&(\exists x_1(Q_1(x_1)\&$$
$$Q_2(x_1)\&\neg Q_3(x_1)\&\neg Q_4(x_1)\&Q_5(x_1))\vee\neg\exists x_2(Q_1(x_2)\&\neg$$
$$\neg Q_2(x_2)\&Q_3(x_2)\&Q_4(x_2)\&\neg Q_5(x_2))\&(\neg Q_1(x_3)\&Q_2(x_3)\&$$
$$Q_3(x_3)\&Q_4(x_3)\&\neg Q_5(x_3))\&\exists x_3(\neg Q_1(x_3)\&\neg Q_2(x_3)\&$$
$$\&Q_3(x_3)\&Q_4(x_3)\&Q_5(x_3))).$$

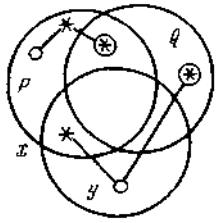


Рис. 80

4.3.2. Операции над диаграммами Венна в логике одноместных предикатов

Операция Q_1. Пусть L_i — ∘ * ⊛-ломаная диаграммы Венна, i=1, ..., l. Преобразуем каждую L_i так, чтобы ни в одной ячейке диаграммы на L_i не находилось двух и более одинаковых знаков. Полученную диаграмму будем называть результатом операции O_1. Операция O_1 соответствует эквивалентности $(A \vee A) \equiv A$ при конъюнктивном соответствии или $(A \& A) \equiv A$ при дизъюнктивном соответствии, где A —или квазибазисное высказывание некоторой формулы Ф, или базисное высказывание формулы Ф, или отрицание базисного высказывания формулы Ф.

Операция O_2. Пусть L_1, ..., L_l — все ∘ * ⊛-ломаные диаграммы Венна B. Если на B существуют такие две ломаные L_k и L_j, $k \neq j$, что в каждой ячейке, в которой лежит некоторый знак ломаной L_k, находится такой же знак ломаной L_j, то ломаную L_j можно уничтожить. Результатом операции O_2 является такая диаграмма, на которой не существует таких пар ломаных. Операция O_2 соответствует закону поглощения $(A \vee (A \& C)) \equiv A$ при дизъюнктивном соответствии или закону поглощения $(A_1 \& (A_1 \vee C_1)) \equiv A_1$ при

конъюнктивном соответствии, где A и $(A \,\&\, C)$— конъюнкции квазибазисных, базисных и отрицаний базисных высказываний некоторой формулы Ф, A_1 и $(A_1 \vee C_1)$ — дизъюнкции квазибазисных, базисных и отрицаний базисных высказываний некоторой формулы Φ_1.

Операция O_3. Операция O_3 состоит в уничтожении на диаграмме таких ломаных, у которых по крайней мере два знака $*$, $\circledast$ лежат в одной ячейке области X_k, и таких ломаных, на которых находится не менее двух кружков, один из которых принадлежит одной из ячеек области $X_j \cap Q_i$, другой —одной из ячеек области $X_j \setminus Q_i$. В случае, когда все $\circ * \circledast$-ломаные уничтожаются, результатом операции O_3 будем считать диаграмму, разбитую на то же число ячеек, что и исходная, в которой расположена только одна $* \circledast$-ломаная, состоящая только из двух знаков $*$ и $\circledast$, находящихся в одной ячейке. Преобразования при применении операции O_3 соответствуют эквивалентностям:

$$(\exists x A(x) \,\&\, \neg \exists x A(x) \,\&\, D) \equiv 0, \quad (B(y) \,\&\, C(y) \,\&\, D) \equiv 0$$

при дизъюнктивном соответствии или

$$(\exists x A(x) \vee \neg \exists x A(x) \vee D) \equiv 1,$$
$$(B(y) \vee \neg B(y) \vee D) \equiv 1$$

при конъюнктивном соответствии, где $A(x)$, $B(y)$ и $C(y)$ — квазибазисные высказывания некоторой формулы Ф и в $B(y)$ входит предикат P, а в $C(y)$ входит его отрицание $\neg P$.

Операция O_4. Пусть $L_1, \ldots, L_h$— все ломаные диаграммы. В ячейках области X_j вместо кружков на всех L_i, $i = 1, \ldots, h$, поставим звездочки; к полученной диаграмме применим последовательно операции O_1, O_2, O_3. Если на диаграмме останутся только $* \circledast$-ломаные, то сотрем границы областей $X_2, \ldots, X_k$; в результате плоскость или таблица разделится на 2^m ячеек; к полученной диаграмме применим последовательно операции O_1, O_2, O_3.

Операция O_4 соответствует эквивалентности

$$\exists x_j ((A \& B) \vee (C \& D)) \equiv ((\exists x_j A \& B) \vee (C \& \exists x_j D))$$

при дизъюнктивном соответствии, где A, D — квазибазисные высказывания некоторой формулы Ф, в которые входит предметная переменная x_j; B, C — конъюнкции квазибазисных, базисных и отрицаний базисных высказываний формулы Ф, не содержащие свободно переменную x_j.

Операция O_5. Пусть B_1, B_2 — диаграммы Венна (у которых, как и обычно при одновременном рассмотрении нескольких диаграмм, предметные и предикатные переменные соответственно совпадают).

Перенесем на B_1 все ломаные диаграммы B_2; к полученной диаграмме применим последовательно операции O_1, O_2, O_3.

Операция O_5 соответствует дизъюнкции двух формул $(A_1 \vee \ldots \vee A_k)$, $(C_1 \vee \ldots \vee C_t)$ при дизъюнктивном соответствии или конъюнкции двух формул

$$(A_1 \& \ldots \& A_k), \quad (C_1 \& \ldots \& C_t)$$

при конъюнктивном соответствии, где A_i, C_i — конъюнкции или дизъюнкции квазибазисных, базисных и отрицаний базисных высказываний некоторой формулы.

Операция O_6. Пусть $L_1, \ldots, L_h$ — все ломаные диаграммы B_1; $M_1, \ldots, M_s$ — все ломаные диаграммы B_2. Перенесем M_1 на B_1 и присоединим к L_j; ... и т. д.; перенесем M_s на B_1 и присоединим к L_j; $j = 1,\ldots, r$. К полученной диаграмме применим последовательно операции O_1, O_2, O_3.

Операция O_6 соответствует конъюнкции двух формул $(A_1 \vee \ldots \vee A_r)$, $(C_1 \vee \ldots \vee C_s)$ при дизъюнктивном соответствии или дизъюнкции двух формул

$$(A_1 \& \ldots \& A_r), \quad (C_1 \& \ldots \& C_s)$$

при конъюнктивном соответствии, где A_i, C_i — конъюнкции или дизъюнкции квазибазисных, базисных и отрицаний базисных высказываний некоторой формулы.

Операция O_7. Пусть $L_1, \ldots, L_n$ — все ломаные диаграммы B. Построим диаграммы D_i, $i = 1, \ldots, n$, каждая из которых содержит только одну ломаную L_i диаграммы B. Предположим, что k_i — число знаков $*$ на L_i, l_i— число знаков $\circledast$ на L_i, l_i— число областей X_s на D_i; содержащих по одному кружку. Вместо ломаной L_i на D_i построим $k_i + p_i + l_i(2^m - 1)$ ломаных, где 2^m — число ячеек каждой области на D_i: M_j $(j = 1, \ldots, k_i)$ — $\circledcirc$-ломаные;

M_t $(t = k_i + 1, \ldots, k_i + p_i)$ — $*$-ломаные,

M_v $(v = k_i + p_i + 1, \ldots, k_i + p_i + l_i \ (2^m - 1))$ — $\circ$-ломаные;

M_j: заменим знак $*$ на знак $\circledcirc$; M_t: заменим знак $\circledast$ на знак $*$; M_v: в каждой ячейке, не содержащей знаков $\circ$, области X_s поставим по одному знаку $\circ$, в каждой области X_s получим $2^m - 1$ $\circ$-ломаных. Построенные диаграммы обозначим E_i, $i = 1, \ldots, n$.

Если $n = 1$, то результат операции O_7 есть диаграмма E_1. Если $n>1$, то результатом операции O_7 является следующая диаграмма B_n: к диаграммам E_1 и E_2 применим операцию O_6, получим диаграмму B_2; к диаграммам B_2 и E_3 применим операцию O_6, получим диаграмму B_3; и т. д.; к диаграммам B_{n-1} и E_n применим операцию O_6, получим диаграмму B_n.

Операция O_7, — операция отрицания диаграммы.

4.3.3. Соответствие между формулами и диаграммами Венна в исчислении одноместных предикатов

Укажем индуктивный способ построения диаграмм Венна для записи информации, заданной в виде произвольной формулы Ф. Диаграмму Венна, соответствующую формуле Ф, будем обозначать [Ф] или [Ф] ∘ ∗ ⊛.

Пусть Ф — произвольная формула. Пусть Q_1, ..., Q_m — все графически неравные предикатные переменные формулы Ф, x_1, ..., x_k — все графически неравные предметные переменные формулы Ф. Индукцию будем вести по логической длине формулы Ф.

1. Пусть Ψ — атомарная подформула $Q_i(x_j)$ формулы Ф. Построим обобщенный символ Венна из $k \cdot 2^m$ ячеек. Во всех ячейках области $X_j \cap Q_i$ поставим по одному кружку. Получим [Ψ] ∘.

2. Пусть *A, B* — любые подформулы формулы Ф. Предположим, что [*A*] и [*B*] построены. Построим [∃x A], [A ∨ B], [A & B], [¬ A], [A ⊃ B], [∀ x A], предполагая, что соответствующие формулы являются подформулами формулы Ф.

1) Применим к [*A*] операцию *Oi*. Получим [∃xA].
2) Применим к [*A*] и [*B*] операцию *Oь*. Получим [A ∨ B].
3) Применим к [*A*] и [*B*] операцию *0%*. Получим [A & B].
4) Применим к [*A*] операцию O_7. Получим [¬ A].
5) [A ⊃ B] определим как [¬A ∨ B].
6) [∀xA] определим как [¬∃x¬A].

В силу соответствия между формулами исчисления одноместных предикатов и диаграммами Венна можно говорить о графическом (диаграммном) построении исчисления одноместных предикатов.

4.3.4. Решение проблемы разрешения в логике одноместных предикатов с помощью диаграмм Венна

Определение. Формулу Ψ, дизъюнктивно соответствующую диаграмме [Ф], будем называть *дизъюнктивной квазиприведенной формулой* и обозначать $\Psi_Ф$, где Ф — формула.

Можно доказать следующие предложения:

3.3. $Ф \equiv \Psi'_Ф$, где Ф — произвольная формула.

3.4. Если диаграмма $[\Phi]$ $\circ * \textcircled{4}$ не содержит кружков, то формула Ψ_Φ есть дизъюнктивная приведенная форма формулы Ф.

3.5. Для каждой формулы Ф существует равносильная (относительно общезначимости) дизъюнктивная приведенная формула Ψ.

Сформулируем предложение 3.2 на языке диаграмм Венна:

Формула Ф тождественно ложна тогда и только тогда, когда на диаграмме [Ψ], где Ψ — равносильная (относительно общезначимости) формуле Ф дизъюнктивная приведенная формула, находится одна $* \bigcirc$-ломаная или содержащая 2^m знаков $\textcircled{4}$, которые лежат в различных ячейках, или на которой находится только два знака $*$, $\textcircled{4}$, которые лежат в одной ячейке.

3.6. Для каждой формулы Ф существует равносильная (относительно общезначимости) конъюнктивная приведенная формула Ψ.

В силу теоремы 3.5. построим $[F] * \textcircled{4}$, применим к $[F] * \textcircled{4}$ следующую операцию O_8.

Операция O_8. Пусть $L_1, \ldots, L_n$ — все $* \textcircled{4}$-ломаные диаграммы. Построим диаграммы $B_{0,i}$, $i = 1, \ldots, n$, каждая из которых содержит только одну ломаную L_i. Уничтожим на каждой диаграмме $B_{0,i}$ прямые, соединяющие знаки $*$, $\textcircled{4}$, получим диаграммы $B_{1,i}$.

К диаграмме $B_{1,1}$ применим последовательно операции O_1, O_2, O_3, получим диаграмму B_1. К диаграммам B_1 и $B_{1,2}$ применим операцию Q_6 получим диаграмму Венна B_2. И так далее. К диаграммам B_{n-1}, $B_{1,n}$ применим операцию O_6, получим диаграмму Венна B_n.

Операция O_8 соответствует переходу от дизъюнктивной приведенной формы к конъюнктивной приведенной форме данной формулы. F эквивалентна Ψ, где Ψ — конъюнктивная приведенная формула, конъюнктивно соответствующая диаграмме B_n. Полученную диаграмму B_n будем называть *приведенной диаграммой Венна.*

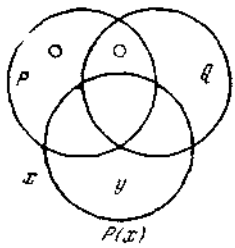
Сформулируем предложение 3.1 для диаграмм Венна:

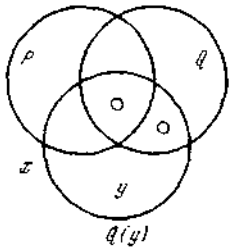
Формула Ф универсально общезначима тогда и только тогда, когда в соответствующей приведенной диаграмме Венна находится только одна $* \textcircled{4}$-ломаная, или содержащая только 2^m звездочек, которые лежат в различных ячейках, или на которой находится только два знака $*$, $\textcircled{4}$, лежащие в одной ячейке.

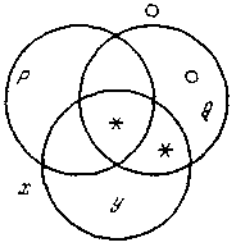
Например, пусть Ф есть формула

$$(\forall x \forall y (\neg P(x) \vee Q(y)) \supset \forall x \exists y (P(x) \supset Q(y))).$$

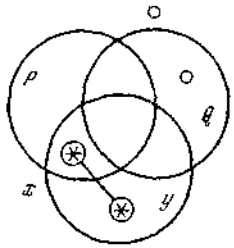
На рис. 81 приведено построение ее диаграммы Венна, приведенная диаграмма дана на рис. 82.



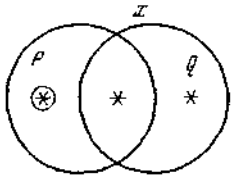




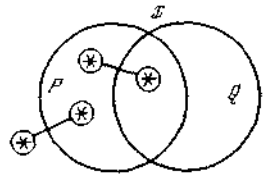
$\exists y(P(x) \to Q(y)) \equiv \exists y(\neg P(x) \vee Q(y))$



$\forall y(\neg P(x) \vee Q(y)) \equiv \exists y \neg(\neg P(x) \vee Q(y))$



$\forall x \exists y(P(x) \to Q(y))$



$\forall x \forall y(\neg P(x) \vee Q(y))$

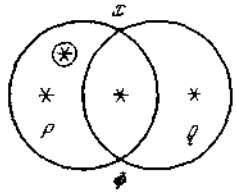


Рис. 81

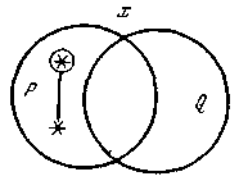


Рис. 82

Полученная приведенная диаграмма Венна содержит только одну ломаную, на которой имеется только два знака $*$ $\textcircled{*}$, лежащие в одной ячейке; следовательно, формула Ф универсально общезначима.

Ясно, что с увеличением числа переменных, особенно предметных, наглядность метода диаграмм уменьшается. Поэтому следует обратить внимание на классы формул, для которых метод диаграмм Венна наиболее прост и нагляден. Эти классы — следующие.

I. В формулу Ф входит только одна предметная переменная и не входят логические знаки $\forall$ и $\exists$; формула Ф фактически есть формула исчисления высказываний.

II. Формула Ф имеет вид $\exists x \Psi$, где Ψ формула класса I с предметной переменной x. Тогда на диаграмме [Ф] находятся только $*$-ломаные, т. е. [Ф] фактически является диаграммой $[\Phi]_*$; на каждой ломаной лежит только один знак $*$, и в каждой ячейке диаграммы может быть только один знак $*$.

III. Формула Ф имеет вид $\forall x \Psi$, где Ψ — формула класса I с предметной переменной x. Тогда на диаграмме [Ф] находятся только $\textcircled{*}$-ломаные, т. е. диаграмма [Ф] совпадает с диаграммой $[\Phi]_{\textcircled{*}}$; на каждой ломаной лежит только один знак $\textcircled{*}$, и в каждой ячейке диаграммы может быть только один знак $\textcircled{*}$.

Для записи информации, заданной прежде всего формулами этого класса, Венн и ввел свои диаграммы.

4.3.5. Обзор простык логических следствий из посылок, выразимых на языке формул исчисления одноместных предикатов, с помощью диаграмм Венна

Пусть $\Phi_1, \ldots, \Phi_k$ — любые формулы. Формулу Ψ будем называть *простым логическим следствием* носылок $\Phi_1, \ldots, \Phi_k$ в том и только в том случае, если Ψ есть дизъюнкция базисных высказываний формулы $(\Phi_1 \& \ldots \& \Phi_k)$ и их отрицаний и если Ψ не поглощается никаким другим логическим следствием посылок $\Phi_1, \ldots, \Phi_k$ того же вида.

Силлогистическая формула. Пусть $\Phi_1, \ldots, \Phi_k$ — любые формулы, все предметные переменные которых связаны; $Q_1, \ldots, Q_m$ — все графически неравные предикатные переменные формул Φ_i, $i = 1, \ldots, k$; $\Psi_1, \ldots, \Psi_{2^m}$ — все графически неравные базисные высказывания формул Φ_i, содержащие только одну

предметную переменную формул Φ_i. Начертим приведенную диаграмму Венна *B* для записи информации, заданной в виде формулы $(\Phi_1 \,\&\, \ldots \,\&\, \Phi_k)$. Пусть Ψ — формула, конъюнктивно соответствующая диаграмме *B*. Формулу *F* будем называть силлогистической, если результат всех возможных выявлений и поглощений, проведенных в формуле Ψ, есть формула *F*.

Теорема 3.7: Формула F_i является простым логическим следствием посылок Φ_j, $j = 1, \ldots, k$, тогда и только тогда, когда F_i является конъюктивным членом силлогистической формулы *F*. Следовательно, на приведенной диаграмме Венна, соответствующей формуле *F*, каждой $*$ ⊛-ломаной соответствует простое логическое следствие посылок $\Phi_1, \ldots, \Phi_k$.

Преобразование формулы Ψ в формулу *F* можно проводить на приведенных диаграммах Венна: операция поглощения есть операция O_2. Операция выявления есть следующая операция O_9.

Операция O_9. Пусть *B* — диаграмма Венна.

1. Предположим, что среди ломаных существует, по крайней мере, одна пара ломаных L_1 и L_2 такая, что в одной ячейке β одновременно находятся знаки $*$, ⊛, принадлежащие разным ломаным. Начертим на *B* новую ломаную L_3: во всех ячейках, кроме β, через которые проходят L_1 и L_2, поставим знаки, графически равные знакам на L_1 и L_2 в ячейке β знаков на L_3 нет.
2. Если среди ломаных не существует таких пар ломаных L_1 и L_2, что в одной ядейке одновременно находятся знаки $*$, ⊛, принадлежащие разным ломаным, то результатом операции O_9 является диаграмма *B*.

Операция O_9 соответствует закону выявления.

В исчислении одноместных предикатов простое логическое следствие посылок $\Phi_1, \ldots, \Phi_k$ определено как дизъюнкция базисных высказываний (и их отрицаний) формулы $(\Phi_1 \,\&\, \ldots \,\&\, \Phi_k)$; при этом базисные высказывания, зависящие от всех графически различных предикатных переменных формулы $(\Phi_1 \,\&\, \ldots \,\&\, \Phi_k)$, играют роль «кирпичей», из которых строятся логические следствия посылок $\Phi_1, \ldots, \Phi_k$. Однако в исчислении предикатов возможны и дальнейшие упрощения, которые зависят от связей между базисными высказываниями формулы $(\Phi_1 \,\&\, \ldots \,\&\, \Phi_k)$. Приведем два характерных примера.

Пример 1. Формула Φ, конъюнктивно соответствующая диаграмме, изображенной на рис. 83, является силлогистической; ее предикатные переменные: *X, Y, Z, W,* предметная — *p*. Простые логические следствия формулы Φ:

1. $\Psi_1 \overline{\circ} \neg \exists p (X(p) \& Y(p) \& \neg Z(p) \& \neg W(p))$,
2. $\Psi_2 \overline{\circ} \neg \exists p (X(p) \& Y(p) \& Z(p) \& \neg W(p))$,
3. $\Psi_3 \overline{\circ} \neg \exists p (X(p) \& Y(p) \& \neg Z(p) \& W(p))$,
4. $\Psi_4 \overline{\circ} \neg \exists p (X(p) \& Y(p) \& Z(p) \& W(p))$.

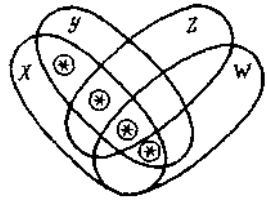


Рис. 83

Так как в каждой ячейке класса *XY* находятся знаки $\circledast$, то конъюнкция формул $\Psi_1, \ldots, \Psi_4$ эквивалентна формуле

$$\Psi \overline{\circ} \neg \exists p (X(p) \& Y(p)), \quad (\Psi_1 \& \ldots \& \Psi_4) \equiv \Psi.$$

Нетрудно убедиться, что Ψ — логическое следствие формулы Φ, а формулы $(\Psi_i \& \Psi)$, $i=1, \ldots, 4$, эквивалентны Ψ. Таким образом, формула Ψ является более «простым» логическим следствием формулы Φ, хотя в Ψ не входят предикатные переменные Z и W формулы Φ.

Заметим, что диаграмму, изображенную на рис. 83, можно рассматривать как приведенную диаграмму конъюнкции посылок:

$$\Phi_1 \overline{\circ} \forall p (X(p) \supset ((Y(p) \& Z(p)) \vee \neg Y(p))),$$
$$\Phi_2 \overline{\circ} \forall p ((X(p) \& Y(p) \& Z(p)) \supset W(p)),$$
$$\Phi_3 \overline{\circ} \neg \exists p (W(p) \& X(p) \& Y(p) \& Z(p)).$$

Эти посылки соответствуют условию задачи 8.1 в п.1.4.1.6. Знаки $\circledast$ диаграммы, изображенной на рис. 83, соответствуют пустым ячейкам диаграммы, представленной на рис. 43.

Пример 2. Формула Φ, конъюнктивно соответствующая диаграмме, изображенной на рис. 84, является силлогистической.

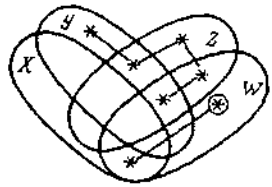


Рис. 84

На диаграмме находится только две $*$ $\circledast$ ломаных, поэтому

$$\Psi_1 \overline{\circ} (\exists p (\neg X(p) \& Y(p) \& Z(p) \& W(p)) \vee \exists p (\neg X(p) \& \neg Y(p) \& Z(p) \& W(p)) \vee \exists p (\neg X(p) \& \neg Y(p) \& Z(p) \& \neg W(p)) \vee \exists p (\neg X(p) \& Y(p) \& Z(p) \& \neg W(p)) \vee \exists p (\neg X(p) \& Y(p) \& \neg Z(p) \& \neg W(p)))$$

и

$$\Psi_2 \overline{\circ} (\exists p (X(p) \& \neg Y(p) \& \neg Z(p) \& W(p)) \vee \neg \exists p (\neg X(p) \& \neg Y(p) \& \neg Z(p) \& W(p)))$$

суть все простые логические следствия формулы Ф. Так как в каждой ячейке класса $\overline{X}Z$ находятся знаки $*$, принадлежащие одной ломаной, то логическое следствие Ψ_1 можно упростить в силу эквивалентности

$$(\exists p (\neg X(p) \& Y(p) \& Z(p) \& W(p)) \vee \exists p (\neg X(p) \& \neg Y(p) \& Z(p) \& W(p)) \vee \exists p (\neg X(p) \& \neg Y(p) \& Z(p) \& \neg W(p)) \vee \exists p (\neg X(p) \& Y(p) \& Z(p) \& \neg W(p))) \equiv \exists p (\neg X(p) \& Z(p)).$$

Мы получим:

$$\Psi_1 \equiv (\exists p (\neg X(p) \& Z(p)) \vee \exists p (\neg X(p) \& Y(p) \& \neg Z(p) \& \neg W(p))).$$

И далее, поскольку для любых формул A и B верно, что $(A \vee (B \& \neg A)) \equiv (A \vee B)$,

то

$$\Psi_1 \equiv (\exists p (\neg X(p) \& Z(p)) \vee \exists p (\neg X(p) \& Y(p) \& \neg W(p)))).$$

[выявления и поглощения проводятся в области действия квантора существования формулы

$$\exists p ((\neg X(p) \& Z(p)) \vee (\neg X(p) \& Y(p) \& \neg Z(p) \& \neg W(p))),$$

при этом можно воспользоваться методами, изложенными в п.1.4.2.6].

4.4. Диаграммы Венна в формальных нейронных схемах

4.4.1. Формальные нейроны Мак-Каллока

Формальный нейрон Мак-Каллока (для краткости — нейрон) является моделью, отражающей деятельность нервной клетки живых организмов. Он описывается следующим образом:

1. Нейрон имеет тело (изображается на рис. 85 и следующих треугольником), входы a_1, ..., a_n (предполагается, что все входы разные) и один выход. Входы нейрона не находятся на теле и не касаются его. Выход расположен на теле нейрона.

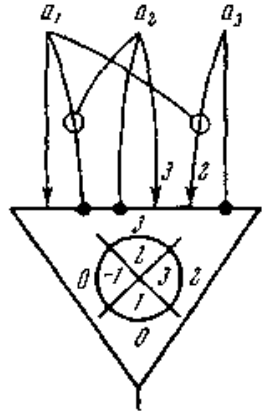


Рис. 85

2. Входы и выход могут находиться только в одном из двух состояний — возбужденном или невозбужденном.

3. Нейрон может иметь возбуждающие, тормозящие и запрещающие волокна, которые начинаются на его входах и кончаются на его теле. Каждое волокно может активироваться (возбуждаться) тогда и только тогда, когда возбужден вход, на котором оно начинается. Возбуждающее волокно (оканчивающееся на рисунках стрелкой на теле нейрона) вводит (в активном состоянии) положительную единицу возбуждения (+1). Тормозящее волокно (на рис. 85 и следующих оканчивается точкой на теле нейрона) вводит отрицательную единицу возбуждения (—1). Запрещающее волокно (на схеме оканчивается петлей на другом, запрещаемом волокне) предотвращает поступление

сигнала (возбуждения) по запрещаемому волокну. В работах Мак-Каллока и его последователей рассматривается только случай, когда запрещающие волокна оканчиваются на возбуждающих и тормозящих волокнах (при этом запрещающее и запрещаемое волокна начинаются на разных входах), и устанавливается, что волокна могут ветвиться, но не могут объединяться.

4. Сигналы могут проходить только в одном направлении — от входов к выходу (на рисунках — сверху вниз). При передаче сигналов через синапс (место контакта волокна с телом нейрона) получается задержка на единицу времени.

5. Нейрон обладает порогом Θ, который выражается целым числом. Изменение порога во времени задается конечным набором различных целых чисел $\{\Theta_1, \ldots, \Theta_r\}$, расположенных (для определенности) в порядке убывания. В каждый данный момент времени t порог Θ принимает одно из значений Θ_j из набора $\{\Theta_1, \ldots, \Theta_r\}$.

6. Выход нейрона возбуждается в момент времени t только в том случае, если сумма положительных и отрицательных единиц возбуждаения, поступивших на тело в момент времени t — 1 (с учетом действия запрещающих волокон), не меньше Θ_i — значения порога в момент t.

В каждый данный момент времени t входы нейрона $A\ (a_1, \ldots, a_n)$ образуют определенную последовательность из n нулей и единиц, где единица на i-ом месте обозначает возбужденность входа a_i, нуль на j-ом месте —невозбужденность входа a_j. Эту последовательность из n нулей и единиц будем называть элементарной, или входной, последовательностью переменных $a_1, \ldots, a_n$.

Каждой элементарной последовательности поставим во взаимно однозначное соответствие ячейку символа Венна n переменных. При этом будем предполагать, что номер ячейки, записанной в двоичной системе, совпадает с рассматриваемой элементарной последовательностью. Например, для $n = 3$ (рис. 29) числа, написанные в ячейках плоскости, можно воспринимать, как соответствующие входные последовательности нейрона, который имеет три входа.

Порог в нейрона в момент времени t принимает определенное значение Θ_i. Нейрон возбуждается в рассматриваемый момент времени t, если сумма возбуждающих и тормозящих единиц, соответствующих входной последовательности в момент времени t — 1 (с учетом действия всех запретов), не меньше Θ_i.

Если значение порога Θ фиксировано ($\Theta = \Theta_0$), то нейрон возбуждается или нет в зависимости только от входных последовательностей. Функционирование нейрона с n входами при фиксированом значении

его порога, $\Theta = \Theta_0$, можно описать на языке диаграмм Венна в классическом исчислении высказываний: если нейрон возбуждается при данной входной последовательности *s,* то в s-ой ячейке символа Венна *n* переменных ставится точка; если нейрон не возбуждается при данной входной последовательности *s,* то s-ая ячейка символа Венна *n* переменных пуста.

При изменении значения порога Θ диаграмма Венна может изменяться, и для каждого значения порога Θ может быть вычерчена своя диаграмма Венна. Например, функционирование нейрона, изображенного на рис. 85 при Θ =3, описывается диаграммой нейрона

(00010010),

а при Θ = 0 — диаграммой

(111110 11).

Из нейронов строятся сети. В сетях Мак-Каллока нейроны располагаются по рангам (рядам); в последнем ранге находится один,— называемый выходным — нейрон; во веех остальных рангах — *n* нейронов, где *n* — число входов каждого нейрона сети; входы первого ранга независимы (их число равно *n);* входами нейронов *(i* + 1)-го ранга являются выходы нейронов *i*-го ранга, *i* = 1, ..., *q* — 1, где *q* — количество рангов в сети *(q*> 0).

Описанные сети Мак-Каллока будем называть одновыходными регулярными сетями. Пример регулярной сети нейронов с одним выходом приведен на рис. 86.

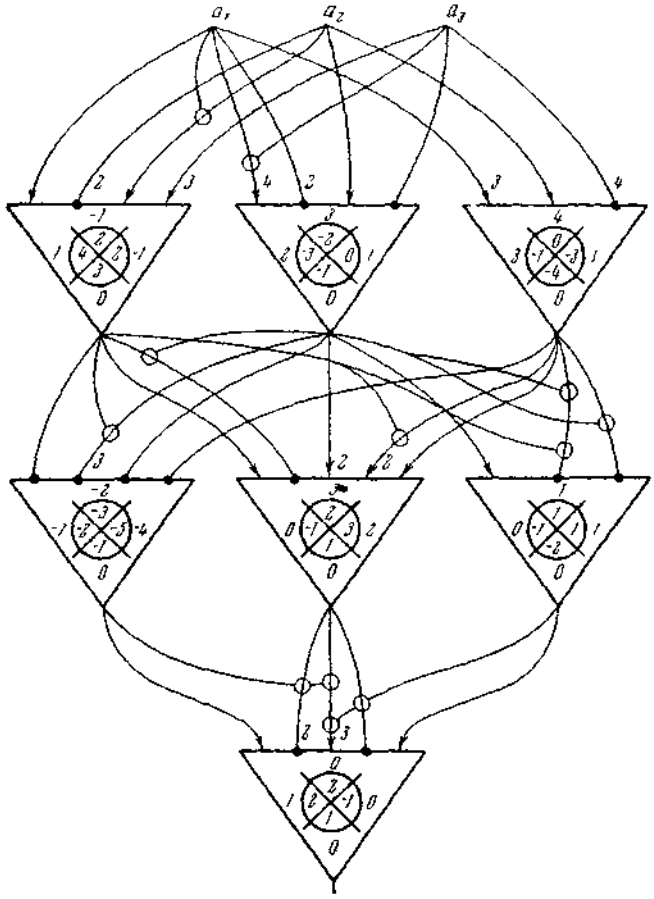


Рис. 86

Определения. 1. Символ Венна *n* переменных, во всех ячейках которого, кроме ячейки номер нуль, стоят (по одному) произвольные целые числа (положительные, отрицательные и нули), а в ячейке номер нуль обязательно находится нуль, будем называть *пороговой диаграммой n переменных.*

Например, на теле нейрона, изображенного на рис. 85, расположена пороговая диаграмма трех переменных.

2. Символ Венна *n* переменных, в каждой ячейке которого находится одно из чисел $1, 2, \ldots, m$, $m \leqslant 2^n$, будем называть *порядковой диаграммой n переменных.*

Обозначения: D^{θ} — пороговая, D^{π} — порядковая диаграммы, иногда в скобках справа будем указывать число переменных или сами переменные, например, $D^{\pi}(n)$ или $D^{\pi}(a_1, \ldots, a_n)$ — порядковая диаграмма n переменных.

Записывать диаграммы будем не только с помощью символа Венна, но и в линейной форме. Например, (0, 2, -3, -1,2, 4, -1,1, —1,1, -4, -2, -1, 1, -2, 0) -линейная запись пороговой диаграммы четырех переменных.

Установим соответствие между пороговыми, порядковыми и вероятностными диаграммами n переменных, с одной стороны, и формальными нейронами — с другой.

а) *Пороговые диаграммы нейронов.* Дан нейрон $A\ (a_1, \ldots, a_n)$.

Построим символ Венна n переменных; в ячейке, соответствующей элементарной последовательности S, поставим число γ_S, равное разности между числами, соответствующими активированным возбуждающим и тормозящим волокнам элементарной последовательности S с учетом действия запретов.

Пороговая диаграмма нейрона характеризуется тем, что в нулевой ячейке символа Венна стоит нуль — $\gamma_0 = 0$ —, а в остальных ячейках — целые числа (возможны и нули). Следовательно, пороговая диаграмма нейрона с n входами есть пороговая диаграмма n переменных.

Если порог нейрона Θ фиксирован: $\Theta = \Theta_0$, то по пороговой диаграмме этого нейрона можно однозначно определить, как те элементарные последовательности, при которых нейрон возбуждается, так и те, при которых он не возбуждается: если $\gamma_S \geqslant \Theta_0$, то нейрон возбуждается; если $\gamma_S < \Theta_0$, то нейрон не возбуждается (см. построение пороговой диаграммы нейрона), $S = 0, 1, \ldots, 2^n - 1$.

Следовательно, пороговая диаграмма нейрона позволяет представить полную картину его функционирования. По пороговой диаграмме можно, кроме того, определить, во-первых, те значения его порога, при которых он всегда возбуждается (при любой из 2^n элементарных последовательностей), во-вторых, те значения его порога, при которых он не возбуждается, и, в-третьих, те значения его порога, при которых для некоторых элементарных последовательностей он возбуждается, а для других — не возбуждается; в первом случае соответствующая диаграмма Венна полна (в каждой ячейке находится точка), во втором случае — пуста, в третьем — имеет как пустые ячейки, так и ячейки, содержащие точки.

б) *Порядковые диаграммы нейронов.* На порядковой диаграмме n переменных можно указать порядок (отсюда название диаграммы) появления точек на диаграммах Венна n переменных,

описывающих работу нейрона при уменьшении значений порога, начиная с пустой диаграммы Венна. Например, порядок появления точек на диаграммах нейрона (рис. 85) при уменьшении значения порога, начиная с $\Theta = 3$, описывается на одной порядковой диаграмме трех переменных: (4, 3, 2, 1 4, 5, 1, 2).

в) *Вероятностные диаграммы нейронов.* Функционирование сети формальных нейронов с входами $a_1, \ldots, a_n$ (в частности, одного нейрона) в зависимости от элементарных последовательностей переменных $a_1, \ldots, a_n$ (совпадающих с входами) можно описывать на вероятностных диаграммах n переменных.

Пусть для каждого нейрона сети задан свой интервал изменения порога. Тогда для данного выхода сети можно построить следующим образом вероятностную диаграмму n переменных. Единица, стоящая в i-ой ячейке символа Венна n переменных, обозначает, что, если входы сети образуют i-ую элементарную последовательность, то для любых значений порогов всех нейронов сети из указанных для них интервалов сеть возбуждается. Нуль, стоящий в j-ой ячейке символа Венна n переменных, обозначает, что если входы сети образуют j-ую элементарную последовательность, то для любых значений порогов всех нейронов сети из указанных для них интервалов сеть не возбуждается. Буква p, стоящая в k-ой ячейке символа Венна n переменных, обозначает, что, если входы сети образуют k-ую элементарную последовательность, то существуют такие значения порогов всех нейронов сети из указанных для них интервалов, что сеть возбуждается, и существуют такие значения порогов у всех нейронов сети из указанных для них интервалов, что сеть не возбуждается.

На языке теории вероятностей единица в i-ой ячейке вероятностей диаграммы означает, что сеть из формальных нейронов возбуждается с вероятностью 1, когда входы сети образуют i-ую элементарную последовательность; нуль в j-ой ячейке вероятностной диаграммы означает, что сеть возбуждается с вероятностью 0, когда входы сети образуют j-ую элементарную последовательность; буква p в k-ой ячейке вероятностной диаграммы говорит, что при k-ой входной последовательности сеть возбуждается с вероятностью r, $0 < r < 1$.

Описанную диаграмму будем называть результирующей вероятностной диаграммой нейронной сети при заданных интервалах изменения нейронов. Методы построения результирующих вероятностных диаграмм разбирались ранее.

Таким образом, понятие «диаграмма» в формальных нейронных схемах получает дальнейшее расширение. Для описания функционирования формальных нейронов используются диаграммы

Венна в классическом исчислении высказываний, вероятностные, пороговые и порядковые диаграммы *n* переменных. Мы не будем подробно описывать теорию формальных нейронных схем, отсылая читателя к соответствующей литературе, остановимся только на некоторых ее аспектах.

4.4.2. Синтез оптимальных формальных нейронов по пороговым диаграммам *n* переменных

Аналитические выражения *n* переменных. В ряде работ задание нейронов тесно связано с их геометрическим представлением. При увеличении количества входов и волокон наглядность геометрического изображения резко уменьшается. Поэтому — а также в целях уточнения понятия «формальный нейрон» — вводятся аналитические выражения *n* переменных.

Определение. Пусть A — набор переменных $a_1, ..., a_n$, H — целое число $(H \neq 0)$.

1. Если $\alpha \in A$, то $H\alpha$, считаем *аналитическим выражением* переменных $a_1, ..., a_n$ (в дальнейшем для краткости — выражением).
2. Если $\alpha \in A$, то (1 — *а)* считаем квазивыражением.
3. Если B — квазивыражение, $\alpha \in A$, то $B\,(1 - \alpha)$ считаем квазивыражением.
4. Если B — квазивыражение, $\alpha \in A$, то $H\alpha B$ считаем выражением, $(1 - \alpha)\ B$ — квазивыражением первого рода.
5. Если B — квазивыражение первого рода, D — или квазивыражение или квазивыражение первого рода, то BD и DB считаем квазивыражениями первого рода.
6. Если B — квазивыражение первого рода, $\alpha \in A$, то $(1 - \alpha B)$ считаем квазивыражением первого рода.
7. Если B — квазивыражение первого рода, $\alpha \in A$, то $H\alpha B$ считаем выражением.
8. Если B, D—выражения, то $B+D$ считаем выражением. Например, $a_1 - 8a_2\,(1 - a_1\ (1 - a_3\,(1 - a_2)))$— аналитическое выражение четырех переменных $a_1,...,a_4$ дерево его построения приведено на рис. 87, в точках ветвления указаны номера шагов индуктивного определения.

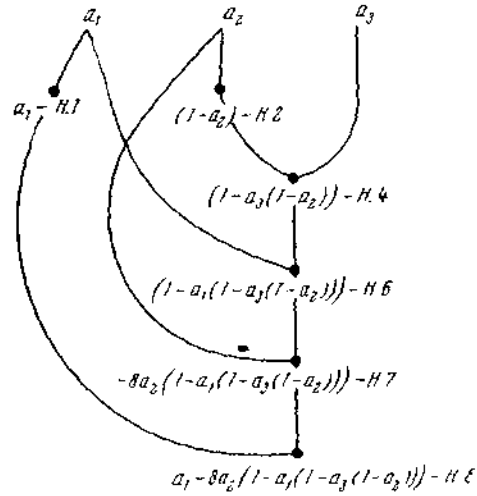


Рис. 87

Между аналитическими выражениями n переменных и формальными нейронами с n входами можно установить взаимно однозначное соответствие. Мы не будем разбирать этот вопрос подробно, заметим лишь, что: (а) можно указать индуктивный способ построения нейронов по заданным аналитическим выражениям, (б) каждому слагаемому аналитического выражения можно поставить в соответствие только одну из ветвей, оканчивающихся на теле нейрона (вместе со всеми связанными с ней запретами); при этом сомножители $(1 - a_j)$ соответствуют запрещающим ветвям, сомножители a_i — ветвям, оканчивающимся на теле нейрона, коэффициенты H указывают веса соответствующих ветвей. Таким образом, задание аналитического выражения позволяет полностью определить нейрон, т. е. указать его входы, типы ветвей волокон и их распределение по входам.

Коэффициенты аналитического выражения Ω_n могут быть неопределены (тогда мы записываем их в виде букв с индексами). Например,

$$H_1a_1 + H_2a_2 + H_3a_3 + H_{1(2)}a_1(1 - a_2) + H_{1(3)}a_1(1 - a_3) + \\ + H_{2(1)}a_2(1 - a_1) + H_{2(3)}a_2(1 - a_3) + H_{3(1)}a_3(1 - a_1) + \\ + H_{3(2)}a_3(1 - a_2) + H_{1(2,3)}a_1(1 - a_2) \times \\ \times (1 - a_3) + H_{2(1,3)}a_2(1 - a_1)(1 - a_3) + H_{3(1,2)}a_3(1 - a_1) \times \\ \times(1 - a_2) + H_{1(2(3))}a_1(1 - a_2(1 - a_3)) + H_{2(1(3))} \times \\ \times a_2(1 - a_1(1 - a_3)) + H_{3(1(2))}a_3(1 - a_1(1 - a_2)) \quad (4.1)$$

аналитическое выражение переменных a_1, a_2, a_3 с неопределенными коэффициентами $H_1, \ldots H_{3(1(2))}$. В (4.1) есть члены, соответствующие ветвям типа «запрет запрета» (запрещающая ветвь оканчивается на другой запрещающей ветви).

Синтез оптимальных формальных нейронов. В п.1.4.4.1 изложен диаграмматический анализ формальных нейронов. Интересна и обратная задача: по данной диаграмме n переменных (одного из рассмотренных выше видов) построить нейрон с n входами, диаграмма — указанного типа — которого совпадает с заданной.

Задача сформулирована для нейронов Мак-Каллока. При ее постановке для обобщенных нейронов приходится задавать несколько диаграмм, соответствующих различным моментам времени.

Предложим общий метод синтеза формальных нейронов по пороговым диаграммам n переменных. Для простоты ограничимся формальными нейронами Мак-Каллока. Если нам задана не пороговая диаграмма, а, скажем, вероятностная, то синтез нейрона можно начинать с построения соответствующей пороговой диаграммы.

При решении задачи синтеза введение аналитических выражений позволяет рассматривать произвольные сочетания ветвей нейронов. Синтез начинается с задания аналитического выражения с неопределенными коэффициентами. Требуется найти числовые значения неопределенных коэффициентов. Перебирая все 2^n различные входные последовательности и подставляя их члены вместо соответствующих переменных a_i в аналитическое выражение с неопределенными коэффициентами, получим систему 2^n —1 линейных алгебраических уравнений (в правых частях уравнений стоят числа γ_j данной пороговой матрицы) и одно тождество 0 = 0 (для входной последовательности из n нулей — 0...0). Например, при n = 3 без ветвей типа «запрет запрета» мы получим систему уравнений с 12-ю неизвестными:

$$\left.\begin{aligned}
H_3 + H_{3(1)} + H_{3(2)} + H_{3(1,2)} &= \gamma_1 \\
H_2 + H_{2(1)} + H_{2(3)} + H_{2(1,3)} &= \gamma_2 \\
H_2 + H_3 + H_{2(1)} + H_{3(2)} &= \gamma_3 \\
H_1 + H_{1(2)} + H_{1(3)} + H_{1(2,3)} &= \gamma_4 \\
H_1 + H_3 + H_{1(2)} + H_{3(2)} &= \gamma_5 \\
H_1 + H_2 + H_{1(3)} + H_{2(3)} &= \gamma_6 \\
H_1 + H_2 + H_3 &= \gamma_7
\end{aligned}\right\} \qquad (4.2)$$

Если полученная таким образом система уравнений противоречива, то задача синтеза при взятом аналитическом выражении не имеет

решения,— необходимо перейти к другому аналитическому выражению.

Предположим, что выбранное аналитическое выражение Ω_n с неопределенными коэффициентами $H_{\alpha_1}, \ldots, H_{\alpha_l}$ таково, что система относительно $H_{\alpha_1}, \ldots, H_{\alpha_l}$ непротиворечива. Тогда для любого ее решения (в силу взаимно однозначного соответствия между нейронами и аналитическими выражениями) конструируется нейрон, пороговая матрица которого совпадает с заданной. Если же система уравнений является неопределенной, то, естественно, возникает вопрос о нахождении среди ее решений оптимального (относительно некоторых параметров). Рассмотрим два критерия оптимальности:

1. *Минимум общего числа волокон нейрона.* Требуется найти такое решение системы уравнений, на котором

$$W = \sum_{i=1}^{l} \overline{\alpha_i} |H_{\alpha_i}|,$$

где $\overline{\alpha_i}$ равно количеству номеров переменных, входящих в α_i (например, если α_i есть 1 (2 (3), 4 (1, 2)), то $\overline{\alpha_i} = 6$), достигает минимума *(W* min).

2. *Минимум общего числа ветвей нейрона.* Требуется найти решение системы уравнений, при котором

$$V = \sum_{i=1}^{l} \overline{\alpha_i} \operatorname{sign}^2 H_{\alpha_i}$$

принимает минимальное значение *V* min.

Заметим, что в обоих случаях речь идет об оптимальных решениях относительно выбранного аналитического выражения с неопределенными коэффициентами $H_{\alpha_1}, \ldots, H_{\alpha_l}$.

В раде работ для построения оптимальных в первом смысле нейронов используются средства линейного программирования. Методы построения оптимальных во втором смысле нейронов еще полностью не изучены.

Изложим способ синтеза оптимальных нейронов путем перебора возможных значений *W* или *V*. Для сокращения перебора могут быть использованы предлагаемые ниже таблицы.

Очевидно: $W_{\min} \geqslant 0$, $V_{\min} \geqslant 0$. Оптимальное решение в одном из указанных смыслов можно найти перебором возможных значений *W* или *V,* начиная с нуля. На каждом шаге проверяется, имеет ли соответствующая система решение. Если при *W=L* (соотв. *V—L)* система непротиворечива, а для любого *W, W < L* (соответственно, *V, V < L)* система противоречива, то $W_{min} = L$ *(*$V_{mim} = L$*).* При переборе

значений W (соответственно, V) можно пользоваться таблицей. Последняя строится по матрице системы уравнений: число строк ее равно числу уравнений, число столбцов — на единицу больше числа различных переменных; если элемент $c_{i,j}$ матрицы равен нулю, то в соответствующей ячейке таблицы (находящейся в j-ом столбце и в i-ой строке) ставится нуль; если $c_{i,j}$ =1, то соответствующая ячейка таблицы оставляется пустой; в последнем столбце располагаются числа γ_i.

Решение системы сводится к заполнению пустых ячеек такими целыми числами, что сумма чисел в каждой i-ой строке равна числу γ_i и в каждом j-ом столбце во всех пустых ячейках одновременно ставятся равные между собою числа.

Например, при n = 3 (без ветвей типа «запрет запрета») общее число ветвей нейрона записывается в виде:

$$V = \operatorname{sign}^2 H_1 + \operatorname{sign}^2 H_2 + \operatorname{sign}^2 H_3 + 2\,(\operatorname{sign}^2 H_{1(2)} + \operatorname{sign}^2 H_{1(3)} + \operatorname{sign}^2 H_{2(2)} + \operatorname{sign}^2 H_{2(3)} + \operatorname{sign}^2 H_{3(1)} + \operatorname{sign}^2 H_{3(2)}) + 3\,(\operatorname{sign}^2 H_{1(2,3)} + \operatorname{sign}^2 H_{2(1,3)} + \operatorname{sign}^2 H_{2(1,2)}).$$

Таблица имеет 13 столбцов (табл. 1 — см. матрицу системы (4.2)).

Таблица 1

H_1	$H_{1(2)}$	$H_{1(3)}$	$H_{1(2,3)}$	H_2	$H_{2(1)}$	$H_{2(3)}$	$H_{2(1,3)}$	H_3	$H_{3(1)}$	$H_{3(2)}$	$H_{3(1,2)}$	γ_i
0	0	0	0	0	0	0	0					γ_1
0	0	0	0					0	0	0	0	γ_2
0	0	0	0			0	0			0	0	γ_3
				0	0	0	0	0	0	0	0	γ_4
		0	0	0	0	0	0		0		0	γ_5
	0		0		0		0	0	0	0	0	γ_6
	0	0	0		0	0	0		0	0	0	γ_7

Если $\gamma_1 = \gamma_2 = \gamma_4 = 1$, $\gamma_3 = \gamma_5 = \gamma_6 = \gamma_7 = 0$, то с помощью таблицы (табл. 2) получаем решение

$$H_{1(2,3)} = H_{2(1,3)} = H_{3(1,2)} = 1,$$

$$H_1 = H_{1(3)} = H_2 = H_{2(1)} = H_{2(3)} = H_3 = H_{3(1)} = H_{3(2)} = 0,$$

по которому синтезируется нейрон с $V = 9$ (рис. 88).

Таблица 2

H_1	$H_{1(2)}$	$H_{1(3)}$	$H_{1(2,3)}$	H_2	$H_{2(1)}$	$H_{2(3)}$	$H_{2(1,3)}$	H_3	$H_{3(1)}$	$H_{3(2)}$	$H_{3(1,2)}$	γ_i
0	0	0	0	0	0	0	0	0	0	0	1	1
0	0	0	0	0	0	0	1	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0

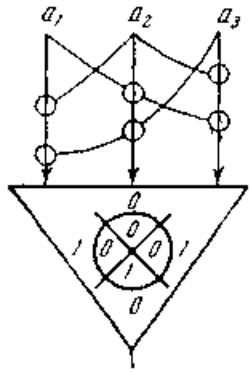


Рис. 88

Если в аналитическое выражение не входят только ветви типа «запрет запрета», то

$$0 \leqslant V\min \leqslant n \cdot 2^{n-1}.$$

Отметим, что введение ветвей типа «запрет запрета» может привести к уменьшению общего числа ветвей нейрона. Например, нейрон с $V = 3$ (рис. 89), имеющий ветвь

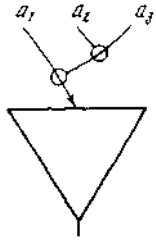


Рис. 89

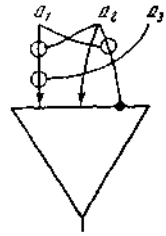


Рис. 90

типа «запрет запрета», и нейрон без ветвей этого типа с $V = 6$ (рис. 90) эквивалентны, т. е. имеют одну и ту же пороговую матрицу

(0, 0, 0, 0, 1, 0, 1, 1).

4.4.3. Надежные сети формальных нейронов

Каждая сеть состоит из нескольких рядов (рангов) нейронов. Входами сети являются все различные входы нейронов первого ранга, не совпадающие с выходами каких-либо нейронов сети; выходами — все выходы нейронов последнего ранга. Входами нейронов, начиная со второго ранга, могут являться только входы сети или выходы некоторых нейронов сети (возможен случай, когда выход нейрона совпадает с его входом). Все входы нейрона нумеруются числами, которые проставляются слева от них. Для однозначности иногда справа от номеров входов нейрона в круглых скобках мы будем ставить номер ранга r и номер i нейрона в ранге. Например, 2 (5,3) — второй вход третьего нейрона в пятом ранге, $r = 5$, $i = 3$ (на рисунках ранги нумеруются сверху вниз, нейроны — слева направо). В случае, когда входы расположены линейно (слева направо) и когда не может быть их разночтения, числа — номера входов — мы будем опускать. В качестве примера см. нейронные сети, изображенные на рис. 85, 86.

Сеть нейронов будем называть *сетью с обратной связью,* если в ней имеется, по крайней мере, один нейрон $A_{r,i}$, $r \geq 1$ с входами, которые совпадают с выходами нейронов $(r+i)$- го ранга, $i \geq 0$.

Например, на рис. 91 изображен трехвходовый нейрон с обратной связью.

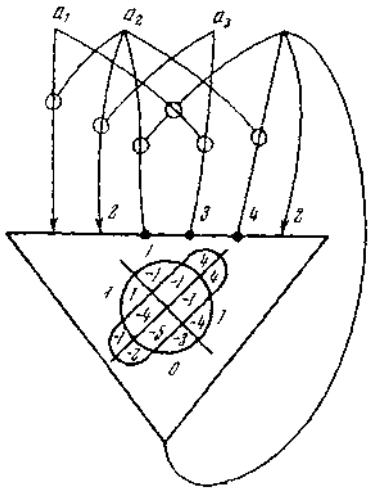


Рис. 91

На теле нейрона расположена пороговая диаграмма четырех переменных (выход нейрона рассматривается как его четвертый вход).

Для каждого нейрона сети можно построить диаграмму Венна в исчислении высказываний (бинарную матрицу) при фиксированном значении порога или вероятностную диаграмму (матрицу) при указанном интервале изменения порога. В результате получаем сеть диаграмм Венна (бинарных матриц) или вероятностных диаграмм (матриц), соответствующую данной седи нейронов. Например, на рис. 58 дана сеть диаграмм Венна, соответствующая двухранговой сети нейронов (рис. 92) при $\Theta_{1,1}=0, \Theta_{1,2}=1$, $\Theta_{2,1}=-1$ (первый ранг сети диаграмм соответствует входам нейронной сети).

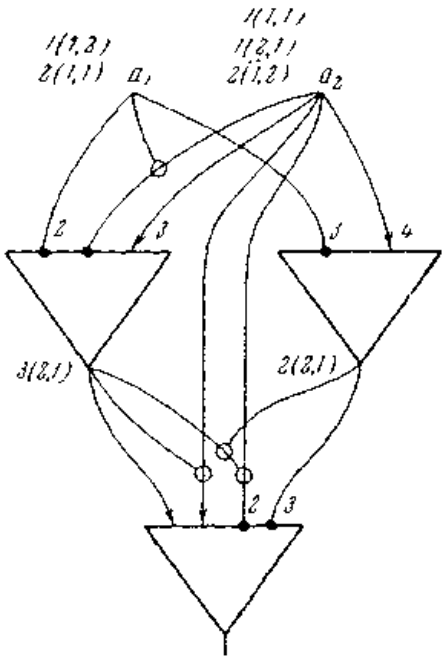


Рис. 92

Правила функционирования сетей бинарных и вероятностных диаграмм описаны ранее. Отметим, что при построении сети диаграмм, соответствующей заданной сети нейронов, можно рассматривать следующие два случая:

1. Прохождение сигнала по волокнам не зависит от времени, т. е. поведение окончания волокна исследуется в тот же момент времени *t,* в который входы образуют данную элементарную последовательность. При этом можно выделить более простой подслучай, когда не учитывается и синаптическая задержка.

2. Вершины волокон играют роль задержек сигналов на единицу времени. При этом в момент времени $t+i$ конструируется определенная сеть диаграмм. Ее поведение можно исследовать или независимо от времени, или в равноотстоящие моменты времени (начальным моментом является момент $t+i$). Таким образом, второй случай в свою очередь разбивается на два подслучая.

Ясно, что первый подслучай, как и случай 1, относится только к сетям без обратных связей. Если не учитывать синаптические задержки, то сети без обратных связей представляют собой однотактные автоматы.

Для любой сети бинарных матриц без обратных связей можно найти эквивалентную ей регулярную сеть (способ регуляризации изложен в п.1.4.2.7).

Следовательно, любая нерегулярная нейронная сеть без обратных связей может быть заменена эквивалентной ей регулярной сетью из формальных нейронов.

Как подчеркивалось ранее, надежность сети из не вполне надежных элементов может быть обеспечена за счет избыточности последних. Так как на одном формальном нейроне можно реализовать несколько неэквивалентных между собой формул исчисления высказываний (без обратных связей число таких формул не превосходит 2^{n+1}, введение обратных связей позволяет увеличить их количество до 2^{n+1}, см. п.1.4.4.4), то формальные нейроны можно использовать в качестве не вполне надежных элементов при конструировании надежных схем описанного ранее вида.

4.4.4. Формальные нейроны с обратными связями

Пороговые матрицы формальных нейронов с обратными связями. Формальный нейрон с обратными связями характеризуется тем, что от его выхода могут быть направлены волокна, принадлежащие этому же нейрону, т. е. выход нейрона в этом случае можно рассматривать как его вход, который мы будем называть *несобственным* (а остальные входы—*собственными).* Будем предполагать, что все входы нейрона $a_1, ..., a_n$ различны; несобственный вход обозначим через a_{n+1}**.** Например, формальный нейрон, изображенный на рис. 91, имеет три собственных входа и один несобственный.

В начальный момент времени t могут быть возбуждены некоторые из собственных входов нейрона. Таким образом, моменту времени t можно ставить в соответствие пороговую матрицу $(n+1)$-ой переменной; при этом несобственный вход a_{n+1} считается фиктивным. Вход a_i нейрона мы называем *фиктивным,* если на этом входе не начинаются волокна нейрона.

Нетрудно видеть, что если вход a_i нейрона фиктивен, то все соседние относительно входа a_i строки или столбцы пороговой матрицы попарно совпадают, и, наоборот, если все соседние относительно влода a_i строки или столбцы пороговой матрицы попарно совпадают, то пороговую матрицу можно реализовать на нейроне, у которого входа$_i$ фиктивен.

Две строки (столбца) матрицы будем называть соседними относительно входа a_i, если соответствующие им входные

последовательности отличаются друг от друга только тем, что в одной из них на *i*-ом месте находится нуль, а в другой — единица.

Пусть все соседние относительно входа a_i строки или столбцы пороговой матрицы *(n* + 1)-ой переменной попарно совпадают. Для построения *(n* + 1)-входового нейрона с фиктивным входом a_i в матрице можно вычеркнуть из каждой пары соседних относительно входа a_i строк или столбцов по одной строке или одному столбцу; в результате мы получим пороговую матрицу *n* переменных; по ней строится нейрон, к входам которого можно добавить фиктивный вход a_i (на вопросах синтеза нейронов по пороговым матрицам остановимся ниже).

Пороговую матрицу, соответствующую моменту времени *t,* мы в дальнейшем будем называть *i*-пороговой матрицей.

В момент времени *t*+1 могут возбуждаться как собственные входы нейрона, так и несобственный вход. Возбуждение несобственного входа зависит от информации, поступившей по возбужденным волокнам на тело нейрона в момент времени *t,* и от значения порога нейрона в момент времени *t* + 1. Т. е. моменту *t* + 1 соответствует *(t+1)*-пороговая матрица *(n* + 1)-ой переменной; вообще говоря, *(t* + 1)-пороговая матрица отличается от *t*-пороговой матрицы. В моменты времени *t+i, i* > 1, поведение нейрона описывается *(t* +1)-пороговыми матрицами, которые, как нетрудно показать, графически совпадают с *(t* + l)-пороговой матрицей. Из сказанного следует, что *n*-входовый формальный нейрон с обратными связями имеет две графически различные пороговые матрицы *(n* + 1)-ой переменной: *t*-пороговую и *(t* + 1)-пороговую. В *t*-пороговой матрице все соседние относительно входа a_{n+1} столбцы попарно совпадают, а все (2*r* — 1)-ые столбцы совпадают с соответствующими столбцами *(t* +1)-пороговой матрицы $(r = 1, \ldots, 2^k$, где 2^k — количество столбцов в матрицах). Следовательно, *t*-пороговую матрицу нейрона можно построить, если известна *(t* +1)-пороговая матрица,

Например, *t*-пороговая и *(t* + 1)-пороговая матрицы трехвходового нейрона, изображенного на рис.91, имеют, соответственно, вид:

$$\begin{pmatrix} 0 & 0 & -3 & -3 \\ 1 & 1 & -4 & -4 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{pmatrix}, \quad \begin{pmatrix} 0 & -2 & -3 & -5 \\ 1 & 4 & -4 & -1 \\ 1 & -1 & 1 & -4 \\ 1 & 4 & -1 & -1 \end{pmatrix}$$

(для сравнения с последней матрицей см. *(t* + 1)-пороговую диаграмму на теле нейрона).

Мы разобрали случай, когда несобственный вход имеет номер n +1. Аналогично можно описать функционирование формального нейрона, когда несобственный вход имеет номер i, $i < n+1$.

Синтез нейронов с обратными связями. Нетрудно заметить, что любую пороговую матрицу $(n + 1)$-ой переменной можно считать $(t + 1)$-пороговой диаграммой n-входового нейрона с несобственным входом a_i, $1 \leqslant i \leqslant n + 1$. Синтез формальных нейронов с обратными связями по заданной $(t$ +1)-пороговой матрице $(n+1)$-ой переменной можно осуществить изложенным выше способом. Несобственным входом можно считать любой из входов $a_1, ..., a_n$.

При фиксированном значении порога на нейроне реализуется некоторая бинарная матрица. Реализуемая на нейроне бинарная матрица может изменяться в зависимости от изменения значения порога. Все числа пороговой матрицы могут быть различными, поэтому на одном n-входовом нейроне без обратных связей можно реализовать (при изменении значения порога) не более 2^n+1 графически различных матриц n переменных.

Введение обратных связей позволяет увеличить количество графически различных бинарных матриц n переменных, реализуемых на одном нейроне, до 2^{nhl}. Действительно, расположим 2^{n+1} различных целых чисел на $(t$ +1)-пороговой диаграмме (матрице) A $(n$ +1)-ой переменной так, чтобы каждое из чисел в ячейках фигуры a_i, соответствующей несобственному входу, было больше каждого из чисел в остальных ячейках и чтобы числа γ^* и γ^{**} находились в ячейках, соседних относительно фигуры a_i (где γ^* — наименьшее из чисел в ячейках фигуры a_i, γ^{**} — наибольшее из чисел в ячейках, принадлежащих дополнению фигуры a_i). Тогда при изменении значения порога от γ^{**} +1 до γ_{min} на n-входовом нейроне $\mathfrak{A}$ с несобственным входом a_i, построенном: по $(t$ +1)-пороговой матрице A, в момент времени t может быть реализовано 2^n+1 графически различных бинарных матриц $(n$ переменных), где γ_{min} — наименьшее из чисел матрицы A. При уменьшении значения порога от γ_{max} до γ^*+1 (где γ_{max} — наибольшее из чисел матрицы A) на построенном нейроне $\mathfrak{A}$ в момент $t + i$, $i \geq 1$, может быть реализовано 2^n — 1 графически различных бинарных матриц $(n$ переменных), не совпадающих ни с одной из матриц, реализуемых в момент времени t.

Следовательно, на n-входовом нейроне с обратными связями, синтезированном по матрице A, можно реализовать 2^{n+1} графически различных бинарных матриц n переменных. Например, на трехвходовом нейроне с несобственным входом a_i и $(t$ + 1)-пороговой матрицей

$$\begin{pmatrix} 0 & 14 & 1 & 12 \\ 2 & 10 & 3 & 11 \\ 4 & 15 & 5 & 13 \\ 6 & 9 & 7 & 8 \end{pmatrix}$$

можно реализовать следующие бинарные диаграммы трех переменных. В момент *t,* управляя порогом в интервале [0, 8]:

(1, 1, 1, 1, 1, 1, 1, 1): (0, 0, 0, 0, 0, 0, 0, 1) (0, 0, 0, 0, 0, 0, 0, 0)
(0, 1, 1, 1, 1, 1, 1, 1); (0, 0, 1, 1, 1, 1, 1, 1) (0, 0, 0, 1, 1, 1, 1, 1)
(0, 0, 0, 0, 1, 1, 1, 1); (0, 0, 0, 0, 0, 1, 1, 1) (0, 0, 0, 0, 0, 0, 1, 1).

В момент *t* + 1, управляя порогом в интервале [9, 15]:

(1, 1, 1, 1, 1, 1, 0, 0); (1, 1, 0, 1, 1, 1, 0, 0); (1, 1, 0, 0, 1, 1, 0, 0)
(1, 0, 0, 0, 1, 1, 0, 0); (1, 0, 0, 0, 1, 0, 0, 0); (0, 0, 0, 0, 1, 0,0,0).

t-пороговая и пороговая матрицы нейрона имеют соответственно вид:

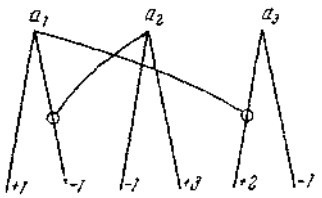
$$\begin{pmatrix} 0 & 0 & 1 & 1 \\ 2 & 2 & 3 & 3 \\ 4 & 4 & 5 & 5 \\ 6 & 6 & 7 & 7 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 2 & 3 \\ 4 & 5 \\ 6 & 7 \end{pmatrix}$$

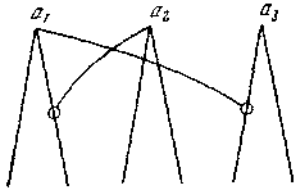
4.4.5. Алгебраические аспекты теории формальных нейронов

Условимся в графических изображениях нейронов опускать «тело» и окончания ветвей — «стрелки» и «точки». Справа от ветвей будем указывать их вес со знаком «плюс» для ветвей из возбуждающих волокон и знаком «минус» — для ветвей из тормозящих волокон (напомним, что «запрет» некоторой «проводящей» ветви веса ρ сам считается ветвью веса ρ). Назовем получившееся изображение *N*-графом нейрона. Опуская же в *N*-графе указания весов ветвей, получим граф нейрона.

На табл. 3 показаны, соответственно, *N*-граф и граф нейрона, изображенного на рис. 85.

Таблица 3





Таким образом, нейрон полностью определяется своим *N*-графом. Граф нейрона характеризует лишь его структуру; результаты работы нейронов с одинаковыми графами в общем случае различны.

В дальнейшем запись пороговой диаграммы нейрона мы будем производить следующим образом. Занумеруем входы справа налево. Состояние возбуждения входов будем символизировать единицей, состояние невозбуждения — нулем. При всякой комбинации состояний входов нули и единицы будем записывать в строку в порядке, соответствующем порядку входов.

Пусть число входов равно *k*. Тогда каждая комбинация состояний входов выражается *k*-значным двоичным числом. Всех таких комбинаций 2^k. Каждой из них соответствует некоторое целое число. Запишем эти числа в вектор-строку размерности 2^k (индексы координат его суть $0, 1, 2, \ldots, i, \ldots, 2^k - 1$). Присвоим числу такой индекс, двоичный код которого выражает комбинацию состояний входов, давшую это число. Это эквивалентно тому, как если бы входы занумеровать числами $2^0, 2^1, 2^2, \ldots 2^{k-1}$ и числу, получаемому при возбуждении некоторой комбинации входов, присваивать индекс, равный сумме номеров возбужденных входов. Итак, запишем диаграмму работы нейрона в виде:

$$(\mu)_k^0 \rightleftharpoons (\mu_0, \mu_1, \ldots, \mu_{2^k-1});$$

вектор $(\mu)_k^0$ получается из матрицы (2.2) при $m = 0$.

Из описания работы нейрона следует, что $\mu_0 = 0$; поэтому мы чаще будем пользоваться вектором

$$(\mu)_k \rightleftharpoons (\mu_1, \ldots, \mu_{2^k-1}).$$

Эти же диаграммы, записанные в виде столбцов, обозначим, соответственно, через $[\mu]_k^0$ (см. (2.2) при $k = 0$), $[\mu]_k$.

Основная решаемая ниже задача состоит в синтезе формального нейрона, реализующего заданную пороговую диаграмму.

1) *Свойства формальных нейронов. А).* Пусть даны *k*-нейроны (нейроны с *k* входами) N_{k1} и N_{k2} с диаграммами $(\mu)_{k1}$ и $(\mu)_{k2}$, соответственно. Определим сложение данных нейронов как объединение их в один *k*-нейрон N_k, при котором одноименные входы данных нейронов, объединяясь, дают вход нейрона N_k с тем же номером; все волокна и связи между ними каждого из данных нейронов переносятся и сохраняются в N_k. Граф нейрона N_k получается наложением графов нейронов N_{k1} и N_{k2} друг на друга с совпадением входов. При таком определении сложения нейронов диаграмма $(\mu)_k$ суммы нейронов равна покоординатной сумме диаграмм:

$$(\mu)_k = (\mu)_{k1} + (\mu)_{k2}.$$

B). Определим умножение *k*-нейрона *N* на целое число *b, bN,* как умножение на это число весов всех ветвей, входящих в данный нейрон. Структура самого нейрона и его граф при этом не меняются. Если нейрон умножить на целое число, то каждый элемент диаграммы нейрона умножится на это же число:

$$b\,(\mu)_k = (b\mu_1, \ldots b\mu_{2^k-1}).$$

Правило умножения на целое число $\pm b$ $(b > 0)$ является следствием правила сложения, примененного к *b* одинаковым *k*-нейронам *N* или —*N.* Нейрон — *N,* который можно назвать *противоположным* данному, получается из *N* заменой всех возбуждающих волокон тормозящими, а тормозящих — возбуждающими. Диаграмма нейрона —*N* есть — $(\mu)_k$:

$$-(\mu)_k = (-\mu_1, \ldots, -\mu_{2^k-1}).$$

Операции *A* и *B* позволяют говорить о линейных комбинациях формальных нейронов. Если нейрон N_k есть формальная линейная комбинация с целыми коэффициентами ρ_j *r* нейронов N_{kj} с диаграммами $(\mu)_{kj} = (\mu_{1,j}, \ldots, \mu_{2k-1,j})$:

$$N_k = \sum_{j=1}^{r} \rho_j N_{kj},$$

то *i*-тый элемент диаграммы $(\mu)_k$ нейрона N_k есть

$$\mu_i = \sum_{j=1}^{r} \rho_j \mu_{i,j}$$

C). Пусть дан нейрон N_k с диаграммой $(\mu)_k$. Образуем нейрон $\widetilde{N}_k$ добавлением нового $(k+1)$-го входа к входам N_k. Вход, от которого не исходит ни одного волокна, является фиктивным. В остальном строение нейрона остается тем же. Покажем, что диаграммы $(\mu)^0_{k+1}$ и $(\mu)_{k+1}$ нового нейрона $\widetilde{N}_k$ образуются из диаграмм $(\mu)^0_k$ и $(\mu)_k$ исходного нейрона N_k следующим образом:

$$(\mu)^0_{k+1} = ((\mu)^0_k,\ (\mu^0_k) = (0,\ (\mu)_k,\ 0,\ (\mu)_k),$$

$$(\mu)_{k+1} = ((\mu)_k,\ (\mu)^0) = ((\mu)_k,\ 0,\ (\mu)_k).$$

Действительно, пусть комбинации состояний входов нейрона N_k

$$x \overline{\circ} \alpha_k \alpha_{k-1} \ldots \alpha_1,$$

где $\alpha_i = 1$, если *i*-й вход возбужден, и $\alpha_i = 0$, если он не возбужден, соответствует в диаграмме $(\mu)_k$ кородината

μ_j (j — десятичный код двоичного числа α). Тогда комбинациям состояний входов в $\widetilde{N}_k$

$$0\alpha_k\alpha_{k-1}\ldots\alpha_1 \text{ и } 1\alpha_k\alpha_{k-1}\ldots\alpha_1$$

будут соответствовать координаты диаграммы $(\mu)_{k+1}$, равные μ_j, а их индексы будут, соответственно, j и $2^k + j$. Утверждение доказано.

D). Пусть теперь в нейроне $\widetilde{N}_k$ с фиктивным (k+1)-м входом наложены от этого нового входа запреты на все проводящие ветви нейрона N_k. Получившийся нейрон обозначим через $\overline{N}_k$, а его диаграммы через $\overline{(\mu)}_k^0$ и $\overline{(\mu)}_k$.

Тогда, как нетрудно видеть,

$$\overline{(\mu)}_k^0 = ((\mu)_k^0,\ (0)_k^0) = (0,\ (\mu)_k,\ 0,\ (0)_k),$$

$$\overline{(\mu)}_k = ((\mu)_k,\ (0)_k^0) = ((\mu)_k,\ 0,\ (0)_k).$$

Действительно, так как (k+1)-й вход находится сначала в невозбужденном состоянии, первые 2^k—1 координаты в $(\mu)_{k+1}$ образуют диаграмму $(\mu)_k$; координата с индексом 2^k нулевая, она соответствует единственному возбужденному (k+1)-му входу, от которого не исходит ни одной возбуждающей ветви; далее, (k +1)-й вход остается возбужденным все время и подавляет сигналы от остальных входов; таким образом, последние 2^k — 1 координат нулевые, вектор из них обозначен через $(0)_k$.

Свойства *A* и *B* составляют основу излагаемых ниже методов синтеза формальных нейронов; свойства *C* и *D* позволяют систематически применять индукцию при распространении этих методов на нейроны с любым числом входов.

2) *Схема-I синтеза формальных нейронов.* Пусть некоторый нейрон имеет диаграмму

$$(\mu)_k \overline{\circ} (\mu_1, \ldots, \mu_i, \ldots, \mu_{2^k-1}).$$

Представим $(\mu)_k$ в виде линейной комбинации:

$$(\mu)_k = + \begin{cases} \mu_1 \cdot (1, 0, 0, \ldots, 0, \ldots, 0) \\ \mu_2 \cdot (0, 1, 0, \ldots, 0, \ldots, 0) \\ \quad \ldots\ldots\ldots\ldots \\ \mu_i \cdot (0, 0, 0, \ldots, 1, \ldots, 0) \\ \quad \ldots\ldots\ldots\ldots \\ \mu_{2^k-1} \cdot (0, 0, 0, \ldots, 0, \ldots, 1). \end{cases}$$

В сжатой форме это представление можно записать так:

$$(\mu)_k = \sum_{i=1}^{2^k-1} \mu_i \cdot (I)_{ki},$$

где $(I)_{ki}$ — вектор размерности $2^k - 1$ со всеми нулевыми координатами, кроме *i*-й, равной +1. Вектор $(I)_{ki}$ будем считать диаграммой некоторого нейрона N_{ki}. Тогда в силу свойств *A* и *B,* образуя формальную линейную комбинацию

$$N_k = \sum_{i=1}^{2^k-1} \mu_i N_{ki},$$

мы получим один из *k*-нейронов, реализующих диаграмму $(\mu)_k$.

Итак, для построения нейрона по приведенной схеме, которую мы назовем «схема-1», достаточно уметь строить нейроны N_{ki}, реализующие диаграммы $(I)_{ki}$. Построение N_{ki} будем вести индукцией по *k*. Заметим, что

$$(I)_{11} \overline{\circ} (1).$$

Пусть при некотором *k* нейроны N_{ki} построены. Построим $N_{k+1,\,j}$. Добавим во всех нейронах N_{ki} по одному $(k+1)$-му входу и образуем нейроны $\tilde{N}_{ki}$ и $\overline{N}_{ki}$. В силу свойств *C, D,* имеем диаграммы этих нейронов:

$$D(\tilde{N}_{ki}) \overline{\circ} ((I)_{ki},\ 0,\ (I)_{ki}),\ D(\overline{N}_{ki}) \overline{\circ} ((I)_{ki},\ 0,\ (0)_k).$$

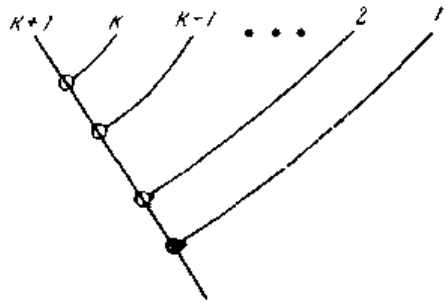
Легко заметить, что диаграммы $D(\overline{N}_{ki})$ суть векторы $(I)_{k+1,j}$ для $j \leqslant 2^k - 1$, а разности диаграмм— $D(\tilde{N}_{ki}) - D(\overline{N}_{ki})$ — суть векторы $(I)_{k+1,j}$ для $j = 2^k + i$. Отсюда следует, что

$$N_{k+1,j} \overline{\circ} \overline{N}_{kj},\ j \leqslant 2^k - 1;$$

$$N_{k+1,j} \overline{\circ} \tilde{N}_{ki} - \overline{N}_{ki},\ j = 2^k + i,\ i \leqslant 2^k - 1.$$

Таким образом построены все нейроны $N_{k+1,j}$, используемые в схеме-1, кроме $N_{k+1,2^k}$, реализующего диаграмму $(I)_{k+1,2^k}$. В этой диаграмме все координаты нулевые, исключая координату с индексом 2^k. Поэтому от *k* первых входов не должно исходить проводящих волокон. Если же возбужден один только $(k+1)$-й вход, то в теле нейрона мы получаем +1; это требует от $(k+1)$-го входа возбуждающего волокна. Так как, далее, при возбуждении любой комбинации из *k* входов вместе с $(k+1)$-м входом в теле нейрона мы получаем нули, то естественно наложить запрет от каждого из *k* первых входов на проводящее волокно от $(k+1)$-го входа. На табл. 4 приведен граф нейрона $N_{k+1,2^k}$.

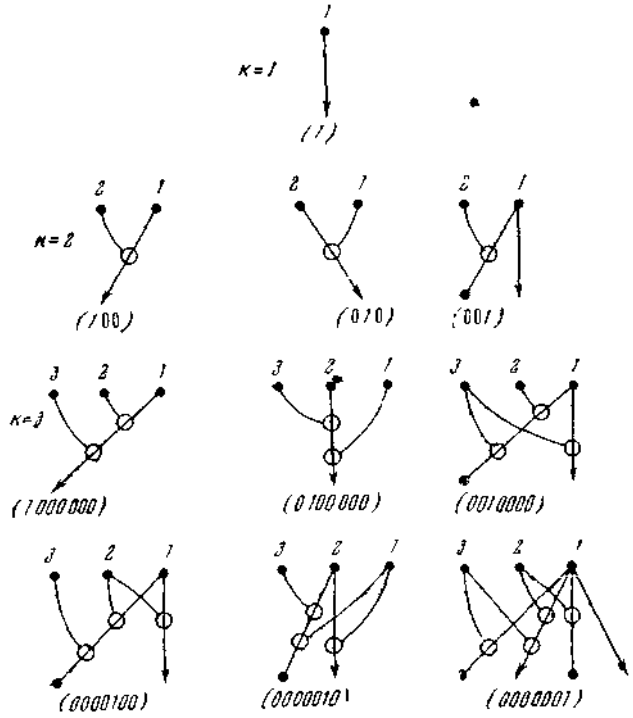
Таблица 4



Полученная система нейронов N_{ki}, строящаяся по индукции, позволяет синтезировать из них нейрон N_k с любой заданной диаграммой. Для синтеза нужно включить в общий нейрон N_k все нейроны N_{ki}, соответствующие ненулевым μ_i диаграммы $(\mu)_k$, и веса проводящих ветвей в N_{kl} взять равными именно μ_i.

На табл. 5 приведены нейроны N_{kl} для схемы-1, вместе со своими диаграммами $(I)_{kl}$ для $k = 1, 2, 3$. Так как веса всех ветвей равны + 1, то для указания знака здесь можно воспользоваться общепринятыми обозначениями — стрелкой и точкой соответственно.

Таблица 5



3) *О структуре формальных нейронов.* Рассмотрим нейроны, имеющие только одно проводящее (возбуждающее или тормозящее) волокно с произвольно наложенными на него запретами от других входов. Такие нейроны назовем *элементарными* или просто *элементами.* В силу свойств *A* и *B* каждый нейрон можно считать линейной комбинацией своих элементов; таким образом, структуру нейрона можно определить набором входящих в него элементов.

Рассмотрим для данного *k* полную систему элементов, то есть систему всех различных элементарных *k*-нейронов. Нейрон, содержащий все элементы полной системы, назовем полным (в смысле структуры). Свойство полноты нейрона геометрически можно выразить так: граф полного *k*-нейрона содержит граф любого *k*-нейрона.

Найдем число элементов L_k полной системы и число ветвей V_k полного нейрона, что дает максимальное число ветвей в *k*-нейроне. Для удобства вычислений возьмем число входов, равное n+1. Зафиксируем в $(n+1)$-нейроне произвольный вход и выберем все элементы с проводящей ветвью от этого входа. Число таких элементов l_{n+1} и число

ветвей в них v_{n+1} не зависят от номера фиксированного входа; очевидно,

$$L_{n+1} = (n+1)\, l_{n+1}, \quad V_{n+1} = (n+1)\, v_{n+1},$$

i запретов можно наложить на проводящую ветвь C_n^i способами, откуда $l_{n+1} = 2^n$ и $L_k = k \cdot 2^{k-1}$.

Число ветвей в элементе, содержащем i запретов, равно $i + 1$, то есть

$$v_{n+1} = \sum_{i=1}^{n} (i+1)\, C_n^i = \sum_{i=0}^{n} iC_n^i + 2^n.$$

$$iC_n^i = i\,\frac{n!}{i!\,(n-i)!} = n\,\frac{(n-1)!}{(i-1)!\,[(n-1)-(i-1)]!} = nC_{n-1}^{i-1},$$

$$v_{n+1} = n2^{n-1} + 2^n = (n+2)\,2^{n-1}, \quad V_k = k(k+1)\,2^{k-2}.$$

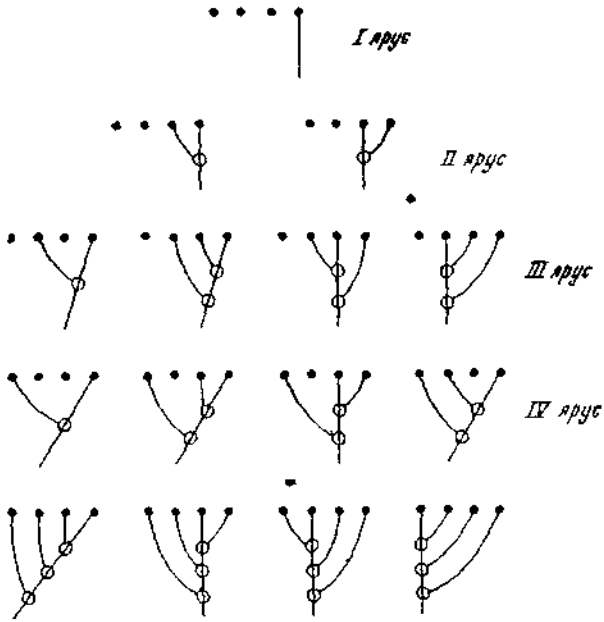
Изучим теперь структуру нейронов, синтезируемых по схеме-1. Речь идет о выделении системы S_k элементов (S_k-системы), входящих в общем случае в нейроны, синтезируемые по этой схеме.

При построении (k+1)-нейронов в схеме-1 использовались уже построенные k-нейроны, реализирующие диаграммы I_{ki}. Пусть система S_k-элементов построена. Рассматривая нейрон N_{ki} как линейные комбинации элементов S_k-системы и получая $N_{k+1,j}$ по указанным в пункте 2 правилам, мы замечаем, что каждый элемент S_k-системы порождает два элемента S_{k+1}-системы. В S_k-систему, кроме того, войдет образующий нейрон $N_{k+1,2^k}$, являющийся, очевидно, элементарным. Таким образом, число элементов p_k в S_k-системе удовлетворяет рекуррентному соотношению $p_{k+1} = 2p_k + 1$, $p_1 = 1$, откуда

$$p_k = 2^k - 1.$$

В таблице 6 показано построение системы S_k элементарных нейронов, используемых в схеме-1 (до $k = 4$). Построение ведется по «ярусам».

Таблица 6



В первом ярусе стоит один элементарный нейрон для $k = 1$. Каждый новый ярус получается из элементов предыдущих добавлением одного $(k+1)$-го входа ко всем элементам предыдущих ярусов и наложением запретов от него на проводящие ветви. В предыдущих ярусах при этом в каждом элементе добавляется новый фиктивный вход. В каждом ярусе последним является нейрон $N_{i,2^{i-1}}$, где i — номер яруса. Полученные таким образом элементы в $(k+1)$-м ярусе образуют S_{k+1} -систему. (Элементы в таблице даны в графах.)

Пусть R_k — максимальное число ветвей в нейроне, строящемся по схеме-1, т. е. R_k есть число ветвей во всех элементарных нейронах S_k -системы. Нетрудно установить, что

$$R_{k+1} = 2R_k + 2^k + k, \; R_1 = 1.$$

Отсюда следует, что

$$R_k = k \cdot 2^{k-1} + 2^k - (k+1).$$

Назовем *экономичностью алгоритма синтеза нейрона по пороговой диаграмме* (обозначение: $Э_k$) отношение числа ветвей в полном нейроне к максимальному числу ветвей в нейронах, синтезируемых по данному алгоритму. Для схемы-1

$$Э_k = \frac{V_k}{R_k}.$$

Для нашего примера:

$$\frac{k+1}{4} < \vartheta_k < \frac{k+1}{2}.$$

Замечание. Условимся каждому элементарному *k*-нейрону ставить в соответствие строку из *k* знаков, принимающих значения $0, \beta, \alpha$. При этом α обозначает вход (α -вход), от которого исходит проводящая ветвь; β — вход (β-вход), от которого наложен запрет на проводящую ветвь; О — фиктивный вход (0-вход). Для иллюстрации приведем элементуаные нейроны, изображенные на табл. 6, в символической записи:

I (000α)
II $(00\beta\alpha)$ $(00\alpha\beta)$
III $(0\beta0\alpha)$ $(0\beta\beta\alpha)$ $(0\beta\alpha\beta)$ $(0\alpha\beta\beta)$
IV $(\beta00\alpha)$ $(\beta0\beta\alpha)$ $(\beta0\alpha\beta)$ $(\beta\beta0\alpha)$ $(\beta\beta\beta\alpha)$
$(\beta\beta\alpha\beta)$ $(\beta\alpha\beta\beta)$ $(\alpha\beta\beta\beta)$.

4) *Схема-2 синтеза формальных нейронов.* Рассмотрим предварительно полное решение задачи о синтезе двухвходового нейрона. Полная система элементов в этом случае состоит из 4-х элементов: $(0\alpha), (\beta\alpha), (\alpha0), (\alpha\beta)$.

Оказывается, любое сочетание трех элементов из полной системы дает решение задачи о синтезе.

Возьмем, например, сочетание $\{(0\alpha), (\beta\alpha), (\alpha 0)\}$ и найдем веса ρ_1, ρ_2, ρ_3, с которыми нужно брать эти элементы в нейроне с диаграммой (μ_1, μ_2, μ_3).

Диаграмма (μ_1, μ_2, μ_3) есть, очевидно, сумма диаграмм взятых элементов, именно,

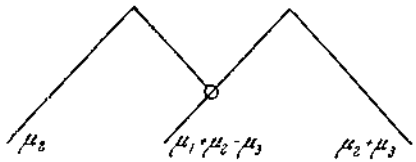
$$\begin{array}{l} \rho_1 0 \rho_1 \\ +\ 0\rho_2\rho_2 \\ \rho_3 00 \\ \hline \mu_1\mu_2\mu_3, \end{array} \quad \text{т. е.} \quad \begin{array}{l} \rho_1 + \rho_3 = \mu_1 \\ \rho_2 = \mu_2 \\ \rho_1 + \rho_2 = \mu_3, \end{array}$$

откуда

$$\begin{aligned} \rho_1 &= -\mu_2 + \mu_3 \\ \rho_2 &= \mu_2 \\ \rho_3 &= \mu_1 + \mu_2 - \mu_3. \end{aligned}$$

N-граф синтезированного нейрона имеет вид, показанный в табл. 7.

Таблица 7



Аналогично устанавливается разрешимость задачи синтеза для остальных сочетаний элементов. Заметим, что приведенное решение требует лишь четырех ветвей, по сравнению с пятью, получающимися в схеме-1. Поэтому предложим иную схему синтеза формальных нейронов — схему-2.

Диаграмму

$$(\mu)_3 \overline{\circ} (\mu_1, \mu_2, \ldots, \mu_7)$$

представим в виде:

$$(\mu)_3 = (\mu)_3^{\bullet} + (\widetilde{\mu})_2 + (\overline{\mu})_2,$$

где

$$(\mu)_3^{\bullet} \rightleftharpoons (0,\ 0,\ 0,\ \mu_4,\ \mu_4,\ \mu_4\ \mu_4),$$

$$(\bar{\mu})_2 \rightleftharpoons (\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3, 0, 0, 0, 0),\ (\widetilde{\mu})_2 \rightleftharpoons (\widetilde{\mu}_1,\ \widetilde{\mu}_2,\ \widetilde{\mu}_3,\ 0,\ \widetilde{\mu}_1,\ \widetilde{\mu}_2,\ \widetilde{\mu}_3),$$

причем

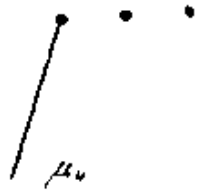
$$(\overline{\mu}_i = \mu_{4+i} - \mu_i\ (i = 1,\ 2,\ 3),\ \overline{\mu}_j = \mu_j - \widetilde{\mu}_j\ (i = 1, 2, 3),$$

т. е.

μ_1	μ_2	μ_3	μ_4	μ_5	μ_6	μ_7	
0	0	0	μ_4	μ_4	μ_4	μ_4	$(\mu)_3^{\bullet}$
+							
$\widetilde{\mu}_1$	$\widetilde{\mu}_2$	$\widetilde{\mu}_3$	0	$\widetilde{\mu}_1$	$\widetilde{\mu}_2$	$\widetilde{\mu}_3$	$(\widetilde{\mu})_2$
+							
$\overline{\mu}_1$	$\overline{\mu}_2$	$\overline{\mu}_3$	0	0	0	0	$(\overline{\mu})_2$.

Единственность такого представления очевидна. Для реализации диаграммы $(\mu)_3^{\bullet}$ достаточно элементарного нейрона $(\alpha 00)$ веса μ_4 (табл. 8).

Таблица 8



Диаграммы $(\tilde{\mu})_2$ и $(\bar{\mu})_2$ в силу свойств *C* и *D* (пункт 1)) реализуются нейронами $\tilde{N}_{21}$ и $\bar{N}_{22}$, соответственно; последние синтезируются в соответствии с изложенным выше решением задачи синтеза для двухвходовых нейронов. В этом случае для синтеза нейрона достаточно 12 ветвей.

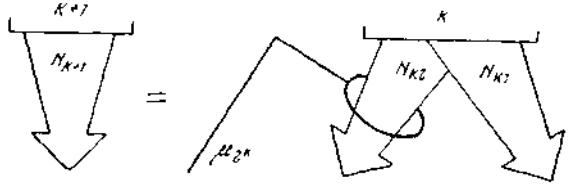
В общем случае произвольная диаграмма $(\mu)_{k+1}$ представляется в виде суммы диаграмм $(\mu)^{*}_{k+1}$, $(\tilde{\mu})_k$, $(\bar{\mu})_k$ следующим образом:

$$\begin{array}{l} \mu_1 \ldots \mu_{2^k-1} \quad \mu_{2^k} \quad \mu_{2^k+1} \ldots \mu_{2^{k+1}-1} \\ \hline 0 \ldots 0 \quad \mu_{2^k} \quad \mu_{2^k} \quad \ldots \mu_{2^k} \\ + \\ \tilde{\mu}_1 \ldots \tilde{\mu}_{2^k-1} \; 0 \quad \tilde{\mu}_1 \quad \ldots \tilde{\mu}_{2^k-1} \\ + \\ \bar{\mu}_1 \ldots \bar{\mu}_{2^k-1} \; 0 \quad 0 \quad \ldots 0, \end{array}$$

$$\tilde{\mu}_i = \mu_{2^k+i} \vee \mu_{2^k}, \quad \bar{\mu}_i = \mu_i \cdot \tilde{\mu}_i \quad (1 \leqslant i \leqslant 2^k - 1).$$

Пусть *k*-диаграммы $(\tilde{\mu}_1, \ldots, \tilde{\mu}_{2^k-1})$ и $(\bar{\mu}_1, \ldots, \bar{\mu}_{2^k-1})$ реализуются, соответственно, нейронами N_{k1} и N_{k2}, синтезированными по схеме-2. Тогда нейрон N_{k+1}, реализующий диаграмму $(\mu)_{k+1}$ и синтезированный также по схеме-П, может быть символически представлен в виде, показанном на табл.9.

Таблица 9



Символ, находящийся на табл. 9 слева от знака равенства, обозначает *(k* +1)-нейрон; общая петля запрета от *(k* + 1)-го входа на нейроне N_{k2} означает, что каждое проводящее волокно этого нейрона запрещено *(k* +1)-м входом.

Максимальное число ветвей в нейронах, строящихся по схеме-2, удовлетворяет, как легко видеть, соотношению:

$$R_{k+1} = 2R_k + 2^k, \quad R_1 = 1.$$

Отсюда $R_k = k \cdot 2^{k-1}$ и $\vartheta_k = \frac{k+1}{2}$, т. е. экономичность схемы-2 при небольших *k* значительно выше экономичности схемы-1.

Замечание 1. *Обобщенная схема.* Вообще говоря, в приведенной схеме можно вместо элементарного нейрона $(\alpha, 00\ldots0)$ веса μ_{2^k} брать произвольный элемент с *(k+1)*-м α-входом, обозначая, как и раньше,

диаграмму этого элементарного нейрона через (μ^*_{k+1}) (его ненулевые координаты, очевидно, равны μ_{2^k}):

$$(\mu)^*_{k+1} = (0, \ldots, 0, \mu^*_0, \mu^*_1, \ldots, \mu_{2^k-1});$$

тогда мы получим следующие соотношения для диаграмм $(\tilde{\mu})_k$ и $(\bar{\mu})_k$:

$$\tilde{\mu}_i = \mu_{2^k+i} - \mu^*_i, \quad \bar{\mu}_r = \mu_i - \tilde{\mu}_i \quad (0 \leqslant i \leqslant 2^k - 1).$$

Таким образом, если схемы синтеза формальных нейронов, отвечающие различным элементам с $(k+1)$-м α-входом, считать различными и обозначать их число через Z_{k+1}, то будем иметь

$$Z_{k+1} = Z_k^2 \cdot 2^k.$$

Как было упомянуто выше, для k=2 имеется четыре различных решения (схемы) задачи синтеза, т. е. Z_2=4. Можно показать индукцией, что

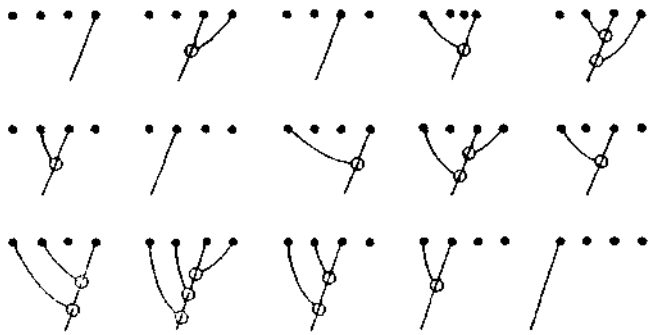
$$Z_k = \frac{2^{5 \cdot 2^{k-2}}}{2^{k+1}}.$$

Например, Z_3 = 64.

Структуру нейронов, синтезируемых по любой из описанных схем, легко получить аналогично тому, как это делалось для схемы-I (которая, кстати, входит в обобщенную схему).

Схема-II выделена благодаря некоторым свойствам, которые будут описаны ниже. Решение задачи синтеза по этой схеме, если двухвходовые нейроны на втором шаге синтеза имеют структуру, изображенную на табл. 6, а построение ведется по табл. 9 (при этом N_{k1} и N_{k2} имеют одинаковую структуру), мы назовем каноническим решением. В табл. 10 приведена система элементарных нейронов (в графах), используемых в каноническом решении до k = 4.

Таблица 10



В символической записи табл. 10 имеет вид:

(000α), $(00\alpha\beta)$, $(00\alpha 0)$, $(0\beta 0\bar{\alpha})$, $(0\beta\alpha\beta)$, $(0\beta\alpha 0)$,
$(0\alpha 00)$, $(\beta 00\alpha)$, $(\beta 0\alpha\beta)$, $(\beta 0\alpha 0)$, $(\beta\beta 0\alpha)$, $(\beta\beta\alpha\beta)$,
$(\beta\beta\alpha 0)$, $(\beta\alpha 00)$, $(\alpha 000)$.

Замечание 2. Рассмотрим случай нулевой диаграммы $(0)_k$. С помощью описанных схем синтеза формальных нейронов можно получить множество нетривиальных реализаций этой диаграммы. Тривиальным нейроном, реализующим нулевую диаграмму, является нейрон, все входы которого фиктивны. Нулевую диаграмму можно рассматривать как разность двух одинаковых произвольных диаграмм: $(0)_k = (\mu)_k - (\mu)_k$. Реализуя диаграмму $(\mu)_k$ дважды различными (то есть в разных схемах) нейронами N_{k1} и N_{k2} и беря их разность $N_k = N_{k1} - N_{k2}$, получим нетривиальный нейрон N_k, реализующий диаграмму $(0)_k$.

5) *Матричный подход.* По существу задача синтеза нейрона по заданной пороговой диаграмме сводится к системам линейных уравнений, решения которых дают значения весов элементарных нейронов, входящих в искомый нейрон. При этом сразу обнаруживается неоднозначность решения (см. п.1.4.4.2, 1.4.4.4). Действительно, полная система элементов k-нейронов содержат $L_k = k2^{k-1}$ элементов, т. е. в общем случае столько неизвестных величин — весов, с которыми нужно брать эти элементы,— сколько надо, чтобы получить данную диаграмму $(\mu)_k$, число координат которой 2^k—1 равно числу уравнений системы. Неравенство $k2^{k-1} > 2^{k-1}$ и влечет неоднозначность решения системы.

Построенные выше схемы синтеза приводят к мысли о том, что существуют минимальные системы элементарных нейронов, с помощью которых можно синтезировать нейрон, реализующий любую данную пороговую диаграмму. Такие системы мы назовем базами. Алгебраически это означает, что диаграммы элементарных нейронов, входящих в базу, образуют максимальную линейно независимую систему векторов размерности $2^k - 1$. Отсюда следует, что в базе в точности $2^k - 1$ элементов. Базами являются, например, системы, приведенные в таблицах 6 и 10.

Пусть дана некоторая база S_k и соответствующая ей максимальная линейно независимая система элементарных диаграмм. Элементы берутся с некоторыми весами ρ_i, что дает нейрон N_k. Если записать диаграммы элементов базы S_k как столбцы матрицы, обозначаемой через B_k, веса ρ_i записать в вектор-столбец $[\rho]_k$, то произведение $B_k [\rho]_k = [\mu]_k$ дает диаграмму нейрона N_k, записанную в столбец. Заметим, что это есть матричное решение прямой задачи: найти

пороговую диаграмму данного нейрона (причем даже для систем, не обязательно являющихся базами). Но обратная задача единственным образом решается только для баз и решение выражается с помощью матрицы M_k, обратной B_k:

$$[\rho]_k = M_k\,[\mu]_k, \quad M_k = B_k^{-1}.$$

Подробно базы изучаются в пункте 7. Так как матрицы B_k для каждой базы бинарные, т. е. состоят из нулей и единиц, то нахождение обратных матриц в каждом конкретном случае не представляет особых затруднений. В некоторых случаях, однако, возможно дать явное выражение матриц M_k, избежав фактического обращения B_k. Запишем в матричном виде каноническое решение задачи синтеза двухвходового нейрона (табл. 7):

$$\begin{aligned}\rho_1 &= -\mu_2 + \mu_3\\ \rho_2 &= \mu_1 + \mu_3 - \mu_3;\ \text{т. е.}\\ \rho_3 &= \mu_2\end{aligned}\quad \begin{pmatrix}\rho_1\\ \rho_2\\ \rho_3\end{pmatrix} = \begin{pmatrix}0 & -1 & 1\\ 1 & 1 & -1\\ 0 & 1 & 0\end{pmatrix}\begin{pmatrix}\mu_1\\ \mu_2\\ \mu_3\end{pmatrix}.$$

Так же легко находятся матрицы для других решений в случае двухвходовых нейронов. Обратимся теперь к схеме-II, символически изображенной в табл. 9. Число μ_2^k (вес единственной проводящей ветви от $(k+1)$-го входа) обозначим для удобства через ρ_3. Пусть нейроны N_{k1} и N_{k2} имеют, соответственно, диаграммы $(\mu)_{k1}$ и $(\mu)_{k2}$. Тогда диаграмма $(\mu)_{k+1}$ нейрона N_{k+1} сокращенно запишется как сумма:

$$\begin{aligned}(\mu)_{k+1} = +\,&((\mu)_{k1},\ 0,\ (\mu)_{k1})\\ &((\mu)_{k2},\ 0,\ (0)_k)\\ &((0)_k,\ \rho_3,\ (\rho_3)_k),\end{aligned}$$

где $(0)_k$ — нулевая строка, $(\rho_3)_k$ — строка чисел ρ_3 (обе строки размерности $2^k - 1$).

Нейроны N_{k1} и N_{k2} независимы в том смысле, что в схеме-2 их можно строить в различных базах. Пусть по предположению N_{k1} и N_{k2} реализованы в элементах баз S_k' и S_k'', соответственно, и веса элементов в выражениях N_{k1} и N_{k2} через эти базы задаются столбцами $[\rho]_{k1}$ и $[\rho]_{k2}$, т. е:

$$[\rho]_{k1} = M_k'\,[\mu]_{k1}, \quad [\rho]_{k2} = M_k''\,[\mu]_{k2},$$

где M_k' и M_k'' — матрицы, соответствующие базам S_k' и S_k'', считаются известными. Образуем из элементов баз S_k' и S_k'' системы $\tilde{S}_k'$ и $\tilde{S}_k''$ для каждого нейрона $N_k' \in S_k'$, беря $\tilde{N}_k' \in \tilde{S}_k'$, и для $N_k'' \in S_k''$, беря $\tilde{N}_k'' \in \tilde{S}_k''$. Системы $\tilde{S}_k'$ и $\tilde{S}_k''$ вместе с элементом веса ρ_3 (с единственной проводящей ветвью от $(k+1)$-го входа) образуют

систему S_{k+1}, являющуюся, очевидно, базой (n+1)-неронов. Элементы этой базы упорядочим следующим образом; этот порядок будем называть естественным.

Элементы системы $\tilde{S}'_k$ старше элементов системы $\tilde{S}''_k$, элемент веса ρ_3 — самый младший в S_{k+1}. Элементы в системах $\tilde{S}'_k$ и $\tilde{S}''_k$ считаются уже упорядоченными соответственно порядку в S'_k и S''_k. Таким образом, весовые коэффициенты элементов базы S_{k+1} в нейроне N_{k+1}, синтезируемом в элементах этой базы, запишутся в вектор-столбец $[\rho]_{k+1}$ так:

$$[\rho]_{k+1} = \begin{Bmatrix} [\rho]_{k1} \\ [\rho]_{k2} \\ \rho_3 \end{Bmatrix} = \begin{Bmatrix} M'_k [\mu]_{k1} \\ M''_k [\mu]_{k2} \\ \rho_3 \end{Bmatrix}.$$

С другой стороны, используя сокращенную запись диаграммы $(\mu)_{k+1}$ и записывая ее в столбец $[\mu]_{k+1}$, имеем:

$$[\mu]_{k+1} = \begin{Bmatrix} [\mu[_{k1} + [\mu]_{k2} \\ \rho_3 \\ [\mu]_{k2} + [\rho_3]_k \end{Bmatrix}.$$

Задача состоит в том, чтобы найти матрицу M_{k+1} по M'_k и M'_k ' из матричного уравнения

$$M_{k+1} [\mu]_{k+1} = [\rho]_{k+1}.$$

Очевидно, что последняя строка матрицы M_{k+1} имеет вид

$$((0)_k,\ 1,\ (0)_k).$$

Матрицу M_{k+1} ищем в виде:

$$M_{k+1} = \begin{pmatrix} M_{k1} & R_{k1} & M_{k2} \\ M_{k3} & R_{k2} & M_{k4} \\ 0_k & 1 & (0)_k \end{pmatrix},$$

где M_{kl} — клетки матрицы размерности $2^k - 1$, R_{kj} — вектор-столбцы той же размерности.

Запишем наше уравнение в развернутом виде:

$$\begin{pmatrix} M_{k1} & R_{k1} & M_{k2} \\ M_{k3} & R_{k2} & M_{k4} \\ (0_k) & 1 & (0_k) \end{pmatrix} \begin{pmatrix} [\mu]_{k1} + [\mu]_{k2} \\ \rho_3 \\ [\mu]_{k1} + [\rho_3]_k \end{pmatrix} = \begin{pmatrix} M'_k [\mu]_{k1} \\ M''_k [\mu]_{k2} \\ \rho_3 \end{pmatrix}.$$

В силу определения умножения матрицы на вектор это матричное выражение распадается на два:

$$M_{k1} [\mu]_{k1} + M_{k1} [\mu]_{k2} + \rho_3 R_{k1} + \\ + M_{k2} [\mu]_{k1} + M_{k2} [\rho_3]_k = M'_k [\mu]_{k1},$$

$$M_{k3}[\mu]_{k1} + M_{k3}[\mu]_{k2} + \rho_3 R_{k2} +$$
$$+ M_{k4}[\mu]_{k1} + M_{k4}[\rho_3]_k = M''_k[\mu]_{k2}.$$

В первом уравнении $[\mu]_{k2}$ в левой части входит со мноителем M_{k1}, в то время как правая часть не зависит от $[\mu]_{k2}$. Поэтому, в силу произвольного $[\mu]_{k2}$ и независимости между $[\mu]_{k1}$ и $[\mu]_{k2}$ имеем $M_{k1} \equiv 0$. Далее, вектор $[\rho_3]_k$ также независим от $[\mu]_{k1}$, откуда $\rho_3 R_{k1} + M_{k2}[\rho_3]_k = \bar{0}$ и $M_{k2} = M'_k$; наконец $R_{k1} = -M'_k[I]_k$, где $[I]_k$ — вектор, все координаты которого равны единице.

Во втором уравнении, рассуждая аналогично, получим:
$M_{k3}[\mu]_{k1} + M_{k4}[\mu]_{k1} = 0$, откуда $M_{k3} = -M_{k4}$; затем
$\rho_3 R_{k2} + M_{k4}[\rho_3]_k = 0$, т. е. $R_{k2} = -M_{k4}[I]_k$; наконец
$M_{k3}[\mu]_{k2} = M'_k[\mu]_{k2}$, т. е. $M_{k3} = M''_k$ и $R_{k2} = M'_k[I]_k$.

Замечание. В случае обобщенной схемы синтеза в базе S_{k+1} младший элемент веса ρ_3 имеет диаграмму не $((0)_k, \rho_3, (\rho_3)_k)$, а $((0)_k, \rho_3, (\rho_3)^*_k)$, где вектор $(\rho_3)^*_k$ имеет некоторые координаты — нулевые и остальные — равные ρ_3: $(\rho_3)^*_k = \rho_3 (I)^*_k$. Матричные уравнения при этом изменяются несущественно, а именно, вектор $[\rho_3]_k$ заменяется на $[\rho_3]^*_k$. В результате этого в решении претерпевают изменения только векторы R_{k1} и R_{k2}; именно, оказывается, что

$$R_{k1} = -M'_k[I]^*_k \text{ и } R_{k2} = M''_k[I]^*_k.$$

Рассмотрим теперь частный случай схемы-II, а именно, каноническое представление нейрона для случая, когда $M'_k = M''_k$, а начальная матрица $M_1 = (1)$. Как легко получить по предыдущему,

$$M_2 = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 1 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

Эта матрица нам уже встречалась. В решении матричного уравнения получаем:

$$M_{k1} = 0;\ M_{k2} = M_{k3} = -M_{k4} = M_k;$$
$$R_{k1} = R_{k2} = M_k[I]_k = R_k.$$

Найдем R_{k+1}, считая R_k известным из уравнения $R_{k+1} = M_{k+1}[I]_{k+1}$. Положим:

$$R_{k+1}=\begin{Bmatrix}[r]_{k1}\\ [r]_{k2}\\ 1\end{Bmatrix};\quad [I]_{k+1}=\begin{Bmatrix}[I]_k\\ 1\\ [I]_k\end{Bmatrix};\quad M_{k+1}=\begin{pmatrix}0 & -R_k & M_k\\ M_k & R_k-M_k & \\ (0)_k & 1 & (0)_k\end{pmatrix}.$$

Действуя, как раньше, имеем: $[r]_{k1}=-R_k+M_k[I]_k$, но $R_k=M_k[I]_k$, т. е. $[r]_{k1}\equiv[0]_k$, $[r]_{k2}=M_k[I]_k+$ $+R_k-M_k[I]_k=R_k$: таким образом,

$$R_{k+1}=\begin{Bmatrix}[0]_k\\ R_k\\ 1\end{Bmatrix}.$$

Так как $R_1=\{1\}$, то R_k можно описать как вектор размерности 2^k-1 с k последними координатами, равными единице.

Приведем вид матрицы M_3 для канонического представления нейрона; разбиение ее на клетки подчеркивает «происхождение» ее из матрицы M_2.

$$M_3=\left(\begin{array}{ccc|c|ccc}0&0&0&0&0&-1&1\\0&0&0&-1&1&1&-1\\0&0&0&-1&0&1&0\\\hline 0&-1&1&0&0&1&-1\\1&1&-1&1&-1&-1&1\\0&1&0&1&0&-1&0\\\hline 0&0&0&1&0&0&0\end{array}\right)$$

Матрицы M_k суть обратные к матрицам B_k, столбцы которых суть диаграммы элементарных нейронов баз S_k. Пусть база S_{k+1} — каноническая. Как указывалось, она строится как объединение систем $\bar{S}_k'$ и $\bar{S}_k''$ (с добавлением младшего элемента), причем должно быть $S_k'=S_k''=S_k$ — канонической. Учитывая естественный порядок элементов в S_k и S_{k+1}, получаем рекуррентное матричное соотношение:

$$B_{k+1}=\begin{pmatrix}B_k & B_k & [0]_k\\ (0)_k & (0)_k & 1\\ B_k & 0 & [I]_k\end{pmatrix},\quad B_1=(1).$$

при $k=2$

$$B_2=\begin{pmatrix}1&1&0\\0&0&1\\1&0&1\end{pmatrix},$$

а при $k = 3$

$$B_3 = \left(\begin{array}{ccc|ccc|c} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

6) *О полной системе элементов.* Рассмотрим полную систему элементарных k-нейронов [см. пункт 3)] и их пороговые диаграммы (т. е. диаграммы с координатой $\mu_0=0$), записанные по строкам в матрицу $\mathfrak{A}_k$ в некотором порядке, который определим ниже. Переходя к полной системе $(k+1)$-нейронов, образуем три группы элементов по типам. В первую группу включим $(k+1)$-элементы, получившиеся из k-элементов добавлением фиктивного $(k+1)$-го входа (т. е. элемента типа $\tilde{N}_k$ с $(k+1)$-ым 0-входом); их диаграммы вида $(\mu)_{k+1} = ((\mu)_k, 0, (\mu)_k)$ назовем симметричными. Во вторую группу включим $(k+1)$-элементы, получившиеся из k-элементов наложением запрета от $(k+1)$-го входа на проводящее волокно, т. е. элементы типа $\overline{N}_k$ с $(k-1)$-м β-входом; сами эти элементы и их диаграммы вида $((\mu)_k, 0, (0)_k)$ назовем косыми.

В третью группу включим элементы, никак не связанные с предыдущей системой, имеющие проводящее волокно от $(k+1)$-го входа, т. е. элементы с $(k+1)$-м α-входом; эти элементы и их диаграммы вида $((0)_k, 1, (\mu)_k)$ назовем отмеченными.

Отмеченных $(k+1)$-элементов будет $2^k = l_{k+1}$. Отбрасывая в отмеченных диаграммах нулевую левую грань, получим векторы размерности 2^k, из которых можно образовать матрицу A_k. В силу сказанного выше, мы можем теперь выразить матрицу $\mathfrak{A}_{k+1}$ из диаграмм $(k+1)$-элементов так:

$$\mathfrak{A}_{k+1} = \begin{pmatrix} \mathfrak{A}_k & \mathfrak{A}_k \\ \mathfrak{A}_k & 0 \\ 0 & A_k \end{pmatrix}.$$

Для того чтобы по этому рекуррентному выражению можно было строить полные системы элементов (в диаграммах), найдем связь между A_{k+1} и A_k. Пусть некоторый отмеченный $(k+1)$-элемент имеет в символической записи вид $(\alpha, \delta_1, \delta_2, \ldots, \delta_k)$, где δ_i

обозначает 0- или β-вход. Образуем из него два *(k* + 2)-элемента: $(\alpha, 0, \delta_1, \ldots, \delta_k)$ и $(\alpha, \beta, \delta_1, \ldots, \delta_k)$. В первом элементе при возбуждении α-входа, любом состоянии 0-входа и некоторой комбинации состояний входов δ_i мы получим тот же результат, который дает исходный *(k* + 1)-элемент при возбуждении *(k* + 1)-го α-входа и той же комбинации состояний входов δ_i. Это приводит к тому, что правая ненулевая грань первого элемента оказывается повторенной дважды правой ненулевой гранью исходного элемента. Ненулевая же грань второго элемента сама состоит из двух граней: левой, равной грани исходного элемента, и нулевой. Объединяя диаграммы первых элементов в одну группу, а вторых — в другую, получаем соотношение:

$$A_{k+1} = \begin{pmatrix} A_k & A_k \\ A_k & 0 \end{pmatrix} \quad A_1 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Определитель матрицы A_k, $k \neq 1$, очевидно, равен +1. Отметим, что $\mathfrak{A}_1 = (0,1)$; $\mathfrak{A}_2$ имеет вид:

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Проведенными рекуррентными выражениями и начальными матрицами порядок диаграмм в $\mathfrak{A}_k$ определяется однозначно. Такой порядок и соответствующий ему порядок в самой системе *k*-элементов мы назовем естественным.

Можно в полной системе *k*-элементов ввести и другой порядок, называемый (α, β)-порядком и определяемый так: а) всякий элемент с *i*-тым α-входом старше элемента с *(k* + 1)-м α-входом для элементов с фиксированным α-входом образуем β-шифр — *k*-значное двоичное число с единицами на местах, соответствующих β-входам, и нулями на остальных, элемент с меньшим β-шифром старше; β-шифр в десятичной системе, очевидно, равен $\sum 2^{k_\beta}$, где k_β — номера β-входов.

Порядок в полной системе элементов влечет порядок в базе, если в базе элементы расположены в той же последовательности, что и в полной системе. Например, при построении схемы-I, и схемы-II, и обобщенной схемы элементы располагались в естественном порядке. (α, β)-порядок удобен для произвольных баз, получаемых по другим схемам.

Представляется интересным следующее свойство отмеченных элементов и связанной с ними матрицы A_k. Нетрудно заметить

тривиальный факт: разность строк матрицы A_1 дает единственную строку матрицы $\mathfrak{A}_1$.

Докажем индукцией общий факт: каждая строка матрицы $\mathfrak{A}_k$ есть разность двух строк матрицы A_k.

Пусть для A_k определена система разностей U_k, дающая полную систему строк $\mathfrak{A}_k$. Возьмем матрицу A_{k+1} как объединение двух матриц: (A_k, A_k) и $(A_k, 0)$. Очевидно, что применение системы разностей U_k как оператора к минорам A_k этих матриц дает системы строк $(\mathfrak{A}_k, \mathfrak{A}_k)$ и $(\mathfrak{A}_k, 0)$. Наконец, образование новых разностей, которые все вместе можно записать как матричную разность $(A_k, A_k) — (A_k, 0)$, дает систему $(0, A_k)$. Итак, вычисляя разности в матрице A_{k+1}, мы получим все составляющие части матрицы $\mathfrak{A}_{k+1}$, входящие в ее рекуррентное представление.

Описанное свойство отмеченных элементов может быть использовано для построения несколько необычной схемы синтеза формальных нейронов. Пусть k-нейрон синтезирован из k-элементов и дает некоторую пороговую диаграмму $(\mu)_k^0$. Введем $(k + 1)$-вход и рассмотрим пары $(k + 1)$-входовых отмеченных элементов, разности которых, дают k-элементы взятого нейрона. Беря члены пар с соответствующими коэффициентами, мы получим $(k+1)$-входовый нейрон, имеющий диаграмму $((0)_\Sigma^0, (\mu)_k^0)$ и отличающийся тем, что все проводящие ветви его исходят из одного $(k + 1)$-го входа. Требуемую диаграмму мы получим на таком нейроне, если всякую комбинацию первых k входов возбуждать вместе с возбуждением $(k + 1)$-го входа. Изучение этой схемы синтеза может, как нам кажется, представить самостоятельный интерес.

7) *Общие свойства баз.* 1°. Пусть дана база S_{k+1}. Образуем из диаграмм элементов базы, записанных в строки, матрицу $\mathfrak{B}_{k+1}$. Эта матрица невырождена. В силу теоремы Лапласа об определителях в $\mathfrak{B}_{k+1}$ можно в первых $2^k — 1$ столбцах найти отличный от нуля минор $\mathfrak{B}_k$ с ненулевым алгебраическим дополнением $\mathfrak{B}_k^*$, элементы которого берутся из остальных 2^k столбцов. Строки матрицы $\mathfrak{B}_{k+1}$ можно расположить так, что выделенные миноры $\mathfrak{B}_k$ и $\mathfrak{B}_k^*$ будут угловыми, то есть:

$$\mathfrak{B}_{k+1} = \left(\begin{array}{c|c} \mathfrak{B}_k & \mathfrak{B}_k'' \\ \hline \mathfrak{B}_k' & \mathfrak{B}_k^* \end{array}\right).$$

Обозначение минора через $\mathfrak{B}_k$ оправдано тем, что его строки являются диаграммами элементарных k-нейронов и в силу невырожденности $\mathfrak{B}_k$ образуют (как будет показано ниже в лемме 2) некоторую базу S_k.

Пусть в $\mathfrak{B}_{k+1}$ числа симметричных, косых и отмеченных диаграмм, соответственно, равны *a, b* и *c*. Очевидно, что $a + b + c = 2^{k+1} - 1$, так как $(2^{k+1} - 1)$ — размерность матрицы $\mathfrak{B}_{k+1}$. Кроме того, имеют место неравенства: $a \leqslant 2^k - 1,\ b \leqslant 2^k - 1,\ 0 < c \leqslant 2^k$.

Действительно, если $a > 2^k - 1$, то в $\mathfrak{B}_{k+1}$ диаграмм вида $((\mu)_k\ 0\ (\mu)_k)$ больше, чем координат в $(\mu)_k$; система таких строк будет линейно зависима, так как будет линейно зависима система из строк $(\mu)_k$. Аналогично в случае $b > 2^k - 1$. Так как в базе обязательно должен быть хотя бы один отмеченный $(k + 1)$-элемент, а всего их 2^k, то последнее неравенство очевидно.

2°. ***Лемма* 1.** Для всякой базы S_k модуль определителя матрицы B_k, а значит $\mathfrak{B}_k$, равен единице.

Доказательство. База по определению суть такая система (диаграмм элементарных нейронов), что любой вектор $[\mu]_k$ она представляет линейной комбинацией с целыми коэффициентами. Из этого следует, что все элементы матрицы $M_k = B_k^{-1}$ в формуле

$$[\rho]_k = M_k [\mu]_k$$

являются целыми числами. Ведь если бы в некоторой строке элемент m_{ij} был нецелым, то беря все μ_i в $[\mu]_k$ равными нулю, кроме $\mu_j = 1$, мы получили бы $\rho_1 = m_{ij}$ — нецелое число. Но определитель (det) B_k — целое число, так как как B_k — бинарная матрица. Отсюда в силу $\det B_k^{-1} = \det M_k = 1/\det B_k$, получаем $|\det |M_k|| = |\det B_k| = |\det \mathfrak{B}_k| = 1$; матрица $\mathfrak{B}_k$—транспортированная к B_k.

До сих пор всюду считалось, что размерность базы S_k равна $2^k - 1$, хотя строгого обоснования не было дано.

***Лемма* 2.** Всякая максимальная линейно независимая система элементарных диаграмм (т. е. диаграмм элементарных *k*-нейронов) является базой в пространстве диаграмм, а соответствующие элементарные *k* -нейроны образуют базу S_k.

Доказательство. Возьмем произвольную максимальную линейнонезависимую систему $\mathfrak{B}_k$ элементарных диаграмм и покажем, что модуль определителя матрицы $\mathfrak{B}_k$ равен единице, что и приведет к доказательству леммы (так как утверждение, обратное к лемме 1, справедливо). Присоединим к $\mathfrak{B}_k$ произвольный вектор $(1, (\mu)_k)$ некоторого отмеченного элемента и рассмотрим матрицу $\mathfrak{B}_k^*$:

$$\mathfrak{B}_k^* = \begin{pmatrix} [0]_k & \mathfrak{B}_k \\ 1 & (\mu)_k \end{pmatrix}.$$

Очевидно, что $\det \mathfrak{B}_k^* = -\det \mathfrak{B}_k$. $\mathfrak{B}_k^*$ может быть получено из A_k: каждая строка $\mathfrak{B}_k^*$, кроме последней, есть разность двух строк A_k, откуда по известным теоремам об определителях и следует

$$|\det \mathfrak{B}_k^*| = |\det \mathfrak{B}_k| = |\det A_k| = 1.$$

Приведем одну элементарную лемму, не относящуюся непосредственно к нейронам.

***Лемма* 3** (о переходе). Для любых двух максимальных линейно независимых систем $\{\alpha\}$ и $\{\beta\}$ векторов одинаковой размерности можно последовательными заменами векторов из $\{\alpha\}$ на вектора из $\{\beta\}$ прийти к системе $\{\beta\}$ так, что все промежуточные системы будут также максимальными линейно независимыми системами.

Для доказательства достаточно показать, что в $\{\alpha\}$ можно заменить хотя бы один вектор на некоторый вектор из $\{\beta\}$ с сохранением линейной независимости. Возьмем разложение некоторого вектора β по векторам системы $\{\alpha\}$: $\beta = \Sigma\, c_i\alpha_i$. Не все $c_i = 0$ одновременно. Выбирая ненулевой c_i и осуществляя замену α_i на β получаем, очевидно, линейно независимую систему.

3°. ***Теорема.*** Во всякой базе S_k общее число ветвей в ее элементах $R_k \geqslant k \cdot 2^{k-1}$.

При $k = 2$ этот факт легко проверяется. Отметим два частных случая.

А). В базе S_{k+i} содержится полная система отмеченных элементов ($c = 2^k$). Тогда разбиение соответствующей матрицы $\mathfrak{B}_{k+1}$ на миноры единственно и таково, что $\mathfrak{B}_k^* = A_k$, $\mathfrak{B}_k' = 0$; $\mathfrak{B}_k$ — матрица, соответствующая по лемме 2 некоторой базе S_k. В полной системе отмеченных элементов число ветвей [пункт 3)] $v_{k+1} = (k + 2)\, 2^{k-1}$. В базе S_k по предположению число ветвей $R_k \geqslant k2^{k-1}$. Отсюда

$$R_{k+1} \geqslant k \cdot 2^{k-1} + (k + 2)\, 2^{k-1} = (k + 1)\, 2^k.$$

В этом случае теорема доказана. Заметим, что число баз Q_{k+1} содержащих полную систему отмеченных элементов, легко выражается через число баз S_k. Ведь каждая строка матрицы $(\mathfrak{B}_k, \mathfrak{B}_k')$ может быть либо симметричной, либо косой диаграммой, поэтому для данной матрицы $\mathfrak{B}_k$ благодаря вариациям в $\mathfrak{B}_k''$, мы можем получить 2^{2^k-1} матриц $\mathfrak{B}_{k+1}$. Если же всех k-баз S_k существует z_k, то $(k + 1)$-баз с матрицей A_k в качестве $\mathfrak{B}_k^*$ будет

$$Q_{k+1} = z_k 2^{2^k-1}$$

Ограничиваясь на каждом шаге только матрицами этого типа (кроме случая $k = 2$, когда берем четыре базы), мы получаем

$$Q_{k+1} = Q_k 2^{2^k-1},\ Q_2=4.$$

Можно вывести, что

$$Q_k = \frac{2^{2^k}}{2^k}.$$

Фактически всем этим развита еще одна конструктивная обобщенная схема синтеза, дающая довольно большое число общих решений. Например $Q_3 = 32$. Обе обобщенные схемы можно комбинировать, так что общее число частных конструктивных схем еще более увеличивается, выражаясь рекуррентной формулой:

$$\tilde{Q}_{k+1} = \tilde{Q}_k 2^{2^k-1} + \tilde{Q}^2 2^k,\ \tilde{Q}_2 = 4.$$

В). Во всей базе S_{k+1} лишь один отмеченный элемент. Оставшихся элементов $2^{k+1} - 2 = 2(2^k - 1) = a + b,$ откуда, в силу неравенств пункта 1°, $a = b = 2^k - 1$. Тогда $2^k - 1$ симметричные диаграммы образуют базу симметричных диаграмм и требуют по предположению не менее $k \cdot 2^{k-1}$ ветвей, а косые диаграммы, также образующие базу, требуют не менее $k2^{k-1}+2^k-1$ ветвей, так как каждый косой элемент имеет дополнительный запрет от $(k + 1)$-го входа. Единственный отмеченный элемент требует хотя бы одной ветви. Итак:

$$R_{k+1} \geqslant k2^{k-1} + k2^{k-1} + 2^k - 1 + 1 = (k+1)2^k,$$

что и требовалось доказать.

Изложим ход доказательства в общем случае.

В системе строк $(\mathfrak{B}_k' \ \mathfrak{B}_k^*)$ матрицы $\mathfrak{B}_{k+1}$ отмеченных трок имеется c $(1 < c < 2^k)$. По лемме 3 о переходе, от A_k можно перейти к $\mathfrak{B}_k^*$ через линейно независимые системы заменой отмеченных строк, не входящих в $\mathfrak{B}_k^*$ строками из $\mathfrak{B}_k^*$. Важно, что в диаграмме элемента с q фиктивными входами содержится ровно 2^q единиц (остальные координаты — нули). Если в промежуточной невырожденной матрице $\tilde{\mathfrak{B}}_k^*$ заменить некоторую отмеченную строку неотмеченной с большим числом единиц, то линейная независимость нарушается. Действительно, как показано в пункте 6), всякая элементарная диаграмма $(\alpha)_k^0$ есть разность двух строк из A_k: $(\alpha)_k^0 = \gamma' - \gamma''$, причем в γ' единиц вдвое больше, чем в $(\alpha)_k^0$, а в γ'' их столько же, сколько в $(\alpha)_k^0$. С другой стороны, по лемме о переходе, для сохранения линейной независимости замену в $\tilde{\mathfrak{B}}_k^*$ на вектор $(\alpha)_k^0$ можно произвести лишь того вектора, который входит в линейное выражение $(\alpha)_k^0$ через векторы системы $\tilde{\mathfrak{B}}_k^0$ с ненулевым

коэффициентом. Однако по предыдущему в выражении $(\alpha)_k^0$ участвуют лишь векторы с тем же или большим числом единиц. Таким образом, заменить на $(\alpha)_k^0$ вектор с меньшим числом единиц с сохранением линейной независимости нельзя.

Возьмем в базе S_{k+1} все отмеченные элементарные нейроны, входящие своими строками в $\mathfrak{B}_k^{\cdot}$, и элементарные k-нейроны, дающие остальные строки $\mathfrak{B}_k^{\cdot}$. Общее число фиктивных входов в этих нейронах не больше числа фиктивных входов во всех отмеченных $(k+1)$-элементах. Последнее равно $(k+1)\,2^k - v_{k+1} = (k+1)\,2^k - (k+2)2^{k-1} = k2^{k-1}$. По предположению в S_k $R_k \geqslant k2^{k-1}$; тогда число фиктивных входов $\Phi_k \leqslant k\,(2^k - 1) - k2^{k-1} = k(2^{k-1} - 1)$. В S_{k+1} число фиктивных входов $\Phi_{k+1} \leqslant k\,(2^{k-1} - 1) + k2^{k-1} + a$, но $a \leqslant 2^k - 1$, т. е. $\Phi_{k+1} \leqslant (k+1) \times 2^k - 1)$, откуда $R_{k+1} \geqslant (k+1)\,(2^{k+1} - 1) - \Phi_{k+1} = (k+1)\,2^k$, что и требуется.

Замечание. Полученный результат показывает, что не существует алгоритма синтеза формальных нейронов с экономичностью выше, чем $\frac{k+1}{2}$; таким образом, выделенное каноническое решение задачи синтеза обладает наивысшей экономичностью.

4°. Пусть дан некоторый нейрон N_k; образуем нейрон N'_k некоторой перестановкой входов в N_k с сохранением всех связей между волокнами. Нейроны N_k и N_k' назовем *изоморфными.* В табл. 11 приведены два изоморфных нейрона, полученные перестановкой 1-го и 3-го входов.

Таблица 11

Каждому k-нейрону соответствует $k!$ изоморфных нейронов. Пороговые диаграммы изоморфных нейронов отличаются также перестановками своих координат. Пусть перестановка входов задана подстановкой

$$\begin{pmatrix} 1, & 2, \ldots, & k \\ p_1, & p_2, \ldots, & p_k \end{pmatrix}$$

Соответствующая перестановка элементов диаграмм дает подстановку

$$\begin{pmatrix} 0,\ 1,\ 2,\ \ldots,\ 2^k-1 \\ q_0,\ q_1,\ q_2,\ \ldots,\ q_{2^k-1} \end{pmatrix},$$

где q_j- — индексы элементов пороговой диаграммы нейрона N_k. Легко установить, что

$$q_{2^{m-1}} = 2^{p_m-1},$$

откуда следует, что

$$q_{2^{m-1}+l} = q_{2^{m-1}} + q_l \quad (1 \leqslant l < 2^{m-1}).$$

Этим и задается правило соответствия перестановок координат диаграмм перестановкам входов.

При перестановке входов в некоторой системе элементов возможно, что какие-либо элементарные нейроны перейдут в нейроны этой же системы или в себя. Если это имеет место для всех элементов некоторой системы, то будем говорить, что система перешла в себя.

Теорема. При перестановке входов никакая база не переходит в себя.

Иначе: всякая система 2^k — 1 диаграмм элементарных нейронов, образующих систему, переходящую в себя, линейно зависима, Изложим ход доказательства. Его достаточно провести для систем, переходящих в себя при перестановке первых двух входов. Система, переходящая в себя, разбивается на две подсистемы: подсистему пар элементов, переходящих друг в друга, и подсистему элементов, переходящих в себя (одинарные элементы).

Возьмем два элемента, переходящих друг в друга при нашей перестановке, и их пороговые диаграммы. Разобьем эти диаграммы на грани (кварты) по четыре координаты в каждой. Очевидно, что если в некоторой кварте нет единиц у одной диаграммы, то их нет в этой же кварте и у другой диаграммы, причем ненулевые кварты в пределах одной диаграммы одинаковы. Беря разность таких диаграмм, можно установить, что разность кварт имеет вид (01 —10). Обозначая такую кварту через *I*, а нулевую (0, 0, 0, 0) через 0, устанавливаем соответствие между разностями переходящих друг в друга диаграмм и векторами размерности 2^{k-2}. Отсюда следует, что система более чем 2^{k-2} пар переходящих друг в друга диаграмм размерности 2^k линейно зависима.

В случае одинарных диаграмм, т. е. переходящих в себя, можно заметить, что у них кварты бывают только двух типов (1000) и (1111). Отсюда следует, что система более чем 2^{k-1}-2 одинарных диаграмм линейно зависима, так как первая грань у всех одинарных диаграмм нулевая и соответствие устанавливается с векторами размерности 2^{k-2}—1. Таким образом, так как $2^k - 1 > 2 \cdot 2^{k-2} + (2^{k-1} - 2)$,

всякая система $2^k - 1$ диаграмм, переходящая в себя, оказывается линейно зависимой.

Итак, полная система баз k-нейронов разбивается на непересекающиеся классы по k- изоморфных между собой баз. Выбирая из некоторого класса базу и находя ее M_k, можно произвести синтез нейронов во всех базах взятого класса, применяя M_k к диаграммам, полученным из данной перестановками, соответствующими всем перестановкам входов.

5°. **Пример**. Назовем правильными базы, приводящие к более экономному синтезу нейронов. Нахождение правильных баз в общем случае весьма трудная (пока не решенная) задача. Для k= 3 правильных баз (с числом ветвей, равным 12), как нетрудно установить, имеется 24. Эта система замкнута относительно перестановок входов и по предыдущему представляется четырьмя базами. Приведем этих представителей. Для краткости представим их как объединения:

$$\{S_3^i\} = \{[a] \vee [bi]\},$$

где элементы подсистем записаны символически

$$[a] \rightleftharpoons [(00x),\ (0x0),\ (\alpha 00),\ (\beta\beta x)],$$
$$[b1] \rightleftharpoons [(0\beta\alpha),\ (\beta 0x),\ (\beta\alpha 0)],$$
$$[b2] \rightleftharpoons [(0\beta x),\ (\alpha 0\beta),\ (\alpha\beta 0)],$$
$$[b3] \rightleftharpoons [(0\beta\alpha),\ (\beta x0),\ (\alpha 0\beta)],$$
$$[b4] \rightleftharpoons [(0\alpha\beta),\ (\beta\alpha 0),\ (\alpha 0\beta)].$$

В базе S^1_3 легко узнать каноническую базу, ее матрица M^1_3 есть матрица M_3, приведенная в пункте 5. Приведем остальные матрицы M^j_3, считая базы $\{S^j_3\}$ упорядоченными в (α, β)-порядке, j = 2, 3, 4.

$$M_3^2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 & 1 & 1 & -1 \\ 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & -1 & 0 & 0 \\ 1 & 1 & -1 & 1 & -1 & -1 & 1 \\ -1 & 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix},$$

$$M_3^3 = \begin{pmatrix} 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 1 & 0 & -1 & 1 \\ 0 & -1 & 1 & 0 & 0 & 1 & -1 \end{pmatrix},$$

$$M_3^4 = \begin{pmatrix} 0 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 1 & 1 & -1 \\ 0 & 1 & 0 & 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 1 & 0 & -1 & 1 \\ 0 & -1 & 1 & 0 & 0 & 1 & -1 \end{pmatrix}.$$

8) *О синтезе оптимальных формальных нейронов.* Критерии оптимальности нейрона могут быть различными в зависимости от преследуемых целей. Выделим две основные задачи (см. также 1.4.4.2, 1.4.4.4).

1. Синтезировать нейрон с минимальным числом волокон, реализующий данную диаграмму.
2. Синтезировать нейрон с минимальным числом ветвей, реализующий данную диаграмму.

Пусть нейрон реализует данную диаграмму и является минимальным по числу ветвей. Тогда элементы, из которых он построен, образуют линейно независимую систему. В ином случае некоторые элементы можно исключить, представляя их через оставшиеся, что противоречит оптимальности. Таким образом, система элементарных нейронов оптимального по числу ветвей нейрона либо сама является базой, либо может быть дополнена (возможно неединственным образом) до некоторой базы S_k. Пусть S_k соответствует матрица M_k. В силу однозначности представления любой данной диаграммы в элементах любой базы, умножение M_k на диаграмму $[\mu]_k$ дает весовые коэффициенты элементов, составляющих именно оптимальный нейрон. Этим, собственно, дается принципиальное решение задачи II оптимального синтеза. Из-за обширности множества баз k-нейронов применение этой методики требует даже при небольших k использования компьютеров.

Можно также принципиально показать существование диаграмм, которые нельзя реализовать *k*-нейронами с числом ветвей, меньшим, чем $k \cdot 2^{k-1}$.

Возьмем произвольную базу S_k и ее матрицу M_k. Нейрон, реализующий данную диаграмму, имеет число ветвей, меньшее числа ветвей в S_k, если некоторые координаты весового вектора

$$[\rho]_k = \mathfrak{M}_k [\mu]_k$$

равны нулю. Если $\rho_i = 0$, то

$$\Sigma m_{ij}\mu_j = \rho_i = 0,$$

где $(m_{ij}) = M_k$ и $\mu_j \in [\mu]_k$. Считая диаграмму $(\mu)_k$ точкой в $(2^k - 1)$-мерном пространстве, в случае $\rho_i = 0$ получаем, что точка $(\mu)_k$ лежит на гиперплоскости размерности $(2^k - 2)$, перпендикулярной вектору $(m_{i,1}, \ldots, m_{i, 2^k-1})$ и проходящей через начало координат. Очевидно, что множество допускающих оптимальную реализацию диаграмм есть множество целочисленных точек, лежащих на описанных гиперплоскостях, и только оно.

При k=2 легко указать систему плоскостей в трехмерном пространстве векторов (μ_1, μ_2, μ_3), на которых лежат точки диаграммы нейронов, допускающих оптимальную реализацию. Эти плоскости суть:

$$\mu_1 = 0, \quad \mu_2 - \mu_3 = 0,$$
$$\mu_2 = 0, \quad \mu_1 - \mu_2 = 0, \quad \mu_1 + \mu_2 - \mu_3 = 0.$$

В заключение приведем подсчет числа всех баз S_3. Запишем систему строк $\mathfrak{A}_3$ (4.5.6) без первого нулевого столбца в три группы: симметричные, косые и отмеченныедиаграммы.

(11) = (1010101),	(21) = (1010000),	(31) = (0001111)
(12) = (1000100),	(22) = (1000000),	(32) = (0001010)
(13) = (0110011),	(23) = (0110000),	(33) = (0001100)
(14) = (0100010),	(24) = (0100000),	(34) = (0001000).

Каждой диаграмме присвоим индекс *(i, j)*: *i* — номер группы, *j* — номер диаграммы. Произведем подсчет минимальных линейно зависимых систем диаграмм; диаграммы в системах указаны своими индексами.

{1°} = {11, 12, 13, 14}	{1} = {12, 13, 14, 21, 31, 32}
{2°} = {21, 22, 23, 24}	{2} = {11, 13, 14, 22, 33, 34}
{1′} = {11, 21, 31, 32}	{3} = {11, 12, 14, 23, 31, 33}
{2′} = {12, 22, 33, 34}	{4} = {11, 12, 13, 24, 32, 34}
{3′} = {13, 23, 31, 33}	{5} = {22, 23, 24, 11, 31, 32}
{4′} = {14, 24, 32, 34}	{6} = {21, 23, 24, 12, 33, 34}

$$\{7\} = \{21, 22, 24, 13, 31, 33\}$$
$$\{8\} = \{21, 22, 23, 14, 32, 34\}$$
$$\{9\} = \{11, 21, 32, 13, 23, 33\}$$
$$\{10\} = \{11, 21, 31, 14, 24, 34\}$$
$$\{11\} = \{12, 22, 34, 13, 23, 31\}$$
$$\{12\} = \{12, 22, 33, 14, 24, 32\}$$

Заметим, что объединения систем $\{p^{\circ}\} \cup \{q'\}$ и $\{1'\} \cup \cup \{3'\}$, $\{1'\} \cup \{4'\}$, $\{2'\} \cup \{3'\}$, $\{2'\} \cup \{4'\}$ содержат системы {1} — {12}. Это дает возможность учесть повторяющиеся линейно зависимые системы. Всех линейно зависимых систем по 7 диаграмм оказывается L=384, откуда линейно независимых $C_{12}^{7} - L = 408$. К тому же результату можно прийти систематическим построением баз S_3, что, однако, гораздо более громоздко.

В заключение отметим, что проблемы вывода логических следствий представляют собой основное содержание курсов математической логики и теории доказательства. Хотя, как известно, в наиболее интересных случаях задача разрешения неразрешима, тем не менее внимание логиков все больше и больше привлекают такие фрагменты логических исчислений, где проблему разрешения удается решить. Как это сформулировал Н. А. Шанин, существуют два разных способа решения этой проблемы: один, который он называет «алгоритм-оракул», и другой, который носит у него название «интеллектуальный партнер.

«Алгоритм-оракул» перерабатывает всякую формулу данного исчисления или его фрагмента в ответы «да» или «нет», т. е. дает ответ на вопрос, доказуема формула или нет. Алгоритм «интеллектуальный партнер» выдает по всякой доказуемой формуле ее вывод (вывод при этом является некоторым словом в определенном алфавите; речь идет об алгоритме поиска «естественного» вывода, т. е. вывода, достаточно похожего на обычный содержательный вывод, осуществляемый человеком). В применении к данной системе посылок в общем случае, однако, отнюдь не ясно, что именно является логическим следствием этих посылок, какую формулу нужно проверять с точки зрения того, является ли она следствием данных посылок или нет, вывод какой формулы нужно искать. Работы логиков XIX века, к числу которых принадлежит и Венн, и теперь представляют еще интерес потому, что эти логики пытались ответить прежде всего на вопрос о том, какую именно информацию мы можем извлечь — и можем ли — из данных

посылок. Если нам рассказано то-то и то-то о таких-то классах, то что именно следует отсюда в применении только к таким-то классам, какие заключения, и притом именно такого-то вида, мы можем сделать о таких-то классах. Именно на эти вопросы и должны были давать ответ алгебраические и геометрические методы этих логиков, именно они и составляют основную проблематику метода исключения неизвестных. Последний поэтому представляет логический интерес и в наши дни. А задачи, приводившиеся в этой связи логиками XIX столетия, заслуживают поэтому и теперь построения — там, где оно возможно,— алгоритмических методов их решения. Особый интерес в этой связи представляет то обстоятельство, что увеличение количества информации, связанной с сигналами, определяющими ячейку в диаграмме Венна (например, добавление информации, относящейся к порогам нейрона), позволяет расширить самую проблематику вывода логических следствий,— позволяет включить в нее, например, вопросы о том, как строить надежные нейронные схемы из не вполне надежных элементов; как обеспечить правильный ответ автомата на некоторый сигнал, даже в случае не вполне исправной работы его элементов. Эти и многие другие аналогичные вопросы, практическое значение которых достаточно ясно, представляют собой наилучшее обоснование того, сколь существенную роль может иметь и в наши дни дальнейшее развитие, уточнение и совершенствование того круга идей, над которыми бились уже Буль, Джевонс, Шредер, Порецкий, Венн, изучение работ которых и до сих пор является поэтому источником новых идей в математической логике и ее все более и более многочисленных и важных применениях.

5. Теорема Геделя о неполноте

5.1. Постановка задачи

Формулировка теоремы о неполноте, которую мы будем уточнять и доказывать, такова: при определенных условиях

в языке существует недоказуемое истинное утверждение.

В этой формулировке едва ли не каждое слово нуждается в разъяснениях. Сделаем такие разъяснения.

1. **Язык**. Мы не будем давать какое бы то ни было определение языка (поскольку не беремся это сделать с достаточной общностью), а ограничимся теми относящимися к языку понятиями, которые единственно и будут нужны нам для дальнейшего. Таких понятий нам

потребуется два: «алфавит языка» и «множество истинных утверждений языка».

1.1. Алфавит. Под *алфавитом* понимается конечный список элементарных (т. е. считающихся не членимыми далее) знаков, называемых *буквами* этого алфавита. Конечная цепочка следующих друг за другом букв некоторого алфавита называется *словом* в этом алфавите. Так, слова русского языка (включая и собственные имена) суть слова в 66-буквенном алфавите (33 строчные буквы, 31 прописная буква (кроме твердого и мягкого знаков), дефис, апостроф); десятичные записи натуральных чисел— слова в десятибуквенном алфавите {0, 1, 2, 3, 4, 5, 6, 7, 8, 9}. Для называния алфавитов используются обычно прописные русские буквы. Множество всех слов в алфавите Б будем обозначать $Б^{\infty}$. Предполагается, что для каждого языка имеется такой алфавит, что все выражения этого языка (т. е. имена тех или иных предметов, утверждения об этих предметах и т. п.) суть слова в этом алфавите; каждую русскую фразу, например, и даже каждый русский текст можно рассматривать как слово в алфавите, представляющем собой расширение указанного выше 66-буквенного алфавита за счет знаков препинания, знака пробела между словами, знака абзацного отступа (и, быть может, еще некоторых знаков). Предполагая, что выражения языка являются словами в некотором алфавите, мы тем самым налагаем запрет на такое «многоэтажное» выражение, как, например, $\int_a^b f(x)\,dx$. Этот запрет, однако, не является слишком ограничительным, поскольку все подобные выражения можно при подходящей кодировке «вытянуть в строку». Всякое множество *M* такое, что $M \subseteq Б^{\infty}$, называется *словарным в* Б. Просто *словарным* называется множество, словарное в каком-либо алфавите. Сделанное только что предположение может быть теперь сформулировано короче: множество выражений всякого языка словарно.

1.2. Множество истинных утверждений. Предполагается, что в множестве $Б^{\infty}$, где Б —алфавит рассматриваемого языка, задано подмножество *T,* называемое множеством «истинных утверждений» (или, короче, просто «истин»). Таким образом, мы опускаем все промежуточные этапы, посредством которых, во-первых, среди слов в алфавите Б выделяются правильно построенные *выражения* языка, получающие определенный смысл при интерпретации (такие, как 2 + 3, *x* + 3, *x* = *y*, *x* = 3, 2 = 3, 2 = 2, — в отличие от таких, как *+=x);* во-вторых, среди выражений выделяются так называемые *формулы,* означающие при интерпретации «утверждения, зависящие, быть

может, от параметра» (такие, как $x = 3$, $x = y$, $2 = 3$, $2 = 2$); в-третьих, среди формул выделяются так называемые *замкнутые формулы,* или утверждения, не зависящие от параметра (такие, как $2 = 3$, $2 = 2$); и лишь, в-четвертых, среди утверждений выделяются *истинные утверждения* (такие, как $2 = 2$).

1.3. Для наших целей будет достаточным считать язык полностью заданным, если задан алфавит Б и подмножество T множества $Б^{\infty}$. Всякую такую пару $\langle Б, T \rangle$ мы будем называть *фундаментальной парой*

2. Недоказуемое. «Недоказуемое» значит не являющееся доказуемым, а «доказуемое» значит имеющее доказательство.

3. Доказательство. Хотя термин «доказательство» является едва ли не самым главным в математике (Н. Бурбаки начинает свои «Начала математики» словами «Со времен греков говорить «математика» — значит говорить «доказательство»), он не имеет точного определения. Понятие доказательства во всей его полноте принадлежит математике не более чем психологии: ведь доказательство — это просто рассуждение, убеждающее нас настолько, что с его помощью мы готовы убеждать других.

3.1. Будучи записанным, доказательство становится словом в некотором алфавите Д (вспомним, что говорилось выше о русских текстах); все доказательства образуют некую (достаточно, впрочем расплывчатую) совокупность в $Д^{\infty}$. Не претендуя на то, чтобы дать точное определение для такого «наивного» или «абсолютного» понятия доказательства (или, что то же самое, для соответствующей совокупности в $Д^{\infty}$), мы займемся его формальным аналогом, для которого сохраним тот же термин «доказательство». Этот аналог в двух существенных чертах будет отличаться от интуитивного понятия: во-первых, мы будем допускать существование разных понятий «доказательства» (что приведет к различным подмножествам в множестве $Д^{\infty}$, да и сам алфавит Д может варьироваться); во-вторых, для каждого из таких понятий мы будем требовать наличия эффективного способа, или алгоритма, проверки, является ли данное слово в алфавите Д доказательством или нет. Далее, будем предполагать наличие алгоритма, который по доказательству определяет, доказательством какого утверждения оно является. (В обычных случаях этим утверждением является последнее утверждение в цепочке, образующей доказательство.)

3.2. Итак, окончательное определение таково:

1° Имеются алфавиты Б *(алфавит языка)* и Д *(алфавит доказательств).*

2° В множестве $Д^{\infty}$ выделено подмножество *D,* элементы которого называются *доказательствами;* предполагается наличие алгоритма, позволяющего но произвольному слову в алфавите Д узнавать, принадлежит оно *D* или пет.

3° Имеется функции δ *(функция выделения доказанного),* у которой область определения Δ удовлетворяет соотношению $D \subseteq \Delta \subseteq Д^{\infty}$ и которая принимает свои значении в $Б^{\infty}$*;* предполагается наличие алгоритма вычисляющего эту функишо; доказательство *d* из *D* называется доказательством слова δ*(d).*

3.3. Тройку $\langle Д, D, \delta \rangle$, удовлетворяющую условиям 1°—3°, назовем *дедуктикой* над алфавитом Б.

3.4. Для читателей, знакомых с обычными способами задания понятия «доказательство» посредством «аксиом» и «правил вывода», поясним, почему эти способы могут быть рассмотрены как частный случай определения из п. 3.2. В самом деле, доказательством обычно называют цепочку выражений языка, в которой каждый член или является аксиомой, или получается из предыдущих по одному из правил вывода. Добавляя к алфавиту языка новую букву *, мы можем записывать доказательства в виде слов в расширенном таким образом алфавите: цепочка $\langle C_1, \ldots, C_n \rangle$ изображается словом $C_1 * C_2 * \ldots * C_n$. Функция выделения доказанного выделяет из каждого слова наибольший не содержащий буквы * конец. Требуемые определением из п. 3.2 алгоритмы могут быть легко построены для любого обычно рассматриваемого уточнения понятий «быть аксиомой» и «получаться по одному из правил вывода».

4. Уточнения первоначальной формулировки.

4.1. Первое уточнение. При определенных условиях для фундаментальной пары $\langle Б, T \rangle$ и дедуктики $\langle Д, D, \delta \rangle$ над Б существует слово из *T,* не имеющее доказательства. Такая формулировка еще слишком неопределенна. К тому же ясно, что можно придумать много дедуктик, в каждой из которых будет очень мало доказуемых слов. В пустой, дедуктике (где $D = \varnothing$) вообще нет ни одного доказуемого слова.

4.2. Второе уточнение. Более естественным является другой подход. Задан некоторый язык в том точном смысле, что задана фундаментальная пара $\langle Б, T \rangle$. Мы теперь ищем дедуктики над Б (содержательно— ищем такие способы доказывания), в которых доказывалось бы как можно больше слов из *T*, в идеале — все слова из *T.* Нас интересует ситуация, когда такой дедуктики (в которой каждое слово из *T* имело бы доказательство) не существует. Итак, нас заинтересовала бы следующая формулировка: *при определенных*

условиях, налагаемых на фундаментальную пару $\langle Б, T\rangle$, *не существует дедуктики над Б, в которой каждое слово из Т имеет доказательство.* Однако пары $\langle Б, T\rangle$ с этим свойством просто не может быть. В. самом деле, достаточно положить $Д = Б$, $D = Д^{\infty}$, $\delta(d) = d$ для всякого d из $Д^{\infty}$; тогда всякое слово из $Б^{\infty}$ окажется доказуемым (его доказательством будет оно само).

5. Непротиворечивость. Естественно потребовать, чтобы доказуемыми были лишь «истинные утверждения», т. е. слова, принадлежащие множеству *Т.* Назовем дедуктику $\langle Д, D, \delta\rangle$ *непротиворечивой относительно* (или *для)* фундаментальной пары $\langle Б, T\rangle$, коль скоро $\delta(D) \subseteq T$. В дальнейшем будем интересоваться лишь непротиворечивыми дедуктиками. Если имеется язык, то представляется необходимым найти такую непротиворечивую дедуктику, в которой каждое истинное утверждение было бы доказуемым. **Теорема Гёделя в интересующем нас варианте именно и утверждает, что при определенных условиях, налагаемых на фундаментальную пару, этого сделать нельзя**.

6. Полнота. Назовем дедуктику $\langle Д, D, \delta\rangle$ *полной относительно* (или *для)* фундаментальной пары $\langle Б, T\rangle$, коль скоро $\delta(D) \supseteq T$. Занимающая нас формулировка приобретает такой вид:

при определенных условиях, налагаемых на фундаментальную пару $\langle Б, T\rangle$, *не существует дедуктики над* Б, *полной и непротиворечивой относительно* $\langle Б, T\rangle$.

На этой формулировке мы и остановимся и в следующих параграфах найдем те условия, о которых в ней идет речь.

5.2. Начальные понятия теории алгоритмов и их применения

Условия, при которых не существует полной непротиворечивой дедуктики, легко формулируются в терминах теории алгоритмов.

Нам вначале сдостаточно лишь самых общих интуитивных представлений об *алгоритме* как о предписании, позволяющем по каждому *исходному данному*, или *аргументу, или* некоторой совокупности *возможных* (для данного алгоритма) *исходных данных* (аргументов) получить *результат* в случае, если таковой существует, или не получить ничего в случае, если для рассматриваемого исходного данного не существует результата (подчеркнем, что возможные исходные данные — это не те данные, в применении к которым алгоритм дает результат, а те, к которым можно его применять (возможно, безрезультатно).) Если для выбранного исход-

ного данного результат существует, говорят, что алгоритм *применим* к этому исходному данному и *перерабатывает* его в этот результат.

Для наших целей будет достаточным — и это позволит избежать лишних обсуждений — считать, что исходные данные и результаты любого алгоритма суть слова. Более точно: для каждого алгоритма можно указать некоторый *алфавит исходных данных,* так что все возможные исходные данные являются словами в этом алфавите, и некоторый *алфавит результатов,* так что все результаты являются словами в этом алфавите. Поэтому, чтобы иметь дело с алгоритмами, применяемыми, скажем, к парам слов или к цепочкам слов, мы должны предварительно записать эти образования в виде слов в каком-нибудь алфавите. Для определенности условимся соотносить с каждым алфавитом Б некоторую не входящую в него букву и обозначать эту букву звездочкой (подчеркнем, что, таким образом, эта звездочка в различных ситуациях обозначает различные буквы). Первоначальный алфавит Б, пополненный этой повой буквой, будем обозначать $Б_*$. *В* п. 3.4 предыдущего параграфа мы уже договорились записывать цепочку $\langle C_1, \ldots, C_n\rangle$ слов в алфавите Б посредством слова $C_1 * \ldots * C_n$ в алфавите $Б_*$; в частности, в том же $Б_*$ запишется в виде слова $C_1 * C_2$ и пара $\langle C_1, C_2\rangle$. Пусть, далее, при фиксированном n $Б_1, Б_2, \ldots, Б_n$ суть произвольные алфавиты; обозначая по-прежнему через $*$ дополнительную букву, соотнесенную с алфавитом $(Б_1 \cup \ldots \cup Б_n)$ и тем самым заведомо не входящую ни в один из $Б_i$, мы будем отождествлять цепочку $\langle C_1, \ldots, C_n\rangle$, где каждое C_i является словом в $Б_i$, со словом $C_1 * \ldots * C_n$ в алфавите $(Б_1 \cup \ldots \cup Б_n)_*$; совокупность всех таких цепочек (и отождествленных с ними слов) будем обозначать через $Б_1^\infty \times \ldots \times Б_n^\infty$.

Совокупность всех исходных данных, к которым алгоритм применим, называется *областью применимости* алгоритма; каждый алгоритм задает функцию, относящую каждому элементу области применимости соответствующий результат; область определения этой функции совпадает, таким образом, с областью применимости алгоритма; говорят, что рассматриваемый алгоритм *вычисляет* функцию, задаваемую указанным способом. Условимся обозначать через $A(x)$ результат применения алгоритма A к объекту x (при этом $A(\langle x_1, x_2, \ldots, x_n\rangle)$ для краткости будем записывать просто как $A(x_1, \ldots, x_n)$). Тогда определение термина «вычисляет» можно переформулировать следующим образом: алгоритм A вычисляет функцию f, коль скоро $A(x) \simeq f(x)$ для всех x. (Знак $\simeq$ есть знак

«условного равенства»; утверждение $A \simeq B$ считается истинным в двух случаях: либо когда выражения *A* и *B оба* не определены, либо когда *A* и *B* оба определены и обозначают одно и то же.)

Функция, которая вычисляется некоторым алгоритмом, называется *вычислимой.* В части 3° определения понятия доказательства (п. 3.2 предыдущего параграфа) говорится, следовательно, о том, что функция выделения доказанного должна быть вычислимой функцией.

В силу сделанных предположений относительно понятия алгоритма для каждой вычислимой функции можно указать такие два алфавита, что все ее аргументы суть слова в первом из этих алфавитов, а все ее значения — слова во втором из этих алфавитов.

Особый интерес представляют функции, аргументы и значения которых суть натуральные числа (число 0 мы также считаем натуральным). Такие функции условимся называть *числовыми.* Чтобы иметь право говорить о вычислимых числовых функциях, мы должны ввести в рассмотрение алгоритмы, имеющие дело с числами, а для этого прежде всего необходимо представить числа в виде слов в каком-либо алфавите, называемом в этом случае *цифровым.* Возможны различные способы такого представления, например:

1) двоичная запись чисел в алфавите {0, 1};

2) десятичная запись чисел в алфавите {0, 1, 2, 3, 4, 5, G, 7, 8, 9};

3) запись чисел и однобуквенном алфавите $\{|\}$, причем число *n* записывается словом $\underbrace{\|\ldots|}_{n\text{ раз}}$;

4) запись чисел в трехбуквенном алфавите $\{|, (,)\}$, причем число *n* записывается словом $\underbrace{(\|\ldots|)}_{n\text{ раз}}$, и т. д.

Для тех или иных целей выбираются наиболее удобные способы записи. Каждая запись числа (в какой-либо фиксированной системе) называется *цифрой.* Допуская вольность речи, говорят об алгоритмах и вычислимых функциях, оперирующих с числами, имея в виду алгоритмы и вычислимые функции, оперирующие с изображающими эти числа цифрами (в какой-либо выбранной системе записи).

Понятие вычислимой числовой функции, таким образом, выглядит зависящим от принятого способа записи чисел. Однако легко обнаружить, что всякая числовая функция, вычислимая при одной системе записи, будет вычислима и при другой, по крайней мере для широкого класса таких систем. Назовем две системы эквивалентными, если существует алгоритм, дающий по записи произвольного числа в первой системе запись этого же числа во второй системе, а также алгоритм, дающий по записи произвольного числа во второй системе

запись этого же числа в первой системе. Приведенные выше примеры систем записи очевидным образом эквивалентны. Покажем, что числовая функция *f*, вычислимая при одной из двух эквивалентных систем записи, вычислима и при другой системе. Пусть *C* и *D* — алгоритмы перехода от первой системы ко второй и обратно, и пусть алгоритм ***A*** вычисляет функцию *f* при первой системе записи (т. е. ***A*** вычисляет функцию на цифрах первой системы, индуцированную функцией *f*). Тогда следующий алгоритм ***B*** будет вычислять *f* при второй системе записи (т. е. вычислять индуцированную функцию на цифрах второй системы):

$$\boldsymbol{B}(x) \simeq \boldsymbol{CAD}(x).$$

Предписание, задающее алгоритм ***B***, может быть словесно выражено следующим образом: «переведи *x* в первую систему счисления, затем примени алгоритм **A**, полученный результат (если таковой получится) переведи обратно во вторую систему счисления». Аналогичным образом вводимое ниже понятие перечислимого числового множества не зависит от выбора системы записи чисел.

Ввиду сказанного мы будем, коль скоро фиксирована какая-либо система записи чисел, не слишком педантично различать числа и цифры; множество и тех и других будем называть натуральным рядом и обозначать буквой $\mathbb{N}$.

Множество называется *перечислимым,* если оно либо пусто, либо является множеством элементов какой-нибудь вычислимой последовательности (т. е. множеством значений какой-нибудь вычислимой функции, определенной на натуральном ряду); о такой функции (последовательности) говорят, что она *перечисляет* рассматриваемое множество. Очевидно, каждое перечислимое множество словарно.

Пример 1. Множество $\mathbb{N}^2$ всевозможных пар натуральных чисел перечислимо: одна из перечисляющих функций — функция

$$\varphi(n) = \langle a, b \rangle, \quad \text{где} \quad n = 2^a(2b+1) - 1.$$

Пример 2. Множество $Ж^\infty$ всех слов в произвольном алфавите Ж перечислимо. Один из возможных способов построения перечисляющей последовательности таков: упорядочиваем произвольным образом элементы Ж; затем слова в Ж упорядочиваем следующим образом: из слов разной длины предшествующим считается то, которое короче, а на словах одинаковой длины вводим словарный порядок (сравнивая два слова, находим первые слева различающиеся буквы и смотрим, какая из них идет раньше в упорядочении алфавита Ж). Выписывая слова в порядке следования друг за другом, получаем требуемую перечисляющую последовательность.

(Может возникнуть вопрос, почему она вычислима, т. е. почему существует алгоритм, дающий по k член a_k этой последовательности с номером k; искомый алгоритм, например, таков: выписывай члены последовательности, пока их не станет $k+1$; последний из выписанных членов и будет a_k (напомним, что последовательность начинается с a_0).)

Пример 3. Вычислимая функция f, перечисляющая $Ж^\infty$ и построенная в примере 2, осуществляет взаимно однозначное отображение $\mathbb{N}$ на $Ж^\infty$; поэтому можно говорить об обратной функции f^{-1} осуществляющей взаимно однозначное отображение $Ж^\infty$ на $\mathbb{N}$. Эта f^{-1} тоже вычислима, поскольку вычисляется следующим алгоритмом: чтобы вычислить $f^{-1}(a)$, вычисляй последовательно $f(0), f(1), f(2), \ldots$ и т. д., пока для некоторого n не получишь $f(n)=a$; это n и есть $f^{-1}(a)$.

Пример 4. Для любых алфавитов $Ж_1$ и $Ж_2$ композиция вычислимых функций, взаимно однозначно отображающих $\mathbb{N}$ на $Ж^\infty_2$ (пример 2) и $Ж^\infty_1$ на $\mathbb{N}$ (пример 3), дает вычислимую же функцию, взаимно однозначно отображающую $Ж^\infty_1$ на $Ж^\infty_2$.

Подмножество S множества A называется *разрешимым* относительно A, коль скоро существует такой алгоритм *(разрешающий S относительно A),* который распознает принадлежность элементов A к S, т. е. такой алгоритм, который все элементы из S перерабатывает в некоторое одно и то же слово x (например, в слово «да»), а все элементы из $A \setminus S$ в некоторое одно и то же, но отличное от x слово y (например, в слово «нет»; разумеется, выбор слов x и y совершенно несуществен). Очевидно, разрешимость множества S относительно A равносильна разрешимости множества $A \setminus S$ относительно того же A. В части 2° определения понятия доказательства требовалось, чтобы множество всех доказательств было разрешимо относительно множества всех слов в алфавите доказательств.

Из определения разрешимости вытекает, что область применимости алгоритма, разрешающего S относительно A, объемлет A. При этом безразлично, что получается в результате применения алгоритма к словам, не лежащим в A. Например, если мы хотим построить алгоритм, отличающий стихи Пушкина от стихов Лермонтова, и тем самым доказать разрешимость множества стихов Пушкина относительно множества стихов Пушкина и Лермонтова, то нам неважно, что получится (и получится ли что-нибудь вообще) в результате применения этого алгоритма к стихотворению Маяковского или к Уставу гарнизонной и караульной служб. Возникает естественный вопрос: что произойдет, если предложить другое, более узкое определение разрешимости, потребовав, чтобы разрешающий алгоритм был применим *только* к элементам множества A? При таком определении разрешимость S

относительно *А* равносильна, очевидно, вычислимости характеристической функции множества S относительно *А* (т. е. определенной на *А* функции, равной 1 на S и 0 на $A \setminus S$). Как будет показано в п..5.5 (следствие 1 аксиомы протокола), область применимости любого алгоритма всегда есть перечислимое множество, и потому лишь перечислимые множества могут обладать разрешимыми — в смысле нового, узкого определения— подмножествами. Если же множество *А* перечислимо, то для него оба определения разрешимого подмножества приводят к одинаковым результатам. В самом деле, пусть вычислимая функция *f* перечисляет *А*, а алгоритм ***В*** разрешает *S* относительно *А* в прежнем, широком смысле. Тогда следующий алгоритм будет также разрешать S относительно *А* и притом иметь *А* своею областью применимости: бери произвольное *а* и вычисляй последовательно *f* (0), *f* (1), *f* (2), ... ; как только получишь *f(n)=a,* применяй к *а* алгоритм ***В***.

Замечание 1. Поскольку каждая вычислимая функция, каждое перечислимое множество и каждое разрешимое подмножество задаются некоторым алгоритмом, существование функций, множеств и подмножеств, не являющихся соответственно вычислимыми, перечислимыми или разрешимыми (имеются в виду утверждения о существовании неперечислрмых словарных множеств, невычислимых функций со словарными областями определения и значений и т. п.), усматриваются из количественных соображений. Действительно, каждый алгоритм может быть записан в конечном счете на русском языке (с добавлением, если надо, необходимых математических символов), т. е., согласно п. 1.1 предыдущего параграфа, в виде слова в некотором достаточно обширном алфавите, а всех слов в произвольном алфавите — счетное множество. Конечно, от такого рассуждения еще далеко до построения индивидуальных примеров неалгоритмических объектов.

Приложим теперь описанные только что понятия теории алгоритмов к исследованию возможности существования полной непротиворечивой дедуктики.

Лемма 1. *Каково бы ни было словарное множество X, множества Ø и X разрешимы относительно X.*

Доказательство. Пусть *X* словарно в Ж. Достаточно взять алгоритм, который каждому слову из $Ж^{\infty}$ ставит в соответствие некоторое одно и то же слово *х.* Этот алгоритм будет разрешать каждое из множеств Ø и *X* относительно *X.*

Теорема 1. *Если Т — перечислимое множество, то для фундаментальной пары* $\langle Б, T \rangle$ *можно ввести полную непротиворечивую дедуктику.*

Доказательство. Требуется задать тройку $\langle Д, D, \delta \rangle$. Замечаем, что Ø и $Д^{\infty}$ разрешимы (относительно $Д^{\infty}$) по лемме 1. Если $T = Ø$, то берем $\langle Д, Ø, \delta \rangle$, где Д и δ — любые. Если $T \neq Ø$, то $T = \{\tau(0), \tau(1), \tau(2), \ldots\}$, где x — вычислимая функция; отождествим число n со словом $\|\ldots\|$ длины n и положим $Д = \{\ |\ \}$, $D = Д^{\infty}$, $\delta = \tau$.

Замечание 2. Это доказательство не такое искусственное, как может показаться на первый взгляд. В самом деле, если множество истин некоторого языка перечислимо, т. е. может быть расположено в вычислимую последовательность, то для того, чтобы убедиться в принадлежности какого-либо выражения к этому множеству (т. е. доказать истинность рассматриваемого выражения), достаточно указать номер этого выражения в этой последовательности (каковой номер поэтому и можно считать доказательством). Обратное к теореме 1 утверждение будет доказано дальше (теорема 3); предварительно нам придется доказать некоторые вспомогательный утверждения.

Лемма 2 (о перечислимости разрешимого подмножества). *Разрешимое подмножество перечислимого множества перечислимо.*

Доказательство. Пусть $S \subseteq A$, причем A перечисляется вычислимой функцией f. Если S пусто, то S перечислимо по определению. Если S непусто, то существует такое s, что $s \in S$. Положим

$$g(n) = \begin{cases} f(n), & \text{если } f(n) \in S, \\ s, & \text{если } f(n) \in A \setminus S. \end{cases}$$

Очевидно, g есть вычислимая функция, перечисляющая множество S.

Из леммы 2 вытекает, что всякое разрешимое подмножество натурального ряда перечислимо. Однако обратное утверждение неверно: в п.1.5.5 будет построен пример перечислимого неразрешимого подмножества натурального ряда. Следующая лемма указывает условия разрешимости перечислимого множества.

Лемма 3. *Подмножество S перечислимого множества X тогда и только тогда разрешимо относительно X, когда перечислимо как S, так и его дополнение $X \setminus S$.*

Доказательство. Если S разрешимо, то разрешимо и $X \setminus S$, и остается применить лемму 2 о перечислимости разрешимого множества. Пусть теперь S и $X \setminus S$ оба перечислимы. Если хотя бы одно из них пусто, то по лемме 1 множество S разрешимо. Если оба они непусты, то, значит, перечисляются некоторыми вычислимыми функциями f и g. Тогда, чтобы ответить на вопрос «$x \in S$?», поставленный для произвольного x из X, достаточно вычислять последовательно

$$f(0),\quad g(0),\quad f(1),\quad g(1),\quad \ldots$$

до тех пор, пока не встретится x (что произойдет непременно, так как образующаяся последовательность исчерпывает все X). Если при этом окажется, что x встретилось среди значений f, то x принадлежит S; если же x встретилось среди значений g, то x не принадлежит S.

Теорема 2. *Множество всех доказательств (для данной дедуктики) перечислимо.*

Доказательство. Множество всех слов в алфавите доказательств перечислимо (см. пример 2). Поэтому достаточно применить лемму 2.

Лемма 4 (об образе перечислимого множества). *Пусть R перечислимо и f — вычислимая функция, определенная на всех элементах множества R. Тогда f(R) перечислимо.*

Доказательство. Если R пусто, то и $f(R)$ пусто. Если R перечисляется вычислимой функцией ρ, то $f(R)$ перечисляется вычислимой функцией $y = f(\rho(x))$.

Пример 5. Пусть $\neg$— символ алфавита $Л$, $A \subseteq Л^\infty$. Обозначим через $\neg A$ множество всех слов вида $\neg a$, где $a \in A$. Полагая в лемме 4 $R = A$, $f(a) = \neg a$, получаем, что из перечислимости A вытекает перечислимость $\neg A$; полагая $R = \neg A$, $f(\neg a) = a$, получаем, что из перечислимости $\neg A$ вытекает перечислимость A.

Пример 6. Для любого алфавита И множество $И^\infty \times И^\infty$ перечислимо. В самом деле, множества $\mathbb{N}^2$ и $И^\infty$ перечислимы (примеры 1 и 2). Пусть $И^\infty$ перечисляется вычислимой последовательностью g. Определим на $\mathbb{N}^2$ вычислимую функцию f, полагая $f(a, b) = \langle g(a), g(b)\rangle$. Очевидно, $f(\mathbb{N}^2) = И^\infty \times И^\infty$ и остается применить лемму 4.

Как обычно, через $K_1 \times K_2 \times \ldots \times K_n$ обозначается прямое произведение множеств $K_1, \ldots, K_n$, т. е. множество всех таких цепочек $\langle k_1, \ldots, k_n\rangle$, что $k_1 \in K_1, \ldots, k_n \in K_n$. В силу соглашений, сделанных в начале параграфа, в случае, если $K_1 \subseteq Б_1^\infty, \ldots, K_n \subseteq Б_n^\infty$, где $Б_1, \ldots, Б_n$ — алфавиты, прямое произведение $K_1 \times \ldots \times K_n$ отождествляется с некоторым множеством слов из $Б_1^\infty \times \ldots \times Б_n^\infty$.

Следствие 1 леммы 4. *Если $K_1, \ldots, K_n$ суть перечислимые множества, то их прямое произведение $K_1 \times K_2 \times \ldots \times K_n$ также перечислимо.*

Доказательство. Для $n=2$ — как в примере 6. Далее — по индукции, применяя лемму 4 к «естественному» вычислимому отображению множества $(K_1 \times \ldots \times K_s) \times K_{s+1}$ на множество $K_1 \times \ldots \times K_s \times K_{s+1}$.

Цепочка $\langle C_{i_1}, \ldots, C_{i_r}\rangle$, где $i_1 \leqslant n, \ldots, i_r \leqslant n$, называется *проекцией* цепочки $\langle C_1, \ldots, C_n\rangle$ на оси $i_1, \ldots, i_r$ и обозначается $\mathrm{пр}_{i_1, \ldots, i_r}\langle C_1, \ldots, C_n\rangle$.

В частности, $\mathrm{пр}_1\langle C_1, \ldots, C_n\rangle = C_1$, $\mathrm{пр}_2\langle C_1, \ldots, C_n\rangle = C_2$.

Если $M \subseteq K_1 \times \ldots \times K_n$, то через $\mathrm{пр}_{i_1, \ldots, i_r} M$ обозначается множество всевозможных проекций $\mathrm{пр}_{i_1, \ldots, i_r} m$, где $m \in M$.

Следствие 2 леммы 4. *Если M — перечислимое подмножество множества* $\text{Б}_1^\infty \times \ldots \times \text{Б}_n^\infty$, *где* $\text{Б}_1, \ldots, \text{Б}_n$ — *некоторые алфавиты, а* $i_1, \ldots, i_r$ — *числа, не превосходящие n, то* $\mathrm{пр}_{i_1, \ldots, i_r} M$ *перечислимо.*

Доказательство. Достаточно рассмотреть вычислимую функцию $x \mapsto \mathrm{пр}_{i_1, \ldots, i_r} x$.

Т е о р е м а 3. *Множество всех доказуемых слов (для данной дедуктики) перечислимо.*

Доказательство. Пусть P — множество всех доказуемых слов для дедуктики $\langle \text{Д}, D, \delta\rangle$. Очевидно, что $P = \delta(D)$. По теореме 2 множество D перечислимо. Остается применить лемму 4.

Таким образом, если T неперечислимо, то для пары $\langle \text{Б}, T\rangle$ невозможно ввести полную непротиворечивую дедуктику; для всякой непротиворечивой дедуктики множество доказуемых слов P будет собственным подмножеством множества T и в разности $T \setminus P$ всегда найдется элемент; этот элемент будет истинным, но не доказуемым утверждением!

Теоремы 1 и 3 в совокупности дают условия, налагаемые на фундаментальную пару и необходимые и достаточные для того, чтобы для этой пары можно было ввести полную непротиворечивую дедуктику. **Это условие — перечислимость множества всех истин.** Можно ожидать (и так и оказывается на самом деле), что в «богатых», «выразительных» языках множества всех истин настолько сложны, что неперечислимы, и потому для этих языков невозможны полные непротиворечивые дедуктики. Найденный критерий, однако, не слишком удобен, поскольку рассмотрение множества T всех истин может оказаться затруднительным. Поэтому мы в следующем параграфе переформулируем этот критерий, сделав его более «применимым».

5.3. Простейшие критерии неполноты

Мы знаем теперь, что неперечислимость множества T необходима и достаточна для того, чтобы для $\langle \text{Б}, T\rangle$ не существовало полной и непротиворечивой дедуктики.

Однако нас могут интересовать не все истины языка, а только истины некоторого вида или из некоторого класса, подобно тому как сдающего экзамен по математике интересует истинность не всех математических утверждений, а лишь тех, которые могут ветретиться на экзамене. Например, может представлять интерес построение дедуктики, в которой выводятся все истинные утверждения длиной не больше 1000 и не выводится ни одного ложного утверждения такой длины. При этом выводимость утверждений большей длины в этой додуктике может быть никак не связанном с их истинностью. Кроме того, в некоторых случаях (язык теории множеств) множество истин в полном объеме совершенно неопределенно. Сказанное оправдывает рассмотрение понятий непротиворечивости и полноты в применении к произвольному подмножеству множества $Б^{\infty}$. Перейдем к формальным определениям.

Пусть $\langle Б, T\rangle$—фундаментальная пара, $\langle Д, D, \delta\rangle$ — дедуктика над Б и P — множество всех доказуемых слов. Пусть $V \subseteq Б^{\infty}$. Скажем, что дедуктика $\langle Д, D, \delta\rangle$

а) *непротиворечива применительно к V,* если $V \cap P \subseteq V \cap T$;

б) *полна применительно к V,* если $V \cap T \subseteq V \cap P$.

Теорема 4. *Если V — перечислимое подмножество множества* $Б^{\infty}$, *а множество истинных утверждений, принадлежащих к V, неперечислимо, то никакая дедуктика не является одновременно непротиворечивой и полной применительно к V.*

Доказательство. По условию $V \cap T$ неперечислимо. Для непротиворечивой и полной применительно к V дедуктики $V \cap T = V \cap P$. Но $V \cap P$ обязано быть перечислимым, как это вытекает из теоремы 3 и следующей леммы.

Лемма 5. *Теоретико-множественные объединение и пересечение перечислимых множеств перечислимы.*

Доказательство. Пусть R и S — перечислимые множества. Сначала докажем перечислимость $R \cup S$. Если одно из множеств пусто, то это тривиально. Если оба множества непусты, то $R = \{\rho(0), \rho(1), \ldots\}, S = \{\sigma(0), \sigma(1), \ldots\}$, где ρ и σ — вычислимые последовательности. Тогда вычислимая последовательность f, заданная соотношениями $f(2n) = \rho(n)$, $f(2n+1) = \sigma(n)$, будет перечислять $R \cup S$. Докажем теперь перечислимость $R \cap S$. Если $R \cap S$ пусто, то оно перечислимо по определению. В противном случае существует некоторое a такое, что $a \in R \cap S$, а R и S перечисляются вычислимыми функциями ρ и σ. Поскольку $\mathbb{N}^2$ перечислимо (пример

1 из п.1.5.2), оно перечисляется некоторой вычислимой функцией *g*. Каждое значение *g(n)* есть некоторая пара натуральных чисел; обозначим через $\xi(n)$ и $\eta(n)$ первый и второй члены этой пары. Функции ξ и η, очевидно, вычислимы. Введем функцию *h:*

$$h(n) = \begin{cases} \rho(\xi(n)), & \text{если } \rho(\xi(n)) = \sigma(\eta(n)), \\ a & \text{в противном случае.} \end{cases}$$

Функция *h* вычислима и перечисляет множество $R \cap S$.

Замечание 1. Условие несуществования дедуктики с определенными свойствами, сформулированное в теореме, является не только достаточным, но и необходимым (причем даже без предположения о перечислимости *V).* В самом деле, если $V \cap T$ перечислимо, то полная непротиворечивая дедуктика для $\langle Б,\ V \cap T\rangle$, существующая в силу теоремы 1, будет в то же время полной и непротиворечивой для $\langle Б, T\rangle$ применительно к *V.*

Очевидно, что дедуктика непротиворечива (полна) относительно $\langle Б, T\rangle$ тогда и только тогда, когда она непротиворечива (полна) применительно к любому подмножеству множества $Б^{\infty}$. Поэтому для обнаружения неполноты относительно $\langle Б, T\rangle$ непротиворечивой (относительно той же $\langle Б, T\rangle$) дедуктики достаточно (и необходимо) указать такое подмножество *V* множества $Б^{\infty}$, применительно к которому эта дедуктика неполна. Следующее построение помогает найти в ряде важных случаев такое подмножество.

Условимся говорить, что посредством фундаментальной пары $\langle Б, T\rangle$ *выразима принадлежность* к множеству *Q* натуральных чисел, если существует такая определенная на натуральном ряду и принимающая значения в $Б^{\infty}$ вычислимая функция *f (выражающая* эту принадлежность), что:

1) если $n \in Q$, то $f(n) \in T$,
2) если $n \in \mathbb{N} \setminus Q$, то $f(n) \in Б^{\infty} \setminus T$.

Для такой функции *f* множество *V* всех ее значений перечислимо. Поэтому (в силу теоремы 4) не будет существовать полной и непротиворечивой применительно к *V* дедуктики, коль скоро множество $V \cap T$ принадлежащих к *V* истинных утверждений неперечислимо. Неперечислимость же множества $V \cap T$, как мы сейчас увидим, будет иметь место, если неперечислимо множество Q.

Лемма 6 (о полном прообразе). *Пусть f — вычислимая функция, область определения которой есть перечислимое множество, и B— произвольное перечислимое множество. Тогда множесто* $f^{1}(B)$ *перечислимо.*

Доказательство. Если $f^1(B)$ пусто, оно перечислимо по определению. Пусть теперь $c \in f^{-1}(B)$, множество B перечисляется вычислимой функцией h, а область определения функции f — вычислимой функцией g. Для построения перечисления $f^1(B)$ мы поступаем так.

Перебирая $\mathbb{N} \times \mathbb{N}$, для каждой пары $\langle k, l \rangle$ проверяем, переводит ли функция f элемент $g(k)$ (*«k-й* элемент в перечислении области определения *f*») в $h(l)$ (*«l-*й элемент в перечислении *B»);* если да, то включаем $g(k)$ в строимое нами перечисление множества $f^1(B)$, если нет, то включаем в него элемент *c*.

Более формально, пусть ξ и η определены, как в доказательстве леммы 5. Положим

$$\varphi(n) = \begin{cases} g(\xi(n)), & \text{если} \quad f(g(\xi(n))) = h(\eta(n)), \\ c & \text{в противном случае.} \end{cases}$$

Легко видеть, что φ — вычислимая функция, перечисляющая множество $f^1(B)$.

Вернемся теперь к рассмотрениям, предшествующим формулировке леммы 6. Заметим, что $Q = f^{-1}(V \cap T)$. Поэтому если Q неперечислимо, то неперечислимо и $V \cap T$ (в противном случае в силу леммы 6 было бы перечислимо и Q). Таким образом (принимая во внимание теорему 4) нами доказана следующая

Теорема 5. *Если посредством фундаментальной пары* $\langle Б, T \rangle$ *выразима принадлежность хотя бы к одному неперечислимому множеству натуральных чисел, для* $\langle Б, T \rangle$ *не может существовать непротиворечивой и полной дедуктики; более того, не существует дедуктики, являющейся одновременно непротиворечивой и полной применительно к множеству значений функции, выражающей указанную принадлежность.*

Замечание 2. Достаточное условие несуществования, сформулированное в теореме 5, является и необходимым. В самом деле, если для $\langle Б, T \rangle$ нельзя ввести полную непротиворечивую дедуктику, то T неперечислимо (теорема 1); $Б^{\infty}$ неречнелимо (пример 2 из п.1.5.2) и перечисляется некоторой вычислимой функцией f. Поскольку $T = f(f^{-1}(T))$, то множество $f^1(T)$ неперечислимо (в силу теоремы 3 об образе перечислимого множества). В то же время функция f выражает принадлежность к $f^1(T)$ посредством пары $\langle Б, T \rangle$.

5.4. Язык арифметики

В этом параграфе мы приложим построения предыдущих параграфов к языку арифметики. Содержательно, **под языком арифметики понимается язык, утверждения которого формулируются (с помощью логических операций и отношения равенства) в терминах натуральных чисел и операций сложения и умножения**. На формальном уровне нам надлежит предъявить соответствующую фундаментальную пару. Разумеется, задача построения такой пары не может иметь однозначного решения: ясно, что возможны различные алфавиты для записи одной и той же сути. Здесь будет избран 14-буквенный алфавит A (арифметический алфавит), буквами которого служат следующие знаки:

1° —2° скобки (и);

3° знак для образования цифр |; ′ ,

4° знак для образования переменных *x;*

5°—6° знаки сложения + и умножения •;

7° знак равенства =;

8°—14° логические знаки $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, $\exists$, $\forall$ (при содержательной интерпретации эти знаки будут иметь следующий смысл: «неверно, что», «и», «или», . «если..., то», «эквивалентно», «существует такой.., что», «для всех»).

Чтобы выделить надлежащее множество истинных утверждений, нам придется предпринять вначале некоторые рассмотрения синтаксического характера: мы должны будем выделить определенные классы слов в A^{∞} и заняться их строением.

Слово вида $\underbrace{a \ldots a}_{n \text{ раз}}$, где a — какая-то буква, будем обозначать через a^n.

При $n = 0$ слово a^n пусто — не содержит ни одной буквы. *Цифрами* будем называть слова вида $(|^n)$, где $n \geqslant 0$, а *переменными* — слова вида (x^n), где $n > 0$. При интерпретации языка слово $(|^n)$ будет служить записью числа n, а слово (x^n) будет одной из переменных, пробегающих натуральный ряд (для записи утверждений арифметики может потребоваться сколь угодно много таких переменных). Введем теперь следующее индуктивное определение *терма:*

1° все цифры и все переменные суть термы;

2° если t и u — термы, то $(t + u)$ и $(t \cdot u)$ суть термы.

Параметрами терма будем называть все переменные, входящие в нею. Терм, не имеющий параметров, будем называть *постоянным.*

Пример 1. Терм $((|||)\cdot(||))$ — постоянный, а термы, $((\ )\cdot(x))$ и $((|||)\cdot(xx))$ — не постоянные: параметром первого из них является *(x),* параметром второго — *(xx).*

Каждому постоянному терму естественно поставить в соответствие число, его *значение,* по следующим правилам:

1° значением цифры $(|^n)$ является число *n;*

2° значением постоянного терма *(t + u)* служит сумма значений постоянных термов *t* и *u,* а значением постоянного терма $(t\cdot u)$ служит произведение значений постоянных термов *t* и *u.*

Пример 2. Значением постоянного терма $((|||)+((|)\cdot(||)))$ служит число 5.

Всякое слово вида *(t — u),* где *t* и *u* суть термы, будем называть *элементарной формулой.* Наконец, введем следующее индуктивное определение *формулы:*

1° все элементарные формулы суть формулы;

2° если α есть формула, то $\neg\alpha$ есть формула;

3° если α и β суть формулы, то $(\alpha\wedge\beta)$, $(\alpha\vee\beta)$, $(\alpha\to\beta)$ и $(\alpha\leftrightarrow\beta)$ суть формулы,

4° если α есть формула, а ξ есть переменная, то $\exists\xi\alpha$ и $\forall\xi\alpha$ суть формулы.

П р и м е р 3 Слово

$$(\exists(x)\forall(xx)\neg((x)=(xx))\leftrightarrow\forall(xx)((x)=(xx)))$$

является формулой.

Для облегчения чтения мы будем в дальнейшем записывать термы и формулы сокращенно, заменяя $(|^n)$ 1 на *n, (x^n)* на x_n и опуская внешние скобки; например, формула из примера 3 может быть сокращенно записана так:

$$\exists x_1\forall x_2\neg(x_1=x_2)\leftrightarrow\forall x_2(x_1=x_2).$$

Среди формул и будут выделяться истинные утверждения. Но сначала нам потребуются сравнительно более технические понятия параметров формулы и подстановки цифры вместо переменной.

Мы сопоставим каждой формуле некоторое конечное множество переменных, элементы которого будут называться *параметрами* формулы. Множества параметров формул определяются индуктивно по следующим правилам:

1° параметрами элементарной форхмулы *(t — u)* являются все параметры терма *t,* а также все параметры терма *u;*

2° у формулы $\neg\alpha$ те же параметры, что у формулы α;

3° параметрами формул $(\alpha \wedge \beta)$, $(\alpha \vee \beta)$, $(\alpha \to \beta)$ и $(\alpha \leftrightarrow \beta)$ являются все параметры формулы α, а также все параметры формулы β;

4° параметрами формул $\exists \xi \alpha$ и $\forall \xi \alpha$ являются все параметры формулы α, отличные от ξ.

Пример 4. Единственным параметром формулы из примера 3 является x_1. В самом деле, параметрами формулы $(x_1 = x_2)$ являются x_1 и x_2, те же параметры у формулы $\neg(x_1 = x_2)$; параметром формулы $\forall x_2 \neg(x_1 = x_2)$ является x_1, формула $\exists x_1 \forall x_2 \neg(x_1 = x_2)$ не имеет параметров; с другой стороны, единственным параметром формулы $\forall x_2 (x_1 = x_2)$ является x_1.

Менее формально параметры можно описать как переменные, входящие в формулу свободно, т. е. не попадающие в область действия одноименных кванторов. Формулы, не имеющие параметров, называются *замкнутыми формулами* или *суждениями*. Формула из примера 3 не является суждением. Суждения интерпретируются как высказывания о свойствах натурального ряда; каждому из них приписывается значение «истина» или «ложь» согласно описанному ниже способу (согласующемуся с подразумеваемым смыслом символов, входящих в формулы). Те суждения, которым окажется приписанным значение «истина», и будут служить «испитыми утверждениями арифметики».

Пример 5. Предложение «для всякого натурального числа, кроме нуля, найдется меньшее натуральное число» может быть переведено следующей формулой арифметики:

$$\forall x_1 (\neg(x_1 = 0) \to \exists x_2 \exists x_3 (\neg(x_3 = 0) \wedge ((x_2 + x_3) = x_1))).$$

Свойство «x_2 меньше x_1» приходится (из-за отсутствия в языке символа <) записывать косвенно:

$$\exists x_3 (\neg(x_3 = 0) \wedge ((x_2 + x_3) = x_1)).$$

Прежде чем мы перейдем к установлению значений замкнутых формул, нам понадобится еще одно, на этот раз последнее, техническое понятие — понятие *результата подстановки цифры n вместо переменной w в формулу* α. Этот результат является формулой, обозначается $S^w_n \alpha$ и определяется индуктивно по следующим правилам:

1° результатом подстановки *n* вместо *w* в элементарную формулу $(t = u)$ является результат замены всех вхождений переменной *w* на цифру *n*;

2° $S^w_n \neg \alpha = \neg S^w_n \alpha$;

3° если λ — любой из символов $\wedge$, $\vee$, $\to$, $\leftrightarrow$, то

$$S_n^w(\alpha\lambda\beta) = (S_n^w\alpha\lambda S_n^w\beta);$$

4° результатом подстановки *n* вместо *w* в формулу $Q\xi\alpha$, где Q—один из знаков $\forall$, $\exists$, а ξ — переменная, является формула $Q\xi S_n^w\alpha$, если переменная *w* отлична от переменной ξ; в противном случае (если *w* и ξ — одна и та же переменная) результат будет совпадать с исходной формулой $Q\xi\alpha$.

Пример 6. Если α — формула из примера 3, то $S_3^{x_1}\alpha$ есть $\exists x_1\forall x_2\neg(x_1 = x_2) \leftrightarrow \forall x_2(5 = x_2)$, а $S_1^{x_2}\alpha$ есть α. Заметим, что если бы мы вместо всех вхождений x_1 в формулу α подставили 5, то получили бы слово $\exists 5\forall x_2\neg(5 = x_2) \longrightarrow \forall x_2(5 = x_2)$, не являющееся формулой. (Согласно нашему определению, при подстановке вместо *w* следует оставлять без изменений те вхождения, которые попадают под действие кванторов $\exists w$ и $\forall w$.)

Л е м м а 7. *Параметрами формулы* $S_n^w\alpha$ *являются параметры формулы* α*, отличные от* w.

Это очевидное утверждение может быть формально доказано индукцией по построению (или по длине) формулы α.

Теперь мы уже в состоянии перейти к снабжению суждений *значениями.* Как уже говорилось, таких значений будет два — «истина» *(И)* и «ложь» *(Л).* Суждения, имеющие значение *И,* будем называть истинными, имеющие значение *Л* — ложными. Значения формул определяются индукцией по построению формул следующим образом:

1° суждение *(t — u)* истинно, если значения постоянных термов *t* и *u* равны; в противном случае оно ложно;

2° суждение $\neg\alpha$ истинно, если суждение α ложно; в противном случае оно ложно;

3° суждение $(\alpha \wedge \beta)$ истинно, если оба суждения α и β истинны; в противном случае оно ложно;

4° суждение $(\alpha \vee \beta)$ истинно, если хотя бы одно из суждений α и β истинно; в противном случае суждение $(\alpha \vee \beta)$ ложно;

5° суждение $(\alpha \rightarrow \beta)$ ложно, если суждение α истинно, а суждение β ложно; в противном случае суждение $(\alpha \rightarrow \beta)$ истинно;

6° суждение $(\alpha \leftrightarrow \beta)$ истинно, если значения суждений α и β одинаковы; в противном случае суждение $(\alpha \leftrightarrow \beta)$ ложно;

7° суждение $\exists\xi\alpha$ истинно, если существует такая цифра *n,* что суждение $S_n^\xi\alpha$ истинно; если такой цифры нет, то суждение $\exists\xi\alpha$ ложно;

8° суждение $\forall\xi\alpha$ истинно, если для любой цифры *n* суждение $S_n^\xi\alpha$ истинно; в противном случае суждение $\forall\xi\alpha$ ложно.

Заметим (это относится к 7° и 8°), что $S_n^{\xi}\alpha$ является суждением, так как формула α не имеет параметров, отличных от ξ (иначе $\exists\xi\alpha$ и $\forall\xi\alpha$ не были бы суждениями).

Пример 7. Формула из примера 5 истинна. Формула из примера 3 не является ни истинной, ни ложной, поскольку не является суждением; однако результат подстановки в нее вместо переменной x_1 любой цифры является истинным суждением.

Истинные суждения мы и объявим истинными утверждениями арифметики. Обозначая их множество буквой *T,* мы приходим к фундаментальной паре $\langle A, T\rangle$ языка арифметики. Нас будет интересовать возможность ввести для этой пары полную непротиворечивую дедуктику. Мы покажем, что это невозможно, ссылаясь на критерий, установленный в предыдущем параграфе.

Итак, нам надо показать, что существует такое неперечислимое множество натуральных чисел, принадлежность к которому выразима посредством только что введенной фундаментальной пары $\langle A, T\rangle$. С этой целью мы введем в рассмотрение некоторый класс множеств, принадлежность к которым заведомо выразима посредством $\langle A, T\rangle$, а затем попытаемся установить наличие в этом классе неперечислимого множества. Класс, о котором идет речь, — класс так называемых арифметических множеств — вводится следующим образом.

Пусть α — формула, не имеющая параметров, кроме, быть может, переменной x_1. Тогда для каждой цифры *n* формула $S_n^{x_1}\alpha$ является суждением — истинным или ложным. Рассмотрим множество всех тех и только тех цифр *n*, для которых $S_n^{x_1}\alpha$ — истинное суждение. Будем говорить, что это множество *сопряжено* с формулой α. Каждое множество цифр (а также соответствующее множество чисел), сопряженное с некоторой формулой языка арифметики, будем называть *арифметическим по Гёделю* или, короче, просто *арифметическим.*

Арифметические множества обладают рядом очевидных свойств:

Свойство 1. Дополнение к арифметическому множеству (до натурального ряда $\mathbb{N}$) есть арифметическое множество. В самом деле, если *M* сопряжено с α, то $(\mathbb{N} \setminus M)$ сопряжено с $\neg\alpha$.

Свойство 2. Объединение и пересечение арифметических множеств суть арифметические множества. В самом деле, если M_1 и M_2 сопряжены с α_1 и α_2, то $M_1 \cap M_2$ сопряжено с $(\alpha_1 \wedge \alpha_2)$, а $M_1 \cup M_2$ — с $(\alpha_1 \vee \alpha_2)$.

Свойство 3. Принадлежность к произвольному арифметическому множеству выразима посредством $\langle A, T\rangle$. В самом деле, пусть множество *M* сопряжено с формулой ее. Определим функцию *f*

следующим образом: значение f на цифре n есть слово $S_n^{x_1}\alpha$. Тогда f будет вычислимой функцией, выражающей принадлежность к множеству M.

Ключевым пунктом излагаемого нами доказательства теоремы Гёделя является следующее утверждение:

$(*)$ *существует неперечислимое арифметическое множество.*

Обоснование этого утверждения $(*)$ мы отложим до следующего параграфа. А сейчас заметим, что из него в силу свойства 3 и теоремы 5 вытекает, что

для фундаментальной пары $\langle A, T\rangle$ *языка арифметики нельзя ввести полной непротиворечивой дедуктики.*

Этот результат может быть назван теоремой Гё деля о неполноте для формальной арифметики. Он показывает, что для любого точно сформулированного понятия доказательства найдется либо доказуемое, по ложное утверждение, формулируемое на языке арифметики, либо истинное утверждение того же языка, не являющееся доказуемым.

Замечание 1. Пусть M — неперечислимое арифметическое множество. Как гласит вторая часть теоремы 5, не существует дедуктнки, одновременно непротиворечивой и полной применительно к множеству V значений произвольной вычислимой функции f выражающей принадлежность к M. Таким образом, для непротиворечивой дедуктики уже среди членов последовательности $f(0),\ f(1),\ f(2),\ \ldots$ непременно встретятся истинные, но не доказуемые утверждения

В качестве f, как мы только что видели, можно взять функцию $n \mapsto S_n^{x_1}\alpha$, где M сопряжено с α. При таком выборе f слово $f(n)$ естественно интерпретируется как утверждение «$n \in M$». Поэтому, говоря неформально, истинное недоказуемое утверждение можно найти (для любой непротиворечивой дедуктики!) среди утверждении вида «$n \in M$». В следующем параграфе мы увидим, что M может быть выбрано так, что его дополнение E до натурального ряда $\mathbb{N}$ окажется перечислимым. Итак, существует такое перечислимое множество E, что среди истинных утверждений вида «n не принадлежит E» для любой непротиворечивой дедуктики найдется недоказуемое (заменить в этой формулировке «n не принадлежит E» на «n принадлежит E» было бы ввиду теоремы 1 невозможно).

Замечание 2. Многие определения в этом параграфе используют индукцию по построению термов и формул. При этом возникает следующая трудность: представим себе, например, что слово X имеет вид $(\alpha \wedge \beta)$ и одновременно имеет вид $(\alpha' \to \beta')$, где α, β,

α', β' — некоторые формулы. В этом случае требования пунктов индуктивного определения, касающихся формул вида $(\alpha \wedge \beta)$ и формул вида $(\alpha' \to \beta')$, могут противоречить друг другу.

Поэтому, давая индуктивные определения, мы должны быть уверены в однозначности анализа термов и формул, т. е. в том, что указанные в определении терма (или формулы) случаи исключают друг друга и что в тех из них, в которых терм или формула получается в результате комбинации двух термов или формул, комбинируемые термы или формулы восстанавливаются однозначно. Именно для этой цели в формулах употребляются скобки. (В естественном языке похожую роль играют знаки препинания. Постановка запятой в известной фразе «казнить нельзя помиловать» есть фактически выбор между «казнить $\wedge$ нельзя помиловать» и «казнить нельзя $\wedge$ помиловать». Впрочем, иногда двусмысленность не может быть устранена расстановкой знаков препинания: в фразе «Он из Германии туманной привез учености плоды» эпитет «туманной» можно отнести и к Германии, и (что менее очевидно) к учености.

Для формального доказательства однозначности анализа полезно следующее вспомогательное утверждение:

число открывающихся скобок в терме или формуле равно числу закрывающихся; если слово X является началом терма или формулы, не совпадающим со всем термом или со всей формулой, то число открывающихся скобок, в X больше числа закрывающихся.

5.5. Три аксиомы теории алгоритмов

5.0. Наша цель теперь — доказать утверждение (*) из предыдущего параграфа. Однако наших расплывчатых представлений об алгоритмах, которыми мы довольствовались до сих пор, недостаточно для этой цели. Традиционный путь состоит в том, чтобы обратиться к одному *т* так называемых «уточнений» понятия алгоритма, т. е. заменить несколько неопределенное, но зато совершенно общее понятие алгоритма, которым мы все время пользовались, достаточно точным, но зато и более узким, понятием «алгоритма специального вида».

Это более узкое понятие провозглашается, впрочем, равносильным первоначальному, широкому, в том точном смысле, что классы вычислимых функции, возникающие на базе каждого из этих понятий, совпадают (а следовательно, совпадают и классы перечистимых множеств). Указанное совпадение воспринимается не как теорема, подлежащая доказательству, а как гипотеза, проверяемая на практике

После этого строится точная математическая теория функций, вычисляемых «алгоритмами специального вида» (технически наиболее сложным при этом оказывается доказательство утверждений, аналогичных утверждениям задач 9 и 10 к приложению В). Недоказываемая догма о совпадении классов таких функции с классом всех вычислимых функций служит лишь для обоснования значимости построенной теории.

Об одном из таких понятий «алгоритмов специального вида» см. далее в приложении В. Здесь мы изберем другой путь: не привязывая изложения к тому или иному специальному классу алгоритмов, мы вместо этого попытаемся сформулировать некоторые ограничения, налагаемые на наши первоначальные представления об алгоритмах. Эти ограничения будут сформулированы в виде трех аксиом: аксиомы протокола, аксиомы программы и аксиомы арифметичности.

5.1. Первая аксиома. Рассмотрим процесс применения какого-либо алгоритма A к исходному данному x с получением результата y. Мы предполагаем, что все промежуточные выкладки, весь процесс вычисления, ведущий от x к y, можно запротоколировать так, чтобы этот протокол содержал исчерпывающую информацию о последовательных этапах процесса.

Пример 1. При работе вычислительной машины, в целях проверки ее работы, часто бывает нужно выдать наружу, «на печать», не только конечный результат, но и все промежуточные результаты. Получаемый таким способом «протокол работы машины» будет словом в выходном алфавите машины — с добавлением, если нужио, знака пробела, знака новой строки и т. и

Пример 2. Желая проверить, правильно ли усвоен обучающимися алгоритм сложения чисел столбиком, мы можем требовать, чтобы в своих письменных работах они не только указывали конечный результат, но и записывали в определенной системе записи все свои действия. Можно договориться о такой системе записи вычислений, чтобы для сложения, например, чисел 68 и 9967 протокол выглядел так:

			1	11	111	1111	1111	
68,	9967	68	68	68	68	68	68	10035
		9967	9967	9967	9967	9967	9967	
			5	35	035	0035	10035	

Каждый из образующих протокол членов есть либо число в десятичной системе (в нашем примере 10035), либо пара чисел (в нашем примере 68, 9967), либо, наконец, четырехэтажное образование вида

11
68
9967
35

(«подвальный» и «чердачный» этажи могут быть и пустыми). Не представляет труда оформить протокол в виде слова в некотором алфавите. Для этого достаточно ввести некоторые дополнительные знаки, с тем чтобы только что изображенный четырехэтажный объект записать прежде в виде таблицы

г

*	*	1	1	*
*	*	*	6	8
*	9	9	6	7
*	*	*	3	5

а затем в виде слова (**11*/***68/*9967/***35).

А весь протокол сложения 68 и 9967 запишем так:

(68 + 9967)(*****/***68/*9967/*-***)(***1*/***68/*9967
/****5)(**11*/***68/*9967/***35)(*111*/***68/*9967
/**035)(1111*/***68/*9967/*0035)(1111*/***68/*9967
/10035)(10035).

При такой системе записи протокол сложения любых двух чисел является словом в 15-буквенном алфавите

$$\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, (,), /, +, *\}.$$

Эти примеры подводят нас к следующим соображениям общего характера. Мы предполагаем, что:

1) для каждого алгоритма ***A*** имеется некоторый алфавит Π_0 (алфавит протоколов) и всевозможные протоколы, фиксирующие работу ***A*** при различных исходных данных из его области применимости, образуют подмножество P_0 множества Π_0^*;

2) существуют такие вычислимые функции α и ω, что для каждого протокола p_0 из P_0 значениями $\alpha(p_0)$ и $\omega(p_0)$ служат соответственно то исходное данное x и тот результат y, для которых составлен данный протокол (т. е. для которых протоколируется переработка x в y);

3) P_0 разрешимо относительно Π_0^*.

Переформулируем сказанное короче, в виде следующей аксиомы, которую и будем называть *аксиомой протокола:*

для каждого алгоритма ***A*** *существуют алфавит* Π_0*, разрешимое подмножество* P_0 *множества* Π_0^∞ *вычислимая функция* α *и вычислимая функция* ω*, обладающие следующим свойством:*

A*(x)=z=y тогда и только тогда, когда существует такое* p_0 *из* P_0*, что* $\alpha(p_0) = x$ *и* $\omega(p_0) = y$.

Эта аксиома имеет непосредственное

Следствие 1. *Область применимости и множество результатов любого алгоритма перечислимы.*

Доказательство. Первое из этих множеств есть $\alpha(p_0)$, а второе — $\omega(p_0)$; оба эти множества перечислимы ввиду лемм 2 и 4 и примера 2 из п.1.5.2.

Следствие 2 (из следствия 1). *Область определения и множество значений любой вычислимой функции перечислимы.*

Следствие 3. *График произвольной вычислимой функции (т. е. множество всех таких пар (x, y), что f(x) = y) есть перечислимое множество.*

Доказательство. Применяем аксиому протокола к алгоритму, вычисляющему *f*, и берем соответствующие множество P_0, функции α и ω. Строим вычислимую функцию ψ, полагая $\psi(p)=<\alpha(p), \omega(p)>$. Замечаем, что график функции *f* совпадает с множеством $\psi(\Pi_0)$, и применяем лемму 4.

Замечание 1. Следствие 2 можно было бы получить из следствия 3, с учетом следствия 2 леммы 4 и того, что область определения функции и множество значений функции представляются соответственно в виде $\text{пр}_1 M$ и $\text{пр}_2 M$, где *M* — график функции.

Замечание 2. Перечислимость графика есть не только необходимое (как это устанавливается следствием 3), но и достаточное условие вычислимости функции. В самом деле, если график пуст, функция нигде не определена и потому вычислима. Если же график функции *f* не пуст и перечисляется вычислимой функцией ψ, то предлагается такой алгоритм, вычисляющий функцию *f:* для того чтобы вычислить значение *f(a),* перебирай пары $\psi(0), \psi(1), \psi(2), \ldots$ до тех пор, пока не получишь пары с первым членом *a;* второй член этой пары и есть *f(a).*

5.2. Вторая аксиома. Функции, аргументы которых лежат в *X,* а значения — в *Y*, принято называть функциями *из X в Y.* Аналогично алгоритмы, у которых возможные исходные данные лежат в *X,* а результаты—в *Y*, будем называть алгоритмами *из X в Y;* в этом случае мы принимаем, что $X = К^\infty$, $Y = Л^\infty$, где К и Л — некоторые алфавиты. Каждый алгоритм на $К^\infty$ в $Л^\infty$ есть предписание, т. е. текст на

русском или каком-либо другом (в частности, искусственном, специально созданном для записи алгоритмов) языке. Хотя в конкретных случаях обычно не возникает сомнений, является или нет данный текст алгоритмом, само понятие предписания слишком неопределенно для того, чтобы мы могли недвусмысленно отличать предписания от непредписаний. Кроме того, у нас нет единого и достаточно точного способа понимать предписания— ведь они могут быть написаны на разных языках, да и в пределах одного языка проблема смысла достаточно сложна. Тем не менее мы предполагаем (и это предположение и составит аксиому программы), что можно выделить четко очерченное множество единообразно понимаемых предписаний (называемых *программами),* причем такое множество, которое было бы представительным в уточняемом ниже смысле. Два алгоритма назовем *равносильными,* если у них совпадают области применимости и для любого объекта из этой области, взятого в качестве исходного данного, совпадают результаты обоих алгоритмов. Множество алгоритмов из $К^{\infty}$ в $Л^{\infty}$ назовем *представительным* (для алфавитов К и Л), коль скоро любой алгоритм из $К^{\infty}$ в $Л^{\infty}$ равносилен некоторому алгоритму из рассматриваемого множества. Под «четко очерченным» множеством будем понимать здесь разрешимое подмножество множества всех слов в некотором алфавите. Под «единообразным пониманием» разумеем наличие алгоритма ***U,*** применяемого к парам (программа *p,* исходное данное *a*> и дающего в качестве своего результата результат применения программы *p* к исходному данному *a.*

Замечание 3. К такой схеме легко сводятся упоминавшиеся уже «уточнения» понятия алгоритма. Каждое такое уточнение состоит, по существу, в том, что указывается некоторое множество $Р_1$ программ, некоторый неформальный алгоритм ***U,*** объясняющий, как применяется программа к заданному исходному объекту; затем провозглашается (в качестве недоказываемой догмы) представительность множества $Р_1$.

Итак, мы предполагаем, что:

1) для каждых двух алфавитов К и Л имеется некоторый алфавит $П_1$ (алфавит программ) и некоторое множество алгоритмов P_1, называемых программами и записанных в алфавите $П_1$ (так что $P_1 \subseteq П_1^{\infty}$);

2) существует алгоритм ***U*** из $П_1^{\infty} \times К^{\infty}$ в $Л^{\infty}$ (алгоритм применения программы) такой, что ***U****(p, a)* есть результат применения *p* к *a;*

3) множество $Р_1$ является представительным;

4) множество $Р_1$ разрешимо относительно $П_1^{\infty}$.

При этом вовсе не предполагается, что алфавит $П_1$ множество P_1 и алгоритм ***U*** могут быть выбраны лишь единственным образом. Всякую

тройку $<\Pi_1, P_1, \boldsymbol{U}>$, где Π_1 — алфавит, P_1 — множество всех программ, записанных в этом алфавите, и $\boldsymbol{U}$ — алгоритм применения программы к аргументу, будем называть *способом программирования из* $К^{\infty}$ в $Л^{\infty}$.
Таким образом, при заданных К и Л возможны различные способы программирования.

Замечание 4. Предположения 1) —4) вовсе не определяют понятия «способ программирования» (это понятие остается понимаемым интуитивно), а лишь указывают некоторые (причем, как показывает более глубокий анализ, еще не все) его свойства и постулируют, что тройка с такими свойствами существует.

Переходим теперь к формулировке второй аксиомы. Но сначала одно обозначение. Пусть $\boldsymbol{G}$ — произвольный алгоритм из $\Pi_1^{\infty} \times К^{\infty}$ в $Л^{\infty}$. Через $\boldsymbol{G}_p$, где $p \in \Pi_1^{\infty}$, обозначим следующий алгоритм из $К^{\infty}$ в $Л^{\infty}$: для любого a из $К^{\infty}$в качестве результата применения $\boldsymbol{G}_p$ к a берем результат применения $\boldsymbol{G}$ к паре $\langle p, a\rangle$ [так что $\boldsymbol{G}_p(a) \simeq \boldsymbol{G}(p, a)$]. С помощью этого обозначения мы можем переформулировать предположения 1) —4) в виде следующей *аксиомы программы:*

для любых двух алфавитов К и Л *существуют алфавит* Π_1, *разрешимое подмножество* P_1 *множества* Π_1^{∞} *и алгоритм* $\boldsymbol{U}$ *из* $\Pi_1^{\infty} \times К^{\infty}$ *в* $Л^{\infty}$, *обладающие следующими свойствами: для всякого алгоритма* $\boldsymbol{A}$ *из* $К^{\infty}$ в $Л^{\infty}$ *найдется такое* p *из* P_1, *что алгоритмы* $\boldsymbol{A}$ *и* $\boldsymbol{U}_p$ *равносильны.*

Эта аксиома также имеет важные следствия. Но прежде ряд определений.

Пусть I, X и Y — некоторые множества, F — функция из $I \times X$ в Y. Если i — элемент множества I, то через F_i мы будем обозначать функцию из X в Y, которая определена на тех x, для которых пара $\langle i, x\rangle$ лежит в области определения функции F; в этом случае значение F_i на x равно $F(i,x)$. С помощью знака условного равенства сказанное можно записать короче: $F_i(x) \simeq F(i, x)$.

Пусть теперь Ф — некоторый класс функций из X в Y; функцию F из $I \times X$ в F назовем *универсальной* для класса Ф, если выполнены следующие два условия:

1° при всяком $i \in I$ функция F_i принадлежит классу Ф;

2° всякая функция из Ф есть F, при некотором i; иными словами, для всякой $\varphi \in \Phi$ существует такое $i \in I$, что для всех $x \in X$ верно условное равенство $\varphi(x) \simeq F(i, x)$.

Следствие 1 аксиомы программы. *Пусть* К *и* Л — *два алфавита,* Ф — *семейство всех вычислимых функций из* $К^{\infty}$ *в* $Л^{\infty}$. *Тогда существует вычислимая функция из* $\mathbb{N} \times К^{\infty}$ *в* $Л^{\infty}$, *универсальная для класса* Ф.

Доказательство. Условие 1 выполнено автоматически для любой вычислимой функции F (если F вычислима, то и все F_i вычислимы). Поэтому достаточно построить вычислимую функцию F из $\mathbb{N} \times К^{\infty}$ в $Л^{\infty}$, удовлетворяющую условию 2. Рассмотрим алфавит $П_1$, разрешимое подмножество $Р_1$ множества $П_1^{\infty}$ и алгоритм $\boldsymbol{U}$ из $П_1^{\infty} \times К^{\infty}$ в $Л^{\infty}$, существующие по аксиоме программы. Будучи разрешимым подмножеством перечислимого множества, множество $Р_1$ перечислимо (лемма 2); пусть I—перечисляющая его функция. Тогда функция F, определенная соотношением

$$F(i, x) \simeq \boldsymbol{U}(f(i), x),$$

будет искомой. В самом деле, пусть φ — любая вычислимая функция из $К^{\infty}$ в $Л^{\infty}$, $\boldsymbol{A}$ — вычисляющий ее алгоритм: $\boldsymbol{A}(x) \simeq \varphi(x)$ для всех $x \in К^{\infty}$. В силу аксиомы программы существует такое p из P_1, что для всех $x \in К^{\infty}$ выполнено условное равенство

$$\boldsymbol{U}(p, x) \simeq \boldsymbol{A}(x).$$

Так как $p \in P_1$, то $p = f(i)$ при некотором i; для этого i имеет место цепочка условных равенств

$$F(i, x) \simeq \boldsymbol{U}(f(i), x) \simeq \boldsymbol{U}(p, x) \simeq \boldsymbol{A}(x) \simeq \varphi(x),$$

показывающая, что для построенной нами функции F выполнено условие 2° определения универсальной функции.

Частным случаем следствия 1 является

Следствие 2. *Существует вычислимая функция F из $\mathbb{N} \times \mathbb{N}$ в $\mathbb{N}$, универсальная для класса всех вычислимых функций из $\mathbb{N}$ в $\mathbb{N}$*

Следствие 2 получается из следствия 1, если взять в качестве К и Л один из цифровых алфавитов для записи чисел, например алфавит {|}.

Будем говорить, что функции f и g из X в Y всюду отличаются, если ни при каком x из X условное равенство $f(x) \simeq g(x)$ не имеет места (это означает, что для всякого x хотя бы одна из функций f и g определена на x и что если обе функции определены на x, то их значения различны).

Следствие 3 (из следствия *2). Существует такая вычислимая функция d из $\mathbb{N}$ в $\mathbb{N}$, что никакая вычислимая функция из $\mathbb{N}$ в $\mathbb{N}$ не может отличаться от нее всюду.*

Доказательство. Пусть F— универсальная функция из следствия 2. Возьмем в качестве d функцию, определяемую соотношением

$$d(i) \simeq F(i, i).$$

В этом случае $d(i) \simeq F_i(i)$, поэтому d и F_i не могут отличаться всюду; так как любая вычислимая функция из $\mathbb{N}$ в $\mathbb{N}$ есть F; при некотором I, то никакая вычислимая функция из $\mathbb{N}$ в $\mathbb{N}$ не может всюду отличаться от d.

Это следствие может поначалу показаться парадоксальным: казалось бы, функция $d_1(x) \simeq d(x) + 1$ всюду отличается от *d*. Разрешение кажущегося противоречия состоит в том, что *d* — не всюду определенная функция, и на тех *x,* на которых *d* не определена, d_1 не определена тоже, и для этих *x* условное равенство $d_1(x) \simeq d(x)$ выполнено. Однако мы можем рассмотреть не саму функцию d_1, а какое-нибудь всюду определенное продолжение D_1 функции d_1 (это значит, что D_1 — всюду определенная функция, совпадающая с d_1 там, где d_1 определена). Теперь уже D_1 всюду отличается от *d:* если *d(x)* определено, то $d_1(x)$ также определено и равно $d(x)+1$, поэтому $D_1(x) = d(x) + 1 \neq d(x)$; если же *d(x)* не определено, то $D_1(x) \neq d(x)$ уже потому, что левая часть этого соотношения определена, а правая — нет. Не вошли ли мы в противоречие со следствием 3? Нет — мы доказали только, что никакое всюду определенное продолжение функции d_1 не может быть вычислимым, получив тем самым

Следствие 4 (из следствия 3). *Существует вычислимая функция с натуральными аргументами и значениями, не имеющая всюду определенного вычислимого продолжения.*

Пусть *q* — вычислимая функция с натуральными аргументами и значениями, не имеющая всюду определенного вычислимого продолжения; может ли область определения *q* быть разрешимым подмножеством $\mathbb{N}$? Легко понять, что нет: в самом деле, если бы она была разрешимым подмножеством $\mathbb{N}$, то функция Q, определяемая равенством

$$Q(x) = \begin{cases} q(x), & \text{если } x \text{ принадлежит} \\ & \text{области определения } q, \\ 0, & \text{если } x \text{ не принадлежит} \\ & \text{области определения } q, \end{cases}$$

была бы вычислимым всюду определенным продолжением *q.* Итак, область определения функции *q* — неразрешимое множество; согласно следствию 2 аксиомы протокола это множество перечислимо. Таким образом, нами доказано

Следствие 5 (из следствия 4). *Существует перечислимое неразрешимое подмножество натурального ряда.*

Факт существования перечислимого неразрешимого подмножества натурального ряда — один из важнейших фактов теории алгоритмов. Так как подмножество натурального ряда разрешимо тогда и только тогда, когда оно и его дополнение перечислимы (лемма 3}, то предыдущее следствие может быть переформулировано так:

Следствие 6 (из следствия 5). *Существует перечислимое подмножество натурального ряда с неперсчислимым дополнением.*

5.3. Третья аксиома. Если отвлечься от того (весьма существенного) обстоятельства, что на вычислительных машинах могут вычисляться лишь функции, определенные на конечных множествах натуральных чисел (поскольку слишком большие аргументы просто не смогут поместиться в машине), то можно считать, что на этих машинах вычисляются вычислимые числовые функции. Как известно, основными операциями, совершаемыми машиной, являются сложение, умножение и логические операции. Опыт работы на машинах приводит к убеждению, что с помощью этих операций можно запрограммировать любую вычислимую функцию. Следовательно, и всякое перечистимое множество натуральных чисел (как множество значений вычислимой функции) может быть записано в терминах сложения, умножения и логических операций. Сказанное делает естественным формулировку следующей аксиомы, которую мы будем называть *аксиомой арифметичности;*

всякое перечислимое множество натуральных чисел является арифметическим.

Непосредственным следствием этой аксиомы и служит интересующее нас утверждение предыдущего параграфа:

существует арифметическое множество, не являющееся перечислимым.

Таковым является дополнение к множеству из следствия 6 п. 5.2: оно является неперечислимым множеством с перечислимым дополнением. Само это множество будет арифметическим, как дополнительное к арифметическому (1-е свойство арифметических множеств) .

Таким образом, доказательство теоремы о неполноте закончено: как уже отмечалось, из существования неперечислимого арифметического множества следует существование неперечислимого множества, принадлежность которому выразима в арифметике; отсюда следует, что не существует непротиворечивой н полной дедуктики для $<A, T>$ применительно к некоторому перечислимому подмножеству V и, следовательно, никакая непротиворечивая дедуктика не может быть полной для $<A, T>$.

Приложение 1

К ВОПРОСУ О ТОМ, ЧТО ЗНАЧИТ РЕШИТЬ ЛОГИЧЕСКОЕ УРАВНЕНИЕ

Ранее были сформулированы основные вопросы проблематики алгебры логики конца XIX в. Среди этих вопросов выделяется задача о логических уравнениях и исключении неизвестных, алгебраические методы решения которой изложены нами в п..4.1.4, а графические — в п..4.1.6. В связи с этой задачей среди логиков возникла дискуссия, которую мы постараемся осветить в настоящем разделе.

Достаточно ясное представление об этой дискуссии дают работы П. С. Порецкого и Э. Шредера, в которых Порецкий обвиняет Шредера в ошибках и догматизме, а Шредер парирует эти обвинения. Так как Порецкий критикует Шредера, то для понимания этой критики нужно начать с изложения работ Шредера. Заметим, что Порецкий не был знаком с основной работой Шредера — с его «Алгеброй логики», в которой точка зрения Шредера изложена с наибольшей ясностью. Поскольку, однако, точка зрения Шредера на то, что значит решить уравнение по существу не изменялась, мы позволим себе пользоваться именно этим более поздним трудом Шредера.

Заметим, прежде всего, что всякое равенство по Шредеру можно привести относительно переменной x к так называемой «нулевой форме»,— к виду

$$Ax + B\bar{x} = 0, \qquad (1)$$

где A и B не содержат переменной x. Действительно, всякое включение класса A в класс B, т. е. отношение $A \subset B$ можно представить в виде $A\overline{B} = 0$. («Представить в виде» — значит заменить эквивалентным отношением, т. е. таким, которое логически следует из данного и из которого данное в свою очередь следует. Но если $A \subset B$, т. е. всякое A есть B, то нет таких A, которые суть не-B, т. е. $A\overline{B} = 0$. Наоборот, если нет таких A, которые суть не-B, т. е. если $A\overline{B} = 0$, то всякое A есть B, т. е. $A \subset B$. Иными словами, отношения $A \subset B$ и $A\overline{B} = 0$ эквивалентны.) Равенство же $A=B$ эквивалентно двум включением: 1. $A \subset B$ и 2. $B \subset A$ (эти соотношения между равенством и включением были известны еще средневековым схоластическим логикам; в то же время вопрос о том, какое из этих двух соотношений следует принимать в качестве элементарного, а какое — вводить по определению через другое и как именно, не имеет точного смысла: любое из них может быть принято за элементарное, существенно только умение переходить

от равенства к включению, и наоборот). Поэтому равенство $A = B$ эквивалентно равенству $A\bar{B} + \bar{A}B = 0$, т. е. всякое равенство может быть заменено эквивалентным ему равенством в нулевой форме. Остается разложить левую часть полученного равенства по *x,* чтобы получить эквивалентное равенство вида (1).

Среди же равенств этого вида Шредер различал «аналитические», или тождественно истинные,— такие, например, как равенство $\bar{x} \cdot x + x\bar{x} = 0$, — и «синтетические», или истинные только для некоторых *x.* «Аналитические» равенства не могли быть, по Шредеру, уравнениями, поскольку уравнение рассматривалось им как условие, которому должна удовлетворять неизвестная *x,* иными словами, как задача найти такие значения переменной *x* (такие выражения для *x* через какие-нибудь термы: постоянные или переменные), которые при подстановке в уравнение обращали бы его в тождество: делали бы его тождественно истинным — «удовлетворяли» бы ему.

Таким образом, решить уравнение (1) по Шредеру — это означает:

1) выяснить, имеет ли оно решение, т. е. существуют ли такие выражения, которые, будучи подставлены на место *x* в уравнение (1), обращают его в тождество;

2) если такие выражения существуют, то найти какую-нибудь их общую форму (из которой при разных значениях входящих в нее переменных могли бы быть получены все решения уравнения (1)).

При этом ясно, почему Шредер не хочет включить в число уравнений аналитические тождества, т. е. выражения, эквивалентные тому, что *x=x:* они не представляют собой какого-либо условия, которому требовалось бы еще удовлетворить.

В «Алгебре логики» Шредер показывает, что на первый из поставленных выше вопросов отвечает результат исключения из уравнения (1), или резольвента, $AB = 0$, представляющая собой необходимое и достаточное условие разрешимости уравнения (1); на второй вопрос дает ответ форма

$$x = B\bar{u} + \bar{A}u \qquad (2)$$

при любом значении *u.* Действительно, нетрудно убедиться в том, что если существует такое *x,* что $Ax + B\bar{x} = 0$, то, поскольку $Ax + B\bar{x}$ эквивалентно $Ax + B\bar{x} + AB$, *A B* также равно нулю. И наоборот, если $AB = 0$, то достаточно, например, положить $x = \bar{A}$, чтобы удовлетворить уравнению (1). Условие $AB = 0$ является, таким образом, необходимым и достаточным условием разрешимости уравнения (1).

Но если это условие $AB=0$ удовлетворено, то подстановка $x = B\bar{u} + \bar{A}u$ обращает уравнение (1) в

$$A(B\bar{u} + \bar{A}u) + B\overline{(B\bar{u} + \bar{A}u)} = 0,$$

т. е. в

$$A(B\bar{u} + \bar{A}u) + B(\bar{B}\bar{u} + Au) = 0,$$

или в

$$AB\bar{u} + A\bar{A}u + B\bar{B}\bar{u} + BAu = 0.$$

Если AB тождественно равно нулю, то здесь написано $0\cdot\bar{u} + 0\cdot u + 0\cdot\bar{u} + 0\cdot u = 0$, т. е. действительное тождество. При этом существенно, что u — произвольный класс.

Наоборот, если какое-нибудь x удовлетворяет условию (1), то, как нетрудно убедиться, для этого же x верно и соотношение $x = B\bar{u} + \bar{A}u$ при $u = x$. Действительно,

$$Ax + \bar{A}x = x;$$

но так как x удовлетворяет условию (1), $Ax = 0$ и $B\bar{x} = 0$. Поэтому $Ax + \bar{A}x = \bar{A}x + B\bar{x}$, т. е. $x = B\bar{x} + \bar{A}x$, что и требовалось показать.

Таким образом, Шредер фактически показал, — если записать это более современным образом,— что

$$\forall x((Ax + B\bar{x} = 0) \equiv ((AB = 0) \,\&\, \exists u(x = B\bar{u} + \bar{A}u))).$$

(Заметим, что поскольку

$$(\exists u\mathfrak{A}(u) \supset \mathfrak{B}) \equiv \forall u(\mathfrak{A}(u) \supset \mathfrak{B}),$$

то утверждение Шредера о существовании u такого, что x, определяемое равенством (2), удовлетворяет уравнению (1), эквивалентно утверждению о том, что **любое** u, удовлетворяющее равенству (3), дает решение уравнения (1).)

Решить уравнение (1), по Шредеру, оказалось, таким образом, эквивалентным тому, чтобы заменить это уравнение парой соотношений: $AB = 0$, $\exists u(x = B\bar{u} + \bar{A}u)$, в своей совокупности эквивалентных уравнению (1). Само собою разумеется, что соотношение (3) есть при этом логическое следствие из соотношения (1) только в том случае, если (2) берется вместе с квантором существования по u, хотя, наоборот, при всяком u соотношение (1) есть логическое следствие из (2).

Этого обстоятельства, по-видимому, не заметил Порецкий, который высказал, однако, целый ряд существенных и интересных соображений на тему о том, что значит решить логическое равенство. Эти соображения, вместе с его критикой в адрес Шредера, мы позволим себе поэтому привести здесь полностью, хотя, как мы увидим ниже. Порецкий не во всем был прав.

Согласно Порецкому, *«Решить* не тождественное логическое равенство (тождества не могут быть решаемы) значит вывести из него

все или некоторые его логические *следствия.* Решение равенства будет *полное* или *частное,* смотря по тому, *все* или только *некоторые* его следствия нами найдены. Если найдено полное решение и представлено в виде одного равенства, то понятно, что это равенство будет только новою формою первоначального равенства, т. е. оба такие равенства *тождественны,* между собою по своему логическому значению (т. е. касательно объема содержащихся в них сведений об отношениях между данными классами). Отсюда видим, что вопрос о полном решении равенства в сущности есть вопрос о нахождении новой его формы, т. е. о тождественном замещении его некоторым другим равенством. Чтобы судить о том, тождественны между собой, или нет, данные равенства, мы дадим особый критерий, а именно, условимся признавать два равенства тождественными между собой, коль скоро первое есть следствие второго и, обратно, второе есть следствие первого. И вообще, две системы логических равенств мы будем считать между собой тождественными, коль скоро все равенства первой системы могут быть выведены из равенств второй (и обратно) при помощи известных нам логических операций сложения, умножения, отрицания...

Полезно прибавить, что ни полные, ни частные решения логического равенства вовсе не обладают свойством, будучи в него подставленными, обращать его в тождество...

Решить равенство сполна относительно класса *a...* значит тождественно заменить его новым равенством, в левой части которого мы имели бы только *a* ..., а в правой некоторую функцию данных классов *a, b, c, d...».*

Таким образом, Порецкий понимает уравнение не как условие, которому надо удовлетворить, а как посылку, из которой требуется вывести все или некоторые ее логические следствия определенного вида. В соответствии с этим у него получается и другое определение того, что значит решить логическое уравнение.

Само собою разумеется, что задача, поставленная Порецким, имеет не меньший смысл, чем задача, которую решал Шредер. С точки зрения логики, это даже более общая и важная задача: ведь основная задача логики и состоит как раз в выводе логических следствий. Однако обвинения в ошибке, которую усмотрел у Шредера Порецкий, являются неоправданными.

Уравнение Шредера (1) Порецкий пишет в форме

$$0 = ax + a_1 y; \tag{3}$$

где x и y не зависят ни от a, ни от a_1 (индекс 1 у Порецкого, как часто и у Шредера, обозначает дополнение), т. е. заменяет шрёдеровские *A, B,*

x на *x, y, a,* соответственно, и добавляет, что, по Шредеру, равенство (3) «тождественно с парой равенств

$$0 = xy,\ a = x_1(u + y),$$

где *u* неопределенный класс».

Того, что второе из этих равенств понимается у Шредера в смысле выполнимости по *u* (т. е. в смысле *существования* для всякого *a,* удовлетворяющего условию (3), такого *u,* что $a = x_1(u + y)$, Порецкий не заметил, почему и сделал в адрес Шредера очень тяжелый упрек. Действительно, Порецкий пишет: «Построив формулу $a = x_1(u + y)$, где *u* есть собственно *неопределенный* класс, и не видя никаких условий к определению *u,* Шредер начинает считать этот класс *произвольным,* допускающим *всевозможные* значения от 0 до 1, и объявляет, что его формула, при изменении и от 0 до 1, доставляет *всевозможные корни a* уравнения $0 = ax + a_1y$. Здесь можно сказать, что каждое слово есть ошибка, а причина всех ошибок есть недостаточное внимание к различию между понятиями неопределенного и произвольного. Этого мало. Можно доказать, что вовсе нет даже надобности считать *u неопределенным* классом, потому что можно доказать, что $u = a$. В этом отношении мы вполне разделяем мнение Джевонса, по которому во всех тех случаях, когда класс *m* содержится в классе *n,* нет надобности писать, подобно Булю и Шредеру, равенство $m = vn$, где *v* неопределенный класс; достаточно писать так: $m=mn$. Например, фразу «Москва есть город» нет надобности передавать непременно так: «Москва есть *некоторый* город»; вполне достаточно будет сказать: «Москва есть тот город, который есть Москва». Сказать же, будто *«Москва* есть *какой угодно город»,* будет положительной нелепостью.

Если бы Шредер не употреблял неопределенных классов, а брал их подлинное значение, то доказательство его формулы сделалось бы крайне простым. Именно, т. к. равенство $0 = ax + a_1y$ тождественно с парой равенств $0 = ax$, $0 = a_1y$, из которых первое показывает, что *a* содержится в x_1, т. е. $a = ax_1$, а второе, что *y* содержится в *a,* т. е. $a = a + y$, то легко заключить, что $a = ax_1 = (a + y)x_1$.

Это и есть формула Шредера, в которой *u* заменено его подлинным значением, т. е. через *a*. После этого делается вполне очевидным, что ни о различных значениях и, ни о всевозможных корнях *a* уравнения $0 = ax + a_1y$ не может быть и речи».

В сноске к этому месту Порецкий добавляет: «Считаю долгом заметить, что я не сразу понял указанную ошибку Шредера и в своем первом сообщении безразлично называл *u* то произвольным, то неопределенным классом. Но во всяком случае, приводя формулу

Шредера, я воздержался еще и тогда от воспроизведения указанного выше чисто фантастического ее толкования».

Конечно, Порецкий не только критиковал Шредера: он признавал за ним большие заслуги и даже писал: «Что касается оценки способа Шредера, то, без сомнения, способ этот вполне достигает цели и имеет тем большее значение, что представляет первое вполне общее и независящее ни от каких гипотез решение вопроса».

Свое решение задачи, состоящее в замене уравнения (3) эквивалентным ему равенством вида $a = f(a, b, c, \ldots)$, Порецкий называл полным, а решение Шредера точным, и писал даже, что вопрос о том, какое из них следует предпочесть, не допускает категорического решения. «Категорически отвечать на этот вопрос в общем виде мы находим неудобным, потому что в одних случаях и для одних целей может быть предпочтена одна пара формул, в других — другая. Для отличия мы будем называть первую пару *полным,* а вторую *точным* определением *а».*

Под полным определением *а* при этом имелась в виду пара формул

$$a = aM\,(1),$$
$$a = a + M_1\,(0),$$

где *M(a)* — правая часть равенства *1=M (a),* эквивалентного равенству *A = B* (Порецкий указывает способ приведения всякого равенства *A = B* к виду *1 = M (a)).*

Под точным определением *а* имелась в виду пара формул: $a = a\ \ M\,(1),\ a = a + M_1(0)M\,(1).$ (Для уравнения $0 = ax + a_1y\ M\ (a)$ есть $ax_1 + a_1y_1$, поэтому полное решение имеет вид

$$\begin{cases} a = ax_1, \\ a = a + y, \end{cases}$$

а точное —

$$\begin{cases} a = ax_1, \\ a = a + yx_1, \end{cases}$$

т. е., как в этом нетрудно убедиться, приведя, например, все эти равенства к нулевой форме,— точное решение эквивалентно равенству $a = x_1\ (u + a)$ при $u = a$ в решении Шредера и получается из этого решения, если отбросить резольвенту $xy = 0$.)

Признание заслуг Шредера сопряжено, однако, у Порецкого с обвинениями Шредера в излишней математичности и даже формализме. «Тем не менее,— пишет Порецкнй,— нельзя не признать за способом Шредера довольно крупного недостатка, это именно: формальность и искусственность решения. Формула Шредера не выведена, как бы следовало, из анализа существа дела, а искусственно подогнана и оставляет место для сомнений в том, не заключается ли в ней лишних членов». «...сила, а в то же время и слабость,— пишет в

другом месте Порецкий,— способа Шредера заключается в слишком формальном, слишком общем, слишком математичном решении задачи. У него речь идет об одной внешней оболочке дела, сущность которого совершенно игнорируется; кроме того, формула Шредера не представляет гарантий относительно отсутствия в ней лишних членов».

Логический смысл полного и точного решения Порецкий видел при этом в том, что полное решение обнимает все сведения задачи, точное же (относительно класса *а)* относится только к тем сведениям, «которые прямо предназначены к характеристике *а»*. При этом точное решение получалось у Порецкого из полного решения путем отбрасывания из последнего всех членов, входящих в резольвенту (и равных поэтому нулю, когда задача имеет решение). Под сомнениями насчет «лишних членов», очевидно, имелось в виду то обстоятельство, что оставался невыясненным вопрос о том, учтены ли все «логические нули» задачи, т. е. все классы, которые в силу условий задачи должны быть равны нулю.

Критика Порецкого не осталась незамеченной Шредером. Во втором томе своей «Алгебры логики» Шредер специально остановился на методах Порецкого, привел ряд задач, предложенных и решенных Порецким, и написал возражения на его критику. Так как мы во всех подробностях осветили критику Порецкого, то должны предоставить слово и Шредеру. Тем более, что он тоже был во многом прав.

Вот что пишет в свою защиту Шредер: г. Порецкий упрекает меня в ошибке, в том, что в моем «Operationskreis» я недостаточно различал *«неопределенное»* и *«произвольное».* Конечно, он справедливо замечает — в применении к нашей теореме 43) $(a \subseteq b) = \Sigma_u (a = ub)$ {«$(a \subseteq b)$ эквивалентно тому, что существует такое *и, и* что *a=ub»*,— что предложение «Москва есть город» никак не совпадает с предложением «Москва есть *произвольный* город». В главной теореме в «Operationskreis» я во всяком случае объяснил *и* как произвольное.

Но там я имел в виду решение уравнения по такой неизвестной, которую я молча предполагал подлежащей определению из уравнения и только из уравнения, а не решение по некоторому классу, уже где-нибудь и как-нибудь данному. Иначе я должен был бы объяснить про *и,* каким его нужно считать: «неопределенным» или «произвольным». Что я, однако, и тогда уже был далек от того, чтобы не замечать этого различия, показывает мое изложение в «Note iiber den Operationskreis des Logikkalkuls», где я на одном из примеров пояснил именно это различие,— чего Порецкий, очевидно, не знал. Но, если г. Порецкий, что о моем утверждении, состоящем в том, что выражение $x = a_1(u+b)$ охватывает для *и,* меняющегося от 0 до 1, все корни *x*

уравнения $ax + bx_1 = 0$, можно сказать «здесь каждое слово — ошибка», то я могу это только оспаривать и теперь, как и раньше, настаивать на том, что здесь каждое слово правильно. (Необходимое, как ошибочно думает Порецкий, допущение $u = x$ дает только один из бесконечного в общем случае числа корней, и при том, когда мы понимаем под x некоторый определенный, уже иным способом данный, класс,— именно этот, данный, класс.) — Сделанный моему методу упрек в том, что он формалистичен, недостаточно близок к естественному мышлению и производит впечатление искусственности, не в меньшей мере относится к способу г. Порецкого, как дуально соответствующему моему; и даже относится к его способу в еще большей мере, поскольку Порецкий работает еще с двойственными выражениями громоздких булевых схем разложения (как было показано, от этого упрека впервые свободен только метод McColl — Peirce'a). — Этим отнюдь не отрицаются заслуги автора как первого исследователя и творца в области логики в большом славянском государстве».

Под «двойственными выражениями для булевых схем разложения» Шредер, очевидно, имеет в виду выражения равенств не в «нулевой», а в двойственной к нулевой — «единичной» форме, т. е. в виде равенства $M = 1$, эквивалентного данному равенству $A = B$. Мы видим, таким образом, что вопрос о том, что значит решить логическое уравнение, действительно, представлял существенные трудности и мог решаться — да и решался на самом деле — разными авторами по-разному. Порецкий, в частности, не считал решения, предлагавшиеся Булем, Джевонсом, Венном и другими логиками, окончательными и полными. Он требовал прежде всего выяснения того, о каком решении: полном или точном, идет речь, и возражал против всяких решений с неопределенными классами. Он, в действительности, много содействовал тому, что задачей логики стало не решение уравнений и исключение неизвестных, а вывод логических следствий.

Естественно возникает вопрос о том, в какой связи находятся между собою решение уравнений и вывод логических следствий. Обычно нас ведь не просто интересуют логические следствия из какой-нибудь системы посылок: перед нами стоят некоторые определенные вопросы, ответ на которые мы ищем. Бывает и так, что мы хотим узнать, существует ли объект, удовлетворяющий определенным условиям, и тогда задача естественно приводится к тому, что мы формулируем это условие и ищем объект, удовлетворяющий ему. Если нам удается сформулировать это условие в виде равенства, то задача приводится к решению уравнения: к поиску такого объекта, который ему удовлетворяет.

Не следует думать, что всякая логическая задача обязательно приводится к решению уравнения или неравенства, но не следует отвергать и такие задачи, которые сводятся к решению уравнений или неравенств. Класс этих задач довольно обширен и в последнее время все более и более расширяется,— см., например, работу Р. Л. Гудстейна, определяющего логические связки через отношение равенства и строящего, таким образом, логику без специфически логических знаков, или работы А. И. Таутса, сводящего разрешимые случаи формул исчисления предикатов к решению уравнений. Но и в случаях, где речь идет просто о выводе логических следствий из некоторой информации, последняя часто может быть выражена с помощью уравнений или неравенств, преобразование которых к некоторому определенному виду может рассматриваться,— теперь уже в точности следуя Порецкому,— как решение уравнений (соответственно, неравенств). Алгоритмы этого решения во многих случаях могут быть сформулированы на диаграммах Венна.

Как именно это происходит, мы далее и попытаемся объяснить на некоторых примерах. Речь прежде всего пойдет о решении так называемой «задачи Венна».

Этой задаче,— приведенной Венном в статье «Логическая система Буля» в 1876 году с целью проиллюстрировать необходимость преподавания общих методов символической логики,— почти все логики конца XIX века уделяли много внимания (она решается в «Основах науки» Джевонса, в «Алгебре логики» Шредера, Порецкий называет ее «известной» и неоднократно возвращается к ней.

Задача формулируется следующим образом:

Известно, что все члены совета акционерного общества являются или владельцами облигаций, или владельцами акций, но не теми и другими одновременно. Кроме того известно, что все владельцы облигаций входят в совет. Спрашивается, какое можно вывести из этого заключение.

Пусть a — члены совета, b и c — владельцы, соответственно, облигаций и акций. Тогда посылки задачи можно записать символически в виде:

$$(1)\quad a = a\,(b\bar{c} + \bar{b}c),$$

$$(2)\quad b = ba\,.$$

Нетрудно построить диаграмму Венна, соответствующую условиям (1) и (2). В предложении (1) говорится, что та часть класса a, которая не входит в классы b и $\bar{c}$ (одновременно) или в $\bar{b}$ и c, пуста, т. е. в классе a пусты ячейки $ab\bar{c}$ и abc. Условие (2) говорит о том, что ячейки класса

b, не входящие в *a*, пусты. Таким образом, мы получаемдиаграмму (рис. 1), в которой заштрихованы ячейки

$\bar{a}b\bar{c}$, abc и $\bar{a}b$.

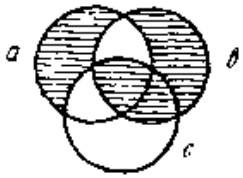


Рис. 1

На рис. 1 видно, что класс *bc* пуст, т. е. лиц, обладающих одновременно облигациями и акциями, нет. Это и есть тот ответ, который удовлетворял Венна и который (как он пишет) получили только пять учеников из 150-ти.

В связи этой задачей Порецкий высказывает критические замечания. Прежде всего он пишет, что «Венн заканчивает задачу *неопределенным* вопросом, вследствие чего нет указания, в *каком направлении* надо мыслить, чтобы получить достойный внимания результат. Очевидно, для получения полного ответа на этот вопрос обязательно исследовать задачу по *всем направлениям»*. И далее по поводу полученного Венном ответа Порецкий продолжает: «... собственно постановка задачи требует, чтобы было доказано, что *никаких других* интересных заключений из данных посылок вывести нельзя» (там же). К этим словам он добавляет в сноске: «На наш взгляд, то обстоятельство, что в данном случае, как мы доказали, простое равенство $b = a\bar{c}$ вполне выражает все условия задачи, также весьма интересно и гораздо более важно, чем результат, полученный самим Венном».

Здесь имеется в виду то обстоятельство, что равенство $b = a\bar{c}$ эквивалентно условию задачи, между тем, как ответ, полученный Венном, есть только логическое следствие из этого условия. Конечно, то обстоятельство, что $bc = 0$ не исчерпывает условий задачи, также может быть обнаружено с помощью диаграмм Венна, т. к. равенству $bc = 0$ соответствует — в случае переменных *a, b* и *c* — диаграмма (рис. 2), отличная от диаграммы, соответствующей посылкам.

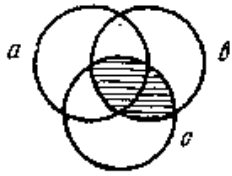


Рис. 2

Из последней видно, что в ячейке *b* остается непустой только часть $a\bar{c}$, т. е. что $b = a\bar{c}$. Диаграмма, соответствующая этому заключению, в точности совпадает, как нетрудно проверить, с диаграммой, соответствующей условию задачи (рис. 1). Нельзя не признать поэтому критику Порецкого вполне оправданной. По-видимому, и Венн заметил, что эта его задача недостаточно хороша и вряд ли может служить целям оправдания преимущества символической логики перед традиционной. Во всяком случае, он не поместил ее среди разнообразных примеров в своей «Символической логике».

Между тем, в действительности диаграммы Венна очень нетрудно применить и для получения полного решения в смысле Порецкого относительно любого класса *x*. Для этого достаточно очертить этот класс на диаграмме Венна, соответствующей условиям задачи, и представить его в виде объединения всех незаштрихованных в нем частей, а затем поступить аналогично склассом $\bar{x}$. Вся информация, содержащаяся в условиях задачи, будет, таким образом, выражена в двух равенствах: $x = \mathfrak{A}$ и $\bar{x} = \mathfrak{B}$, или же в равенствах: $x = \mathfrak{A}$ и $x = \bar{\mathfrak{B}}$, откуда $\mathfrak{A} = \bar{\mathfrak{B}}$. Ясно, что и, наоборот, если $x = \mathfrak{A}$ и $\mathfrak{A} = \bar{\mathfrak{B}}$, то $x = \bar{\mathfrak{B}}$ или $\bar{x} = \mathfrak{B}$, т. е. что равенства $x = \mathfrak{A}$, $\mathfrak{A} = \bar{\mathfrak{B}}$ также дают полное решение задачи в смысле Порецкого. Если за класс *x* мы возьмем в данном случае *b*, то, как это видно из рис. 1, получим (т. к. $a\bar{b}\bar{c} = 0$, $bc = 0$ и $\bar{a}b\bar{c} = 0$) $b = a\bar{c}$, $\bar{b} = \bar{a} + c$, т. е. окажется, что решение $b = a\bar{c}$ является полным в смысле Порецкого. Полное решение для *a* имеет вид: $a = ab\bar{c} + a\bar{b}\bar{c}$, $\bar{a} = \bar{a}\bar{b}$; для с оно дает $c = \bar{b}\bar{c}$, $c = ab\bar{c} + \bar{a}\bar{b}\bar{c}$; таким образом, полное решение для *b* более короткое (по записи).

Кроме полного решения равенства относительно класса *x*, Порецкий рассматривает точное решение этого равенства относительно класса *x*. Остановимся на этом вопросе более подробно. Уравнение $Ax + B\bar{x} = 0$ представим на диаграмме Венна трех переменных *A, B, x* (рис. 3).

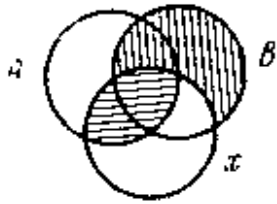


Рис. 3

Это уравнение имеет решение относительно *x,* как мы видели выше, тогда и только тогда, когда $AB = 0$ (диаграмма для резольвенты $AB = 0$ в случае трех переменных *A, B m x* показана на рис. 4).

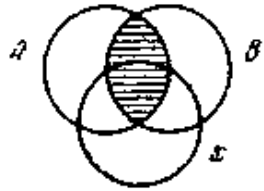


Рис. 4

Полное решение для *x* равенства (1) легко определяется графически (рис. 3) в виде пары равенств:

$$\begin{cases} x = \bar{A}x, \\ \bar{x} = \bar{B}\bar{x}, \end{cases} \tag{4}$$

или

$$\begin{cases} x = \bar{A}x, \\ x = B + x. \end{cases} \tag{5}$$

У Порецкого пара (5) записывается, как отмечалось, в форме (относительно класса *x).*

$$\begin{cases} x = xM\,(1), \\ x = x + M_1\,(0), \end{cases}$$

и показывается, что эта пара эквивалентна («вполне тождественна») с равенством

$$x = xM\,(1) + x_1 M_1\,(0). \tag{6}$$

(Для сравнения см. решение Шредера (2), из которого (6) получается после замены *u* на *x, B* на M_1 (0) и $\bar{A}$ на M (1).)

Последнее утверждение (об эквивалентности) может быть получено графически как следствие более общего предложения об эквивалентности пары равенств одному.

Пусть

$$\begin{cases} x = \mathfrak{A}, \\ x = \mathfrak{B} \end{cases} \tag{7}$$

— пара равенств. Диаграмма трех переменных $\mathfrak{A}$, $\mathfrak{B}$, x, соответствующая (7), построена на рис. 5, где горизонтальная штриховка соответствует первому равенству, вертикальная — второму.

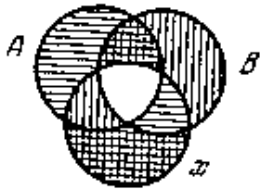


Рис. 5

Предположим, что равенства (7) таковы, что горизонтальная и вертикальная штриховка на диаграмме (рис. 5) не пересекаются, т. е., что $\mathfrak{A}\mathfrak{B}\bar{x}$ и $\bar{\mathfrak{A}}\bar{\mathfrak{B}}x$ тождественно равны нулю $(\mathfrak{A}\mathfrak{B}\bar{x} + \bar{\mathfrak{A}}\bar{\mathfrak{B}}x = 0)$.

Равенства, входящие в полное решение, определяемое по диаграмме, этим свойством обладают.

Тогда информация, записанная на рис. 5, может быть представлена равенством

$$\mathfrak{A}\bar{\mathfrak{B}} + \bar{\mathfrak{A}}\mathfrak{B} = 0. \tag{8}$$

Равенство (8) эквивалентно, как легко убедиться, равенству

$$x = (\mathfrak{A}\mathfrak{B} + \bar{\mathfrak{A}}\bar{\mathfrak{B}})x + (\mathfrak{A}\bar{\mathfrak{B}} + \bar{\mathfrak{A}}\mathfrak{B})\bar{x}, \tag{9}$$

которое в нашем случае (когда $\bar{\mathfrak{A}}\mathfrak{B}\ x = 0$) можно переписать в виде

$$x = \mathfrak{A}\mathfrak{B}x + (\mathfrak{A}\bar{\mathfrak{B}} + \bar{\mathfrak{A}}\mathfrak{B})\bar{x}. \tag{10}$$

Таким образом, пара равенств (7) при условии

$$\mathfrak{A}\mathfrak{B}\bar{x} + \bar{\mathfrak{A}}\bar{\mathfrak{B}}x = 0$$

эквивалентна равенству (10), имеющему вид $x = f(x, a, \ldots, u)$.

На рис. 3 горизонтальная штриховка соответствует первому равенству в паре (5), вертикальная — второму. Выражению $\mathfrak{A}\mathfrak{B}$ из (10) на диаграмме (рис. 3) соответствует объединение незаштрихованных ячеек, выражению $\mathfrak{A}\bar{\mathfrak{B}}$ — объединение ячеек с вертикальной штриховкой, а выражению $\bar{\mathfrak{A}}\mathfrak{B}$ — с горизонтальной штриховкой. Так по диаграмме (рис. 3) строится равенство вида (10), эквивалентное (5):

$$x = \bar{A}x + B\bar{x} \tag{11}$$

$(\bar{\mathfrak{A}}\mathfrak{B}\bar{x} = 0$, т. к. все ячейки с горизонтальной штриховкой находятся в классе *x)*.

Нетрудно проверить, что равенство (11) является другой формой записи равенства Порецкого (6). Отметим также, что в результате подстановки в (1) вместо x выражения $\bar{A}x + B\bar{x}$ мы получаем $AB = 0$ (т. е. (11) есть решение уравнения (1) в смысле Шредера).

Как уже отмечалось, точное решение в смысле Порецкого уравнения (1) относительно x

$$\begin{cases} x = \bar{A}x \\ x = x + \bar{A}B \end{cases} \tag{12}$$

эквивалентно равенству $x = \bar{A}(u + B)$ при $u = x$ в решении Шредера и получается из этого решения, если отбросить резольвенту $AB = 0$. Диаграмма точного решения (12) построена на рис. 6.

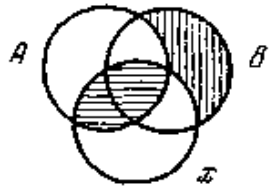


Рис. 6

По этой диаграмме в силу (10) мы получаем

$$x = \bar{A}x + \bar{A}B\bar{x},$$

т. е.

$$x = \bar{A}x + \bar{A}B \tag{13}$$

(то самое, что получается и из решения Шредера). Точное решение равенства вида (1) относительно x характеризуется следующими свойствами:

1. Оно есть следствие из информации, выражаемой равенством вида (1) (на языке диаграмм Венна: заштрихованные ячейки диаграммы точного решения не принадлежат к числу незаштрихованных ячеек диаграммы исходного равенства).
2. Если к точному решению (на диаграмме) присоединить резольвенту (равенство вида $AB = 0$), то получим диаграмму, совпадающую с диаграммой исходного равенства.
3. Никакое логическое следствие равенства (1), не зависящее от x, не может логически следовать из точного решения уравнения (1) относительно x (см. диаграммы самого сильного логического следствия, не содержащего x, и точного решения — рис. 4 и 6).

4. Все логические следствия уравнения (1), в которые входит *x,* но не входит $\bar{x}$, являются логическими следствиями точного решения (относительно *x)* равенства (1). В этом смысле точное решение относительно *x* (12) можно называть «самым сильным» логическим следствием, содержащим *x,* из уравнения (1).

5. Если по диаграмме точного решения в силу (10) написать эквивалентное ему выражение вида $x = f(x, a, \ldots, u)$, то результат подстановки в равенство (1) вместо *x* выражения *f (x, a, ..., и)* будет эквивалентен при условии $AB=0$ тождеству 0 = 0. [При одновременном рассмотрении нескольких диаграмм Венна мы предполагаем, если нет специальной оговорки, что все они построены из одних и тех же переменных.]

Нетрудно проверить, что диаграмма, изображенная на рис. 7, представляет собой точное решение (в смысле Порецкого) уравнения (1) относительно $\bar{x}$ — самое сильное логическое следствие, содержащее $\bar{x}$ (и не содержащее *x),* из уравнения (1).

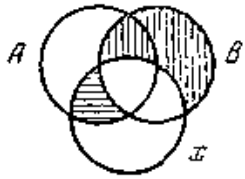


Рис. 7

По диаграмме (рис. 7) находится пара равенств относительно *x:*

$$\begin{cases} \bar{x} = \bar{x} + A\bar{B}, \\ \bar{x} = \bar{B}\bar{x}, \end{cases} \tag{14}$$

или эквивалентная пара равенств (относительно *x):*

$$\begin{cases} x = (\bar{A} + B)\,x, \\ x = B + x. \end{cases} \tag{15}$$

В силу (10) каждая из пар (14), (15) эквивалентна

$$x = \bar{A}x + B$$

или

$$\bar{x} = \bar{B}\bar{x} + A\bar{B}.$$

Следует обратить впимание также на диаграмму, помещенную на рис. 8.

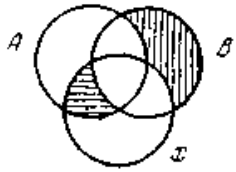


Рис. 8

Этой диаграмме соответствует пара равенств

$$\begin{cases} x = Bx + \bar{A}x, \\ x = x + \tilde{A}B, \end{cases} \tag{16}$$

или

$$\begin{aligned} \tilde{x} &= \tilde{x} + A\tilde{B}, \\ \tilde{x} &= (A + \tilde{B})\,\tilde{x}. \end{aligned} \tag{17}$$

Из диаграммы (рис.8) видно, что равенства, соответствующие этой диаграмме, удовлетворяют приведенным выше условиям 1—3, 5.

В силу (10) решение, представленное этой диаграммой, можно переписать в виде

$$x=(B+\bar{A})x+\bar{A}B$$

или

$$x=(A+\bar{B})\bar{x}+A\bar{B}$$

Кроме того, можно заметить, что все логические следствия уравнения (1), не вытекающие из резольвенты $AB = 0$, являются логическими следствиями решения, представленного диаграммой, изображенной на рис. 8.

Непосредственно из диаграмм (рис. 3, 6—8) следует также, что полное решение (рис. 3) уравнения (1) относительно x эквивалентно каждому из трех рассмотренных решений (см. диаграммы на рис. 6—8) только в случае, когда $AB = 0$.

Так, в задаче Венна уравнение посылок имеет следующий вид относительно b:

$$(\tilde{a} + c)\,b + a\tilde{c}\tilde{b} = 0,$$

и

$$(\tilde{a} + c)\,a\tilde{c} = 0.$$

Поэтому решение $b = a\tilde{c}$ является не только полным, но и точным относительно как b, так и $\bar{b}$ в смысле Порецкого.

Ранее разбиралась задача Порецкого. При этом остался невыясненным вопрос об отношениях между различными группами девиц,

присутствующих на балу. Найдем сначала полное определение класса благовоспитанных девиц, т. е. класса *a.* Из диаграммы Венна, соответствующей этой задаче (рис. 41), мы получаем: $a = a\bar{b}c\bar{d}$, $\bar{a} = \bar{a}bc\bar{d}$ (то же самое можно получить в силу (14)), или

$$a = a\bar{b}c\bar{d} \quad a = a + \bar{b} + c + \bar{d}. \tag{18}$$

Так как равенство $A = AA_1 \ldots A_k$ эквивалентно системе $A = AA_i$ (i =1, …, k), а равенство $A = A + A_1 + \ldots + A_k$ эквивалентно системе $A = A + A_i$ (i = 1, . . . , k), то равенства (18) эквивалентны системе равенств

$$a = a\bar{b}, \quad a = ac, \quad a = a\bar{d}, \tag{19}$$

$$a = a + \bar{b}, \quad a = a + c, \quad a = a + \bar{d}.$$

Следовательно, полное определение a из условия задачи можно представить в виде шести равенств (19). Первое и четвертое равенства из (19) показывают, что a эквивалентно (тождественно) $\bar{b}$, $a = \bar{b}$. Второе и пятое,— что a = с. Третье и шестое из них говорят, что $a = \bar{d}$. Итак, получаем $a = \bar{b} = c = \bar{d}$, $\bar{a} = b = \bar{c} = d$.

Таким образом, благовоспитанные, молодые, невеселые и некрасивые девицы образуют один класс; верно, что и девицы неблаговоспитанные, немолодые, веселые и красивые были одни и те же.

К тому же самому мы придем, если станем искать полное определение любого из остальных признаков: *b, c* или *d*. Однако точное решение относительно некоторого класса не дает ответа на поставленный вопрос. Например, $a = \bar{b}c\bar{d}$ — точное решение относительно класса *a.* Из него нельзя прямо заключить, что $a = \bar{b} = c = \bar{d}$.

ПРИЛОЖЕНИЕ А

СИНТАКСИЧЕСКАЯ И СЕМАНТИЧЕСКАЯ ФОРМУЛИРОВКИ ТЕОРЕМЫ О НЕПОЛНОТЕ

1. Постановка задачи. Доказанную нами формулировку теоремы Гёделя естественно называть семантической, так как в ней шла речь об истинности суждений арифметики. Вообще, семантикой называется та часть науки о языке (языке арифметики в нашем случае), которая интересуется смыслом выражений, их истинностью и ложностью, — в отличие от синтаксиса, который изучает выражения языка как комбинации знаков, в отрыве от их смысла (иногда слово «синтаксис» употребляют в более узком смысле, обозначая им часть грамматики, изучающую сочетания слов в предложениях естественного языка.). Мы хотим перейти к синтаксической

формулировке теоремы о неполноте, т. е. устранить по возможности упоминания об истинности суждений.

Полностью удовлетворительное решение этой задачи требует конкретизации понятия доказательства, выходящей за рамки этого приложения; тем не менее мы сделаем некоторые шаги в этом направлении.

2. Синтаксическая непротиворечивость и синтаксическая полнота. Пусть $\langle Д, D, \delta \rangle$ — дедуктика над алфавитом А языка арифметики. (В этом приложении мы будем рассматривать только дедуктики над А, не оговаривая этого специально.) Назовем ее *синтаксически непротиворечивой,* если не существует такого суждения α, для которого и α, и $\neg\alpha$ доказуемы в этой дедуктике. Назовем ее *синтаксической полной,* если для всякого суждения α хотя бы одно из суждений α и $\neg\alpha$ доказуемо в этой дедуктике. Эти определения можно сформулировать короче, введя предварительно понятие суждения, опровержимого в данной дедуктике, — такого суждения α, что суждение $\neg\alpha$ доказуемо в ней. Теперь можно сказать так: **дедуктика синтаксически непротиворечива, если никакое суждение не является доказуемым и опровержимым одновременно, и синтаксически полна, если всякое суждение либо доказуемо, либо опровержимо**.

Следующая лемма устанавливает связь между этими понятиями и понятиями непротиворечивой и полной (относительно $\langle A, T \rangle$) дедуктики. Напомним, что дедуктика называется непротиворечивой, если все доказуемые суждения истинны, и полной, если все истинные суждения доказуемы.

Л е м м а А.1. А) *Непротиворечивая дедуктика синтаксически непротиворечива.*

Б) *Полная дедуктика синтаксически полна.*

В) *Если дедуктика непротиворечива, то полнота ее равносильна синтаксической полноте.*

Доказательство. А) Если α и $\neg\alpha$ доказуемы в непротиворечивой дедуктике, то α и $\neg\alpha$ истинны, что противоречит определению истинности.

Б) Одно из суждений α и $\neg\alpha$ должно быть истинным, а следовательно, и доказуемым, если дедуктика полна.

В) Если α — истинное суждение, то $\neg\alpha$—ложное, поэтому $\neg\alpha$ не может быть доказуемо в непротиворечивой дедуктике и — если дедуктика синтаксически полна — α должно быть доказуемым.

Учитывая лемму, естественно предложить в качестве синтаксического варианта теоремы о неполноте такое утверждение:

не существует синтаксически непротиворечивой и синтаксически полной дедуктики для языка арифметики.

Этот вариант хорош тем, что, во-первых, из него вытекает доказанный нами семантический вариант теоремы о неполноте и, во-вторых, тем, что в нем совсем ничего не говорится об истинности. Однако так сформулированное утверждение неверно — дедуктика, в которой доказуемы те и только те суждения, в которые четное число раз входит символ ¬ (такая существует в силу теоремы 1), будет синтаксически непротиворечива и синтаксически полна. Поразмыслив о постигшей нас неудаче, мы приходим к выводу, что причина ее как раз в том, что рассмотренная формулировка никак не связана с обычным пониманием знаков алфавита А — в построенной дедуктике одновременно доказуемы, например, формулы (2·2) = 4 и (2·2) = 5. Мы выйдем из создавшегося положения, потребовав от дедуктики, чтобы некоторые суждения обязательно были доказуемыми в ней. Уточним сказанное.

Пусть D_0, D — некоторые дедуктики. Будем говорить, что D является *расширением* D_0, если всякое суждение, доказуемое в D_0, доказуемо и в D. (В этом случае, очевидно, всякое опровержимое в D_0 суждение опровержимо и в D.) Будем говорить, что дедуктика D_0 *пополнима,* если существует ее *пополнение,* т. е. расширение, являющееся синтаксически непротиворечивой и синтаксически полной дедуктикой.

Приведенный выше пример устанавливает пополнимость пустой дедуктики — дедуктики, в которой ни одно утверждение не доказумо. Используя понятие пополнимости, мы можем предложить в качестве синтаксического варианта теоремы Гёделя такое утверждение:

существует непополнимая дедуктика.

Однако это утверждение бессодержательно, так как всякая синтаксически противоречивая дедуктика непополнима. Кроме того, нам хотелось бы, чтобы из синтаксического варианта теоремы о неполноте вытекал доказанный нами семантический вариант. Мы удовлетворим этому требованию, выбрав такую формулировку:

существует непротиворечивая непополнимая дедуктика.

(Отсюда следует несуществование полной непротиворечивой дедуктики, так как такая дедуктика являлась бы пополнением любой непротиворечивой.) На этой формулировке мы и остановимся. Но прежде чем доказывать сформулированное утверждение, объясним, чем оно лучше исходной семантической формулировки,— ведь в нем мы говорим о непротиворечивости, определение которой апеллирует к истинности. Дело в том, что непротиворечивую непополнимую дедуктику можно указать явно и утверждение о непополнимости этой явно заданной дедуктики уже никак не апеллирует к понятию истинности. (Конечно, ценность этого утверждения в наших глазах

определяется нашей верой в непротиворечивость этой дедуктики.) Перейдем теперь к доказательству сформулированного утверждения. Для этого нам понадобятся некоторые новые понятия из теории алгоритмов.

3. Неотделимые множества. Пусть К — алфавит, *A* и *B* — непересекающиеся подмножества K^∞. Будем говорить, что множество *C отделяет A* от *B,* если $A \subset C$ и $B \cap C = \varnothing$. Если множество *C* отделяет *A* от *B,* то его дополнение (до K^∞) отделяет *B* от *A.* Будем говорить, что *A* и *B отделимы,* если существует разрешимое подмножество *C* множества K^∞, отделяющее *A* от *B.* (В этом случае дополнение *C* является разрешимым подмножеством K^∞, отделяющим *B* от *A.)*

Л е м м а А.2. *Непересекающиеся множества A и B отделимы тогда и только тогда, когда функция из* K^∞ *в* $\mathbb{N}$, *определенная соотношением*

$$f(x) = \begin{cases} 1, & \text{если } x \in A, \\ 0, & \text{если } x \in B, \\ \text{не определена}, & \text{если } x \notin A \cup B, \end{cases}$$

имеет всюду определенное вычислимое продолжение.

Доказательство. Если *g* — всюду определенное вычислимое продолжение *f*, то пазрешимое множество $\{x \mid g(x) = 1\}$ отделяет *A* от *B.* Наоборот, если разрешимое множество *C* отделяет *A* от *B,* то вычислимая функция *g,* равная 1 на элементах *C* и 0 вне *C,* продолжает *f.*

Лемма А.3. *Существуют перечислимые неотделимые подмножества* N.

Доказательство. Согласно предыдущей лемме достаточно доказать, что существует вычислимая функция *h* из $\mathbb{N}$ в $\mathbb{N}$, принимающая лишь два значения — 0 и 1 — и не имеющая всюду определенного вычислимого продолжения. В этом случае перечислимые (согласно лемме 6 и следствию 1 аксиомы протокола) множества

$$\{x \mid h(x) = 1\}$$

и

$$\{x \mid h(x) = 0\}$$

будут неотделимы. Чтобы построить функцию *h* с указанными свойствами, рассмотрим (следуя доказательству следствия 4 аксиомы программы в п1.5.5) функцию *d,* от которой никакая вычислимая функция не может отличаться всюду. Функцию *h* определим так:

$$h(x)=\begin{cases}1, & \text{если } d(x)=0,\\ 0, & \text{если } d(x) \text{ определено и не равно 0,}\\ \text{не определено,} & \text{если } d(x) \text{ не определено.}\end{cases}$$

Всякое всюду определенное продолжение функции *h* всюду отличается от *d*, поэтому не может быть вычислимым.

С помощью понятия неотделимости мы сформулируем признак непополнимости дедуктики.

Теорема А.1. *Если множества доказуемых и опровержимых в данной дедуктике суждений неотделимы, то эта дедуктика непополнима.*

Доказательство. Если эта дедуктика имеет пополнение, то множества доказуемых и опровержимых в этом пополнении суждений — непересекающиеся перечислимые множества, дающие в объединении множество всех суждений. В силу леммы 3 каждое из них, в частности, множество *S* доказуемых суждений, является разрешимым подмножеством множества всех суждений и, следовательно, разрешимым подмножеством множества A^{∞}. Множество *S* отделяет множество доказуемых в исходной дедуктике суждений от множества суждений, опровержимых в ней, что противоречит предположению.

4. Построение непополнимой дедуктики. Мы построим непополнимую дедуктпку, применяя теорему А.1. Пусть *P* и *Q* — перечнелимые неотделимые подмножества N (их существование установлено в лемме А.3). Множество *P* арифметично по аксиоме арифметичности (п.1.5.5); пусть α — формула, с которой оно сопряжено. Обозначим через $[n \in P]$ формулу $S_n^{x_1}\alpha$ (*n* — цифра); формула $[n \in P]$ истинна тогда и только тогда, когда $n \in P$. Для каждого *n* из *P* рассмотрим (истинную) формулу $[n \in P]$; для каждого *n* из *Q* рассмотрим (также истинную) формулу $\neg[n \in P]$. Рассмотренные формулы образуют перечислимое множество. Согласно теореме 1 существует дедуктика, в которой доказуемы эти формулы и только они. Эта дедуктика непротиворечива. Докажем, что она непополнима. Согласно теореме А.1 для этого достаточно доказать, что множества доказуемых и опровержимых в ней формул неотделимы; покажем это. Если $n \in P$, то формула $[n \in P]$ доказуема, если $n \in Q$, то формула $[n \in P]$ опровержима. Поэтому, если бы разрешимое множество S отделяло доказуемые формулы от опровержимых, то разрешимое множество $\{n \mid [n \in P] \in S\}$ отделяло бы *P* от Q, что невозможно. Итак, непополнимая дедуктнка построена.

Приложение Б.

АРИФМЕТИЧЕСКИЕ МНОЖЕСТВА И ТЕОРЕМА ТАРСКОГО О НЕАРИФМЕТИЧНОСТИ МНОЖЕСТВА ИСТИННЫХ ФОРМУЛ ЯЗЫКА АРИФМЕТИКИ

Как объяснялось в п.5.4, суждения языка арифметики являются высказываниями о свойствах натурального ряда и операций сложения и умножения. Они бывают истинными и ложными. Для формулы с параметрами вопрос «истинна она или ложна?» лишен смысла. Если мы вместо параметров формулы подставим цифры, то получим суждение, истинность которого зависит от того, какие именно цифры мы подставили. Таким образом, **формулы с параметром можно интерпретировать как свойства натуральных чисел**.

Пример 1. Результат подстановки n вместо x_1 в формулу $\exists x_2((x_2 + x_2) = x_1)$ является истинным суждением тогда и только тогда, когда n четно. Поэтому можно сказать, что эта формула выражает свойство «x_1 четно». Говорят также (не вполне корректно), что эта формула истинна при четных значениях x_1 и ложна при нечетных значениях x_1.

П р и м е р 2. Формула

$$\exists x_3((x_1 + x_3) = x_2)$$

выражает свойство «$x_1 \leqslant x_2$».

Пример 3. Формула

$$\exists x_2((x_1 \cdot x_2) = x_3)$$

выражает свойство «x_1 делит x_3».

Пример 4. Обозначим формулу из примера 3 через [x_1 делит x_3]. Тогда формула

$$\forall x_1([x_1 \text{ делит } x_3] \to ((x_1 = 1) \vee (x_1 = x_3)))$$

выражает свойство «x_3 — простое или x_3= 1».

Пример 5. Обозначим формулу из примерз 1 через [x_1 четно]. Тогда формула

$$\forall x_1([x_1 \text{ делит } x_3] \to ([x_1 \text{ четно}] \vee (x_1 = 1)))$$

выражает свойство «всякий делитель x_3 или четен, или равен 1», т. е. свойство «x_3 есть степень числа 2».

Свойства, выражаемые формулами языка арифметики, назовем арифметическими. Отождествляя свойство с множеством удовлетворяющих ему объектов, приходим к определению арифметического подмножества $\mathbb{N}^k$, частным случаем которого (при

k = 1) будет данное в п. 5.4 определение арифметических подмножеств $\mathbb{N}$.

Дадим точные» определения. Пусть α — формула языка арифметики, w_1, ..., w_p — переменные, c_1 ... ..., c_p — цифры. *Результатом подстановки* c_1, ..., c_p *вместо* w_1, ..., w_p *в* α назовем формулу

$$S_{c_1 \dots c_p}^{w_1 \dots w_p}\alpha = S_{c_p}^{w_p} \dots S_{c_2}^{w_2} S_{c_1}^{w_1}\alpha,$$

получающуюся из α последовательной подстановкой c_1, ..., c_p вместо w_1, ..., w_p. (Нетрудно понять, что результат последовательного выполнения нескольких подстановок не зависит от порядка, так что можно было бы, например, определить $S_{c_1 \dots c_p}^{w_1 \dots w_p}\alpha$ как $S_{c_1}^{w_1} \dots S_{c_p}^{w_p}\alpha$ — получилось бы то же самое.) Пусть α — формула арифметики, не имеющая параметров, отличных от x_1, ..., x_k. Рассмотрим подмножество $\mathbb{N}^k$, состоящее из тех $\langle c_1, \dots, c_k\rangle$, для которых суждение

$$S_{c_1 \dots c_k}^{x_1 \dots x_k}\alpha$$

истинно. Будем говорить, что оно *сопряжено* с формулой α. Множества, сопряженные с формулами языка арифметики, будем называть *арифметическими.* При $k = 1$ мы приходим к (данному в п. 5.4) определению арифметических подмножеств натурального ряда. Используя упоминавшееся отождествление свойств с множествами удовлетворяющих им объектов, мы будем говорить также об арифметичнести свойств натуральных чисел.

Пример 6. Множества

$$\{\langle x_1, x_2, x_3\rangle \mid x_1 + x_2 = x_3\},$$

$$\{\langle x_1, x_2\rangle \mid x_1 = x_2\},\ \{\langle x_1, x_2, x_3\rangle \mid x_1 \cdot x_2 = x_3\}$$

являются арифметическими, так как сопряжены с формулами

$$x_1 = x_2,\quad (x_1 + x_2) = x_3,\quad (x_1 \cdot x_2) = x_3.$$

Пример 7. Множество $\{\langle x_1, x_2\rangle \mid x_1 \leqslant x_2\}$ сопряжено с формулой примера 2 и потому является арифметическим.

Пример 8. Множество $\{\langle x_1, x_2\rangle \mid x_1 \text{ делит } x_2\}$ является арифметическим. Для того чтобы построить формулу, с которой оно сопряжено, нужно слегка переделать формулу из примера 3, заменив в ней x_2 на x_3 и наоборот.

Пример 9. Множество простых чисел и множество степеней числа 2 — арифметические подмножества натурального ряда. (См. примеры 4 и 5.)

Свойства арифметических подмножеств $\mathbb{N}$, указанные в п. 5.4, остаются верными и для арифметических подмножеств $\mathbb{N}^k$. В частности, верна следующая

Лемма Б.1. а) *Дополнение к арифметическому подмножеству* $\mathbb{N}^k$ $(до\ \mathbb{N}^k)$ *арифметично;*

б) *пересечение и объединение арифметических подмножеств* $\mathbb{N}^k$ *арифметичны.*

Следующая лемма показывает, что арифметичность сохраняется при перестановке координат.

Лемма Б.2. *Пусть* σ *— перестановка множества* $\{1, \ldots, k\}$ *(т. е. взаимно однозначное отображение его на себя), M — арифметическое подмножество* $\mathbb{N}^k$. *Тогда множество*

$$M^\sigma = \{\langle x_1, \ldots, x_k\rangle \mid \langle x_{\sigma(1)}, \ldots, x_{\sigma(k)}\rangle \in M\}$$

арифметично.

Доказательство. Если множество *M* сопряжено с формулой α, то множество M^σ сопряжено с формулой α^σ, которая получится, если в формуле α всюду заменить все переменные из списка $x_1, \ldots x_k$ на соответствующие переменные из списка $x_{\sigma(1)} \ldots\ldots x_{\sigma(k)}$.

Следующие леммы связывают классы арифметических подмножеств $\mathbb{N}^k$ при различных *k.*

Лемма Б.3. *Если M — арифметическое подмножество* $\mathbb{N}^k$, *то множество* $M \times \mathbb{N}^l$ *—арифметическое подмножество* $\mathbb{N}^{k+l}$.

Доказательство. В самом деле, $M \times \mathbb{N}^l$ сопряжено с той же формулой, что и *M.*

Лемма Б.4. *Если множество* $M \subset \mathbb{N}^{k+l}$ *арифметично, то его проекция M' на первые k осей, равная*

$$\{\langle x_1, \ldots, x_k\rangle \mid \exists x_{k+1} \ldots \exists x_{k+l} (\langle x_1, \ldots, x_{k+l}\rangle \in M)\},$$

является арифметическим подмножеством $\mathbb{N}^k$.

Доказательство. В самом деле, если *M* сопряжено с формулой α, то *M'* сопряжено с формулой

$$\exists x_{k+1} \ldots \exists x_{k+l} \alpha.$$

Сочетая леммы Б.2 и Б.4, можно доказать арифметичность проекции арифметического множества на любые оси.

Пусть $M \subset \mathbb{N}^2$ — арифметическое множество. Для каждого $n \in \mathbb{N}$ рассмотрим M_n — «*n*-е сечение множества М», множество тех *x,* для которых $\langle n, x\rangle \in M$. Будучи проекцией» множества $(\{n\} \times \mathbb{N}) \cap M$, оно арифметично. Назовем множество $\dot{M} \subset \mathbb{N}^2$ *универсальным* арифметическим множеством, если любое

арифметическое подмножество N является его сечением. Оказывается, такого быть не может.

Теорема Б.1. *Универсальных арифметических множеств не существует: каково бы ни было арифметическое множество* $M \subset \mathbb{N}^2$, *существует арифметическое множество* $Q \subset \mathbb{N}$, *которое отлично от всех сечений множества M.*

Доказательство. Множество $Q = \{x \mid \langle x, x\rangle \notin M\}$ арифметично, так как является проекцией множества $(\mathbb{N}^2 \setminus M) \cap \{\langle x, y\rangle \mid x = y\}$. Оно не может быть сечением *M:* если бы Q равнялось M_n, то по определению M_n мы имели бы $n \in Q \Leftrightarrow \langle n, n\rangle \in M$, но по определению *Q* имеет место соотношение

$$n \in Q \Leftrightarrow \langle n, n\rangle \notin M.$$

(Другими словами, множества Q и M_q по-разному ведут себя по отношению к числу *n,* поэтому не могут совпадать.)

Назовем функцию *f* из $\mathbb{N}^k$ в $\mathbb{N}^l$ *арифметической,* если ее график— арифметическое подмножество $\mathbb{N}^{k+l}$.

Лемма Б.5. *Образы и прообразы арифметических множеств при арифметических функциях арифметичны.*

Доказательство. Рассмотрим, например, образ арифметического множества $A \subset \mathbb{N}$ при арифметической функции *f* из $\mathbb{N}$ в $\mathbb{N}$. Этот образ есть проекция множества $(\text{график } f) \cap (A \times \mathbb{N})$ и поэтому арифметичен. Другими словами, если через $[f(x_1)=x_2]$ обозначить формулу, с которой сопряжен график *f,* а через $[x_1 \in A]$ — формулу, с которой сопряжено *A,* то формула

$$\exists x_1 ([f(x_1) = x_2] \wedge [x_1 \in A])$$

будет истинна для тех и только тех значений x_2, которые принадлежат образу *A.* Чтобы указать формулу, сопряженную с образом *A,* достаточно переименовать переменные (поменять всюду x_1 и x_2). Прообраз множества *A* при функции *f* будет сопряжен с формулой

$$\exists x_2 ([f(x_1) = x_2] \wedge [x_2 \in A]),$$

где $[x_2 \in A]$ обозначает формулу, получающуюся из $[x_1 \in A]$ переименованием переменных.

Мы стремимся доказать теорему Тарского, утверждающую, что *множество истинных формул арифметики неарифметично.*

Чтобы придать смысл этой формулировке, надо объяснить, что мы имеем в виду, говоря об арифметичности множества формул — некоторого подмножества A^∞. Можно определить это понятие так: выбрать взаимно однозначное соответствие между A^∞ и $\mathbb{N}$ *(нумерацию* A^∞), сопоставляющее каждому слову *X* из A^∞ некоторое

натуральное число (номер слова *X* при этой нумерации), и называть множество $M \subset \mathrm{A}^{\infty}$ арифметическим (относительно данной нумерации), если множество номеров слов из *M* является арифметическим подмножеством $\mathbb{N}$.

Конечно, это определение зависит от выбора нумерации слов алфавита А. Назовем две нумерации *арифметически эквивалентными,* если функции, дающие по номеру слова в одной из них номер того же слова в другой, арифметичны.

Лемма Б.6. *Если множество* $M \subset \mathrm{A}^{\infty}$ *арифметично относительно данной нумерации* π_1*, то оно арифметично относительно любой нумерации* π_2*, арифметически эквивалентной* π_1.

Доказательство. По условию множество $\pi_1(M)$ (состоящее из π_1- номеров слов из *M)* арифметично. Множество $\pi_2(M)$ является образом $\pi_1(M)$ при арифметической (по условию) функции, дающей π_2-номера по π_1-номерам, поэтому оно также арифметично.

Окончательное определение арифметических множеств слов в алфавите А таково: *арифметическими* называются множества, арифметические относительно вычислимых нумераций множества A^{∞}. (Нумерация называется *вычислимой,* если функция, сопоставляющая слову его номер, вычислима. В этом случае вычислима и обратная функция, сопоставляющая числу *n* слово с номером *n;* существование вычислимых нумераций A^{∞} установлено в примере 3 из п. 5.2.)

Чтобы доказать корректность этого определения, достаточно показать, что все вычислимые нумерации арифметически эквивалентны. Если π_1 и π_2 — вычислимые нумерации, то функция, сопоставляющая номеру слова относительно π_1 номер того же слова относительно π_2, вычислима. (Она вычисляется следующим алгоритмом; получив аргумент *x,* перебирай все слова алфавита А, вычисляй их π_1-номера и жди появления слова, у которого π_1-номер равен *x;* найдя это слово, вычисли его π_2-номер.) Поэтому корректность будет доказана, если мы установим, что верна следующая лемма.

Лемма Б.7. *Всякая вычислимая функция из* $\mathbb{N}^p$ *в* $\mathbb{N}^q$ *арифметична.*

Доказательство. График вычислимой функции из $\mathbb{N}^p$ в $\mathbb{N}^q$ есть перечислимое подмножество $\mathbb{N}^{p+q}$ (следствие 3 аксиомы протокола), поэтому требуемое утверждение вытекает из следующей усиленной формы аксиомы арифметичности:

всякое перечислимое подмножество $\mathbb{N}^k$ *арифметично.*

(В п. 5.5. аксиомой арифметнчности был назван частный случай этого утверждения, возникающий при $k = 1$.)

Теорема Б.2. *Множество Т истинных формул арифметики неарифметично.*

Доказательство. Покажем, что если бы T было арифметично, то в противоречии с теоремой Б.1 существовало бы универсальное арифметическое множество. Следуя Гёделю, назовем формулы, не имеющие отличных от x_1 параметров, *классовыми.* Множество всех классовых формул — разрешимое подмножество перечислимого множества A^{∞} и потому перечислимо. Зафиксируем какое-нибудь вычислимое перечисление $\alpha_0, \alpha_1, \alpha_2, \ldots$ множества классовых формул. Рассмотрим множество

$M :=$ {$\langle n, m\rangle$ | результат подстановки m вместо
x_1 в a_n — истинное суждение}.

Очевидно, n-е сечение этого множества сопряжено с формулой α_n, а потому сечениями этого множества являются все арифметические подмножества $\mathbb{N}$. Остается показать, что если бы T было арифметично, то и M было бы арифметично. Вспоминая определение арифметичностц множества слов, зафиксируем произвольную вычислимую нумерацию слов алфавита A^{α}. Пусть T' — множество номеров слов из T при этой нумерации. Функция S, сопоставляющая паре $\langle n, m\rangle$ номер слова, являющегося результатом подстановки m вместо x_1 в α_n, вычислима и, следовательно, арифметична (лемма Б.7). Множество M является прообразом множества T при функции S. Поэтому из арифметичности T вытекает арифметичность M (лемма Б.5). Теорема Б.2 доказана.

Анализ доказательства теоремы Б.2 показывает, что оно связано с «парадоксом лжеца». Коротко скажем об этой связи.

Парадокс лжеца состоит в следующем. Некто заявляет: «То, что я сейчас говорю, ложно». Истинно или ложно его высказывание? Любой из ответов ведет к противоречию. Если предположить, что оно истинно, то в силу своего собственного смысла оно должно быть ложным и наоборот. Изложим теперь рассмотренное доказательство теоремы Б.2 в форме, близкой к этому парадоксу.

Пусть множество номеров истинных суждений арифметики арифметично; обозначим через [слово с номером x_3 истинно] формулу, единственным параметром которой является x_3 и которая выражает свойство «слово с номером x_3 принадлежит T», т. е. свойство «$x_3 \in T'$». Функция S арифметична; обозначим через

[x_3 есть номер результата подстановки x_2
в x_1-ю классовую формулу]

формулу, с которой сопряжен график функции S. Формула

$\exists x_3$ ([слово с номером x_3 истинно] $\wedge$
$\wedge$ [x_3 есть номер результата подстановки x_2,
в x_1-ю классовую формулу])

имеет параметрами x_1 и x_2; с этой формулой сопряжено множество *M;* обозначим ее

[результат подстановки x_2
в x_1-ю классовую формулу истинен].

Дальше рассуждение следует доказательству теоремы о несуществовании универсального арифметического множества. Рассмотрим формулу

$\neg\exists x_2\{(x_1 = x_2) \wedge$ [результат подстановки x_2
в x_1-ю классовую формулу истинен]);

обозначим ее

[результат подстановки x_1
в x_1-ю -ю классовую формулу ложен];

она имеет параметром x_1 и отвечает своему обозначению в том смысле, что результат подстановки числа *n* вместо x^1 в эту формулу истинен тогда и только тогда, когда результат подстановки *n* в *n*-ю классовую формулу ложен. (Рассмотрение этой формулы соответствует рассмотрению множества Q в доказательстве теоремы Б.1.) Построенная формула является классовой и, следовательно, имеет некоторый номер (обозначим его *n)* в перечислении классовых формул. Подставим *n* вместо x_1 в построенную нами формулу, результат подстановки обозначим

[результат подстановки *n*
в *n*-ю классовую формулу ложен];

это — суждение, истинное тогда и только тогда, когда результат подстановки цифры *n* в *n*-ю классовую формулу ложен. Но этот результат представляет собой не что иное, как само рассматриваемое нами суждение. Мы получаем, что суждение

[результат подстановки *n*
в *n*-ю классовую формулу ложен]

истинно тогда и только тогда, когда ложно. (Это суждение можно было бы с полным основанием обозначить [Я ложно].) Полученное противоречие доказывает, что множество истинных формул арифметики неарифметично.

Приложение В.

ЯЗЫК АДРЕСНЫХ ПРОГРАММ, РАСШИРЕННЫЙ АРИФМЕТИЧЕСКИЙ ЯЗЫК И АКСИОМА АРИФМЕТИЧНОСТИ

В этом приложении мы постараемся обосновать аксиому арифметичности. План наших рассуждений таков. Сначала мы опишем некоторый конкретный класс алгоритмов — класс адресных программ; функции, вычисляемые с помощью алгоритмов из этого класса, естественно назвать адресно вычислимыми. Затем мы докажем, что область значений всякой адресно вычислимой функции представляет собой арифметическое множество. Тем самым аксиома арифметичности будет обоснована — если поверить в то, что всякая вычислимая функция является адресно вычислимой.

Вспомогательным средством для нас будет служить расширенный арифметический язык, который отличается от описанного в п.1.5.4 языка арифметики наличием некоторых дополнительных выразительных средств. Мы покажем, что множество значений всякой адресно вычислимой функции можег быть описано формулой расширенного арифметического языка. Затем мы покажем, что это расширение на самом деле несущественно и что для всякой формулы расширенною языка можно найти заменитель в обычном языке арифметики. Отсюда будет следовать, что множество значений любой адресно вычислимой функции может быть описано формулой языка арифметики, т. е. арифметично.

Начнем со следующего простого замечания: для обоснования аксиомы арифметичности и даже ее усиленного варианта, рассматриваемого и используемого в приложении Б, достаточно уметь доказывать, что

график всякой вычислимой функции из $\mathbb{N}$ *в* $\mathbb{N}$ *является арифметическим подмножеством* $\mathbb{N}^2$. (Понятие арифметического подмножества $\mathbb{N}^2$ введено в приложении Б.) В самом деле, пусть это утверждение верно. Тогда всякое перечислимое подмножество $\mathbb{N}$ арифметично, так как оно является проекцией графика перечисляющей его вычислимой функции (лемма Б.4). Докажем теперь усиленный вариант аксиомы арифметичности. Пусть $M \subset \mathbb{N}^k$ — перечислимое множество, g— перечисляющая его функция из $\mathbb{N}$ в $\mathbb{N}^k$. Значение функции g на числе n представляет собой кортеж из k чисел: $g(n) = \langle g_1(n), \ldots, g_k(n) \rangle$. Функции g_1, ..., g_k суть вычислимые функции из $\mathbb{N}$ в $\mathbb{N}$, поэтому в силу сделанного нами предположения их графики арифметичны. Обозначим через $[g_i(x_1) = x_2]$ формулы, с

которыми эти графики сопряжены. График g является арифметическим подмножеством $\mathbb{N}^{k+1}$, так как он сопряжен с формулой

$$[g_1(x_1) = x_2] \wedge ([g_2(x_1) = x_3] \wedge (\ldots \wedge [g_k(x_1) = x_{k+1}]) \ldots).$$

(Здесь через $[g_i(x_1) = x_{i+1}]$ обозначена формула, получаюшаяся из формулы $[g_i(x_1) = x_2]$ переименованием переменных, при котором переменные x_2 и x_{i+1} заменяются друг на друга.) Множество M является проекцией графика g на оси $x_2, \ldots, x_{k+1}$ и потому является арифметическим.

Итак, наша цель — доказать, что график всякой вычислимой функции из $\mathbb{N}$ в $\mathbb{N}$ является арифметическим подмножеством $\mathbb{N}^2$. Однако эта задача далеко не тривиальна — читатель может убедиться в этом, попробовав доказать, например, арифметичность показательной функции с основанием 2, т. е. арифметичность множества

$$\{\langle x, y\rangle \mid y = 2^x\}$$

В.1. Язык адресных программ. Мы опишем сейчас некоторый класс алгоритмов специального вида, которые будут называться адресными программами. Эти программы будут напоминать «программы в машинных кодах» для реально существующих ЭВМ.

Адресная программа представляет собой последовательность команд, пронумерованных по порядку. Каждая команда имеет один из следующих видов:

1° $R(a) \leftarrow b$ (присвоение значения);
2° $R(a) \leftarrow R(b)$ (пересылка);
3° $R(a) \leftarrow R(b) + R(c)$ (сложение);
4° $R(a) \leftarrow R(b) \cdot R(c)$ (умножение);
5° **ИДТИ К** n (безусловный переход);
6° **ЕСЛИ** $R(a) = R(b)$ **ТО ИДТИ К** m **ИНАЧЕ К** n (условный переход);
7° **СТОП** (останов).

Здесь a, b и c — произвольные натуральные числа («номера регистров»), m, n — натуральные числа, являющиеся порядковыми номерами некоторых команд программы. Последней командой программы должна быть команда вида 7°. В скобках указаны названия видов команд. Вот простой пример адресной программы:

Пример 1.

1 $R(1) \leftarrow 1$
2 $R(2) \leftarrow 1$
3 $R(3) \leftarrow 1$
4 $R(2) \leftarrow R(2) \cdot R(1)$
5 $R(1) \leftarrow R(1) + R(3)$
6 **ЕСЛИ** $R(1) = R(0)$ **ТО ИДТИ К 7 ИНАЧЕ К 4**
7 $R(0) \leftarrow R(2)$
8 **СТОП**

Адресные программы могут выполняться на (воображаемых) адресных машинах.

Адресная машина имеет бесконечное число устройств, предназначенных для хранения (запоминания) натуральных чисел. Эти устройства называются регистрами. В каждом регистре в каждый момент времени хранится (запоминается) ровно одно число. Регистры снабжаются номерами 0, 1, 2, ... и обозначаются соответственно $R(0)$, $R(1)$, $R(2)$ и т. д.

Адресная машина выполняет программу в порядке номеров команд; этот порядок нарушается лишь при выполнении команд условного и безусловного переходов. Прежде чем давать точные определения, опишем работу адресной машины по программе из примера 1. Пусть до начала работы в регистре *R(0)* находится число 100, в остальных — нули. Первые три команды задают начальные значения регистров *R*(1)— *R*(3).Содержимое регистра *R*(3) не меняется во время дальнейшего выполнения программы, содержимое регистра *R*(1) время от времени увеличивается на 1 (команда 5; напомним, что в *R(3)* всегда хранится 1), содержимое *R*(2) время от времени умножается на значение, хранящееся в *R(1).* Выполнение программы заканчивается, когда содержимое *R(1)* становится равным содержимому *R(0).* Изменение содержимого регистров с течением времени отражено в следующей таблице:

Номер команды	R (0)	R (1)	R (2)	R (3)	R (4)
1	100	0	0	0	0
2	100	1	0	0	0
3	100	1	1	0	0
4	100	1	1	1	0
5	100	1	1	1	0
6	100	2	1	1	0
4	100	2	1	1	0
5	100	2	2	1	0
6	100	3	2	1	0
4	100	3	2	1	0
5	100	3	6	1	0
6	100	4	6	1	0
4	100	4	6	1	0
. . . . .					
6	100	99	98!	1	0
4	100	99	98!	1	0
5	100	99	99!	1	0
6	100	100	99!	1	0
7	100	100	99!	1	0
8	99!	100	99!	1	0

В результате работы этой программы число 99! (=1·2· ... ·99) помещается в регистр *R(0).* Если вначале в *R{0)* хранилось не 100, а 200, то в регистр *R(0)* будет по окончании работы помещено число 199! (=1·2· ... ·199). Если же до начала работы программы во всех регистрах хранились нули, то выполнение программы никогда не закончится.

Дадим точные определения. *Состоянием* адресной машины называется бесконечная последовательность натуральных чисел $s = (s_0, s_1, \ldots)$, в которой почти все (все, кроме конечного числа) члены равны 0 (такие последовательности называются *финитными).* Если $s_0 = 0$, состояние называется *заключительным* (состоянием остановки); если $s_0 \geqslant 1$, состояние называется *рабочим,* а s_0 — *номером выполняемой команды.* Число s_{i+1} называется *содержимым i-го регистра.*

Пусть *p*— некоторая адресная программа, *s* = = (s_0, s_1, ...)—рабочее состояние. Будем говорить, что программа *p применима* к состоянию *s,* если s_0 — номер одной из команд программы *p* (команды с номером s_0 может не быть, если s_0 слишком велико). В этом случае мы определим некоторое состояние *s',* называемое *непосредственным результатом применения программы p* к состоянию *s.* Состояние $s' = (s'_0, s'_1, \ldots)$ определяется следующим образом:

1° если команда с номером s_0 имеет вид $R(a) \leftarrow b$, то $s'_0 = s_0 + 1$, $s'_{i+1} = s_{i+1}$ при $i \neq a$, $s'_{a+1} = b$' (содержимое всех

регистров, кроме *a*-го, не меняется, в *a*-й регистр помещается число *b;* машина переходит к выполнению следующей команды);

2° если команда с номером s_0 имеет вид $R(a) \leftarrow \leftarrow R(b)$, то $s_0' = s_0 + 1$, $s_{i+1}' = s_{i+1}$ при $i \neq a$, $s_{a+1}' = s_{b+1}$ (содержимое всех регистров, кроме *a*-го, не меняется; в *a*-й регистр помещается содержимое *b*-го регистра; машина переходит к выполнению следующей по порядку команды);

3° если команда с номером s_0 имеет вид $R(a) \leftarrow R(b) + R(c)$, то $s_0' = s_0 + 1$, $s_{i+1}' = s_{i+1}$ при $i \neq a$, $s_{a+1}' = s_{b+1} + s_{c+1}$ (содержимое всех регистров, кроме *a*-го, не меняется, в *a*-й регистр помещается сумма содержимого *b*-го и *c*-го регистров; машина переходит к выполнепию следующей команды);

4° если команда с номером s_0 имеет вид $R(a) \leftarrow R(b) \cdot R(c)$, то $s_0' = s_0 + 1$, $s_{i+1}' = s_{i+1}$ при $i \neq a$, $s_{a+1}' = s_{b+1} \cdot s_{c+1}$ (этот случай отличается от предыдущего лишь заменой сложения на умножение);

5° если команда имеет вид ИДТИ К *n,* то $s_0' = n$, $s_{i+1}' = s_{i+1}$ при всех *i* (содержимое регистров не меняется, машина переходит к выполнению команды номер *n);*

6° если команда с номером s_0 имеет вид ЕСЛИ *R(a)= R(b)* ТО ИДТИ К *m* ИНАЧЕ К *n*, то $s_{i+1}' = s_{i+1}$ при всех *i*, $s_0' = m$, если $s_{a+1} = s_{b+1}$; если же $s_{a+1} \neq s_{b+1}$, то $s'_0 = n$ (содержимое регистров не меняется, машина переходит к выполнению команды номер *m,* если содержимое *a*-го регистра равно содержимому *b*-го регистра, и к выполнению команды номер *n* в противном случае);

7° если команда имеет вид СТОП, то $s_{i+1}' = s_{i+1}$ при всех *i* и $s'_0 = n$ (машина переходит в заключительное состояние).

Определение непосредственного результата применения адресной программы к состоянию закончено. Заметим, что если программа *p* применима к состоянию *s,* то либо непосредственный результат применения является заключительным состоянием, либо к нему применима программа *p*.(мы предполагаем, что номера команд в операторах перехода в программе *p* являются номерами команд программы *p* и что последняя команда—-команда останова).

Протоколом применения адресной программы *p* называется последовательность состояний s^0, s^1, ..., s^k, в которой каждое следующее состояние является непосредственным результатом применения *p* к предыдущему, а последнее состояние является заключительным. Состояние s^0 называется *начальным* состоянием протокола. Существует не более одного протокола данной адресной программы *p* с данным начальным состоянием; такого протокола не существует, если *p* неприменима к s^0 или если в последовательности,

возникающей при многократном применении p к s^0, не встречается заключительное состояние.

Пусть p — адресная программа, k — натуральное число. Рассмотрим функцию f из $\mathbb{N}^k$ в $\mathbb{N}$, определяемую так: значение f на наборе $\langle a_1, \ldots, a_k \rangle$ есть b если существует протокол применения адресной программы p, начальным состоянием которого является $(1, a_1, a_2, \ldots, a_k, 0, 0, \ldots)$, а b есть содержимое 0-го регистра в заключительном состоянии этого протокола; другими словами, значение f на $\langle a_1, \ldots, a_k \rangle$ есть содержимое регистра $R(0)$ после выполнения программы, если перед ее выполнением числа $a_1, \ldots, a_k$ были помещены в регистры $R(0), \ldots, R(k-1)$, остальные регистры заполнены нулями и программа начала выполняться с первой команды. Функцию f мы будем называть функцией, *k-вычисляемой* программой p (или просто *вычисляемой* программой p, если значение k ясно из контекста).

Пример 2. Пусть p — адресная программа из примера 1. Функция f_1, 1-вычисляемая этой программой, такова: $f_1(0)$ не определено, $f_1(i) = (i-1)!$ при $i \geq 1$. Функция f_2, 2-вычнсляемая этой программой, такова:

$$f_2(i, j) = \begin{cases} (i-1)!, & \text{если } i \geqslant 1, \\ \text{не определена}, & \text{если } i = 0. \end{cases}$$

Функции, k-вычисляемые адресными программами, назовем *адресно вычислимыми* функциями k аргументов. Очевидно, что все адресно вычислимые функции вычислимы; с другой стороны, все вычислимые функции, известные в настоящее время, оказываются адресно вычислимыми. Таким образом, есть основания принять гипотезу о совпадении классов вычислимых функций из $\mathbb{N}^k$ в $\mathbb{N}$ и адресно вычислимых функций.

Приняв эту гипотезу, мы в следующих пунктах докажем утверждение аксиомы арифметичности.

В.2. Расширенный арифметический язык. Вспомогательным средством при доказательстве арифметичности адресно вычислимых функций будет служить расширенный арифметический язык. Чтобы определить его, следует внести в определение языка арифметики из п. 5.4 некоторые изменения.

В алфавит языка добавим два новых символа v (символ для образования одноместных функциональных переменных) и w (символ для образования двуместных функциональных переменных). Слова вида (v^n) будут называться *одноместными функциональными переменными,* а слова вида (w^n)*—двуместными функциональными переменными.* (Здесь $n \geq 1$.) Сокращенными обозначениями для

одноместных и двуместных функциональных переменных будут служить v_n и w_n. Подразумеваемыми значениями одноместных и двуместных функциональных переменных будут всюду определенные функции от одного и соответственно двух натуральных аргументов с натуральными значениями. Переменные x_n мы будем называть *числовыми* переменными.

Понятие *терма* расширенного арифметического языка определим так:

1° числовая переменная есть терм;

2° если t и u — термы, то $(t+u)$ и $(t\cdot u)$ — термы;

3° если p — одноместная функциональная переменная, а t — терм, то $p(t)$ —терм;

4° если r — двуместная функциональная переменная, а t и u — термы, то $r(t, u)$ — терм.

Пример 1. Слова $(v_4(x_1)+x_2)$, $v_4(v_5(w_2(x_1, x_2)))$, $w_2(w_2(x_1, x_4), x_7)$ являются термами расширенного арифметического языка.

Как и раньше, э*лементарной формулой* называется равенство двух термов; разумеется, имеются в виду термы расширенного арифметического языка. *Формулы* расширенного арифметического языка определяются, как в п. 5.4., при этом в 4° переменная ξ может быть числовой, одноместной функциональной или двуместной функциональной переменной.

Пример 2.

Слова $\forall v_1(v_1(x_1)=v_1(x_2))$,

$\forall x_1\forall x_2(v_1(x_1)=v_1(x_2))$,

$\forall w_1\forall x_1\forall x_2(w_1(x_1, x_1)=$

$=w_1(x_2, x_1))$ являются формулами.

Параметры термов и формул определяются, как в п. 5.4; в качестве параметров могут выступать как числовые, так и функциональные переменные.

Пример 3. Параметрами термов из примера 1 являются v_4, x_1 и x_2 (первый терм), v_4, v_5, w_2, x_1 и x_2 (второй терм), w_2, x_1, x_4 и x_7 (третий терм). Параметрами формул из примера 2 являются x_1 и x_2 (первая формула), v_1 (вторая формула); третья формула не имеет параметров.

Формулы, не имеющие параметров, называются *суждениями* расширенного арифметического языка.

Теперь мы должны определить, какие суждения расширенного арифметического языка мы объявляем истинными. У нас будет два варианта определения истинности для формул расширенного арифметического языка — две *интерпретации* этого языка. В одной из

них функциональные переменные будут принимать в качестве значений все (всюду определенные) функции от одного и от двух натуральных аргументов, в другой их значениями будут лишь финитные функции, т. е. функции, отличные от 0 лишь на конечном множестве аргументов. Чтобы не повторять определение дважды, мы будем говорить о классе *допустимых* функций, подразумевая под ним либо класс всех функций, либо класс всех финитных функций.

Определение истинности будет аналогично определению из п. 5.4. Новым для нас будет случай формул, начинающихся с кванторов по функциональным переменным. Здесь мы сталкиваемся со следующей проблемой: хотелось бы назвать, например, формулу вида $\forall \upsilon \alpha$ истинной, если для всех допустимых значений υ формула, получающаяся из α подстановкой этих значений вместо υ, является истинной. Но в нашем языке нет ничего, что можно было бы подставлять вместо функциональных переменных. Выход из этого положения таков: мы должны ввести в язык, помимо функциональных переменных, и функциональные константы — по одной для каждой допустимой функции.

Дадим точные определения. Выберем некоторый набор символов, находящийся во взаимно однозначном соответствии с множеством допустимых функций. Символы этого набора назовем *функциональными константами,* изображающими соответствующие им функции. Функциональные константы делятся на *одноместные* и *двуместные* в зависимости от числа аргументов у соответствующей функции. Мы будем подставлять функциональные константы вместо функциональных переменных с тем же числом аргументов. Результат подстановки произвольных функциональных констант вместо всех функциональных параметров и произвольных цифр вместо всех числовых параметров некоторого терма (формулы) расширенного арифметического языка назовем *оцененным термом (оцененной формулой).* Подстановка производится таким же образом, как в п. 5.4, — заменяются не все вхождения переменных, а лишь не попадающие в зону действия одноименного квантора. Частным случаем оцененного терма является постоянный терм, т. е. терм, не имеющий параметров (и, следовательно, являющийся термом языка арифметики). Частным случаем оцененной формулы является суждение расширенного арифметического языка.

Теперь мы можем дать определения *значений* оцененных термов и формул, вполне аналогичные определениям значений постоянных термов и замкнутых формул (суждений) языка арифметики. Значения оцененных термов определяются так:

1° значением цифры $(|^n)$ является число n;

2° значением оцененного терма *(t + u)* служит сумма значений оцененных термов *t* и *u,* а значением оцененного терма $(t \cdot u)$ служит произведение значений оцененных термов *t* и *u;*

3° значением оцененного терма вида $\gamma(t)$, где γ — одноместная функциональная константа, а *t* — оцененный терм, служит значение функции, изображаемой константой γ, на числе, равном значению оцененного терма *t;*

4° значением оцененного терма вида $\delta(t, u)$, где δ — двуместная функциональная константа, а *t* и *u* — оцененные термы, служит значение функции из $\mathbb{N}^2$ в $\mathbb{N}$, изображаемой константой δ, на паре ⟨значение *t,* значение *u*⟩.

Теперь для определения значений оцененных формул мы можем воспользоваться данным в п. 5.4 определением значений суждений языка арифметики, заменив слово «суждение» на «оцененная формула», «постоянный терм» на «оцененный терм», оговорив в 7° и 8°, что ξ является числовой переменной, и добавив два следующие пункта:

9° оцененная формула $\exists\xi\alpha$, где ξ — функциональная переменная, истинна, если существует такая функциональная константа γ с тем же числом аргументов, что у ξ, что оцененная формула $S^{\xi}_{\gamma}\alpha$ истинна; если такой константы нет, то оцененная формула $\exists\xi\alpha$ ложна;

10° оцененная формула $\forall\xi\alpha$, где ξ — функциональная переменная, истинна, если для всякой функциональной константы γ с тем же числом аргументов, что у ξ, оцененная формула $S^{\xi}_{\gamma}\alpha$ истинна; в противном случае оцененная формула $\forall\xi\alpha$ ложна.

Дав определения значений оцененных формул, мы, в частности, определили значения суждений расширенного арифметического языка. Этот частный случай будет для нас в дальнейшем особенно важен.

Пример 4. Суждение

$$\forall x_1 \forall x_2 (\forall v_1 (v_1(x_1) = v_1(x_2)) \rightarrow (x_1 = x_2)),$$

которое можно прочесть так: «если значения всех допустимых функций на x_1 и x_2 совпадают, то $x_1 = x_2$» истинно при любом из двух пониманий допустимости — считаем ли мы допустимыми все функции или только финитные функции.

Пример 5. Суждение

$$\forall v_1 \exists x_1 \forall x_2 (v((x_1 + x_2)) = 0),$$

которое можно прочесть так: «всякая допустимая функция равна 0 для всех достаточно больших значений аргумента», истинно, если допустимыми считать финитные функции, и ложно, если считать все функции допустимыми.

П р и м е р 6. Суждение

$$\forall w_1 \exists v_1 \forall x_1 (v_1(x_1) = w_1(x_1, x_1))$$

истинно при любом из двух пониманий допустимости.

Пример 7. Следующее суждение утверждает, что существует допустимое взаимно однозначное соответствие между $\mathbb{N}^2$ и $\mathbb{N}$:

$$\exists w_1 (\forall x_1 \exists x_2 \exists x_3 (w_1(x_2, x_3) = x_1) \wedge \wedge \forall x_2 \forall x_3 \forall x_4 \forall x_5 ((w_1(x_2, x_3) = = w_1(x_4, x_5)) \rightarrow ((x_2 = x_4) \wedge (x_3 = x_5)))).$$

Оно истинно, если считать допустимыми все функция, и ложно, если считать допустимыми лишь финитные функции.

Пример 8. Утверждение «при всех x_1 верно неравенство $2^{x_1} \geqslant x_1$» может быть переведено следующей формулой расширенного языка арифметики (при любом из двух пониманий допустимости):

$$\forall x_1 \forall v_1 (((v_1(0) = 1) \wedge \forall x_2 ([x_2 \leqslant x_1] \rightarrow \rightarrow (v_1((x_2 + 1)) = (2 \cdot v_1(x_2))))) \rightarrow [x_1 \leqslant v_1(x_1)]).$$

Здесь $[x_2 \leqslant x_1]$ обозначает $\exists x_3 ((x_2 + x_3) = x_1)$, а $[x_1 \leqslant v_1(x_1)]$ обозначает $\exists x_3 ((x_1 + x_3) = v_1(x_1))$. Суждение может быть прочитано так: если v_1 — последовательность натуральных чисел, первый член которой равен 1 и каждый следующий, вплоть до x_1+1-го, вдвое больше предыдущего, то $v_1(x_1) \geqslant x_1$. Оговорка «вплоть до x_1+1-го» необходима, если допустимыми являются лишь финитные функции.

Подобно тому как в приложении Б мы интерпретировали формулы языка арифметики с параметрами как выражающие свойства натуральных чисел, мы можем рассматривать формулы расширенного языка как выражающие свойства чисел и функций.

П р и м ер 9. Формула

$$\exists x_3 ((v_1(x_2) + x_3) = v_1(x_1))$$

выражает следующее свойство: значение допустимой функции v_1 на числе x_1 не меньше ее значения на числе x_2. Последнее предложение является (не вполне корректным, так как v_1 — функциональная переменная, а не функция, а x_1, x_2 — числовые переменные, но не числа) сокращением для следующего утверждения: результат подстановки вместо x_1 и x_2 некоторых цифр n_1 и n_2 и вместо v_1 некоторой функциональной константы, изображающей допустимую функцию, тогда и только тогда является истинной оцененной формулой расширенного языка арифметики (при любом из двух пониманий допустимости), когда значение этой допустимой функции на числе n_1 не меньше ее значения на числе n_2.

Пример 10. Формула

$$\forall x_1 \forall x_2 \exists x_3 ((v_1(x_1) + x_3) = v_1((x_1 + x_2)))$$

выражает свойство «допустимая функция v_1 является неубывающей».

Даже если ограничиться формулами расширенного языка, не имеющими функциональных параметров, мы все равно приобретаем новые возможности по сравнению с языком арифметики.

Пример 11. Формула

$$\exists v_1 (((v_1(0) = 1) \wedge \forall x_3 ([x_3 \leqslant x_1] \rightarrow \rightarrow (v_1((x_3 + 1)) = (2 \cdot v_1(x_3))))) \wedge (v_1(x_1) = x_2)),$$

где $[x_3 \leqslant x_1]$ является (уже привычным для нас) сокращением для $\exists x_2 ((x_3 + x_2) = x_1)$, выражает упоминавшееся вначале приложения В свойство: «$x_2 = 2^{x_1}$». (Это справедливо при любом из двух пониманий допустимости; если считать допустимыми все функции, то оговорка $[x_3 \leqslant x_1]$ является излишней.)

Свойства натуральных чисел, выражаемые формулами расширенного арифметического языка, назовем *аналитическими* или *слабо аналитическими* в зависимости от того, считаем ли мы допустимыми все функции или только финитные. Отождествляя свойства с множествами удовлетворяющих им объектов, мы будем говорить также об аналитических и слабо аналитических множествах.

Более точно, пусть α — формула расширенного арифметического языка, не содержащая функциональных параметров, а также числовых параметров, отличных от x_1, ..., x_k. Пусть n_1, ..., n_k — набор цифр. Подставив их вместо переменных x_1, ..., x_k, мыполучим суждение расширенного арифметического языка. Множество тех $\langle n_1, \ldots, n_k \rangle$, при которых это суждение истинно, будем называть *сопряженным* с формулой α (если допустимыми считать все функции) или *слабо сопряженным* с ней (если допустимыми считать лишь финитные функции). Множества, сопряженные (слабо сопряженные) с некоторыми формулами расширенного арифметического языка, будем называть *аналитическими* (соответственно *слабо аналитическими).*

Пример 12. Как показывает пример 11, множество $\{\langle x_1, x_2 \rangle \mid x_2 = 2^{x_1}\}$ является аналитическим, а также слабо аналитическим.

Всякое арифметическое множество, очевидно, является и аналитическим, и слабо аналитическим. Впоследствии мы докажем, что все слабо аналитические множества арифметичны; это обстоятельство оправдывает выбор длинного выражения «слабо аналитические» в качестве временного термина для их обозначения. Но не всякое аналитическое множество арифметично. Можно доказать, что множество номеров истинных суждений языка арифметики при

любой вычислимой нумерации слов алфавита А аналитично, в то время как, согласно теореме Тарского (см. приложение Б), оно не арифметично. (Заметим в скобках, что рассуждение, аналогичное доказательству теоремы Тарского, позволяет установить, что множеству суждений расширенного языка арифметики, истинных, если допустимыми считать все функции, аналогичным не является.) В следующем пункте мы покажем, что график всякой адресно вычислимой функции является слабо аналитическим множеством; в сочетании с упоминавшимся результатом об арифметичности всех слабо аналитических множеств это даст нам утверждение об арифметичности всех адресно вычислимых функций.

В.3. Выразимость адресно вычислимых функций в **расширенном арифметическом языке. В** этом пункте мы докажем, что график всякой адресно вычислимой функции является слабо аналитическим множеством Мы докажем впоследствии (В.4—В.6), что всякое слабо аналитическое множество арифметично, и это завершит доказательство арифметичности адресно вычислимых функций и, следовательно, арпфметичности множеств, перечисляемых такими функциями.

Пусть p — некоторая адресная программа. Докажем, что некоторые свойства, связанные с программой p, выразимы в расширенном языке арифметики. Говоря в дальнейшем об истинности оцененных формул расширенного арифметического языка, мы будем считать допустимыми лишь финитные функции, не оговаривая этого особо.

Напомним, что состояния адресных машин представляют собой последовательности натуральных чисел, в которых все члены, начиная с некоторого, равны 0; такие последовательности суть не что иное, как финитные функции.

Лемма В.1. *Свойство «состояние v_2 является непосредственным результатом применения программы p к состоянию v_1» выразимо в расширенном языке арифметики. (Это означает, что существует такая формула α расширенного арифметического языка, параметрами которой являются одноместные функциональные переменные v_1 и v_2, что результат подстановки вместо v_1 и v_2 двух констант для финитных функций тогда и только тогда является истинной оцененной формулой, когда непосредственным результатом применения программы p к состоянию, изображаемому первой константой, является состояние, изображаемое второй константой.)*

Доказательство. Пусть задана программа p, опишем способ построения требуемой формулы. (Эта формула будет, конечно, зависеть от выбора p.) Нужная нам формула α будет иметь вид $\alpha_1 \wedge \ldots \wedge \alpha_t$ (опущенные скобки можно поставить любым обра-

зом). Число n будет равно числу команд программы, и формула α_i будет соответствовать i-й команде. Каждая из формул α_i может быть одного из семи типов, в соответствии с семью типами команд, возможных для адресных машин. Эти формулы строятся в соответствии с семью пунктами определения непосредственного результата применения; способ их построения поясним на примерах.

Пример 1. Пусть 37-я команда программы имеет вид

$$37 \quad R(16) \leftarrow R(2) \cdot R(16).$$

В этом случае формула α_{37} будет такой:

$$(v_1(0) = 37) \to (\forall x_1(\neg(x_1 = 16) \to (v_2((x_1 + 1)) = v_1((x_1 + 1)))) \wedge (v_2(17) = (v_1(17) \cdot v_1(3)))).$$

Пример 2. Пусть 81-я команда программы имеет вид

81 ЕСЛИ $R(3) = R(4)$ ТО ИДТИ К 7 ИНАЧЕ К 23.

В этом случае формула α_{81} будет такой:

$$(v_1(0) = 81) \to (\forall x_1(v_2(x_1 + 1) = v_1(x_1 + 1)) \wedge (((v_1(4) = v_1(5)) \to (v_2(0) = 7)) \wedge (\neg(v_1(4) = v_1(5)) \to (v_2(0) = 23)))).$$

Доказательство леммы закончено.

Следующим шагом будет построение формулы, выражающей свойство «быть протоколом применения адресной программы p длины $x_1 + 1$». Протокол является последовательностью состояний, т. е. последовательностью финитных функций.

В нашем языке нет последовательностей функций, но есть объекты, их заменяющие: мы отождествим последовательность функций s^0, s^1, ... от одного натурального аргумента с функцией $S(n, m) = s^n(m)$ от двух натуральных аргументов. В соответствии с этим отождествлением мы будем называть всюду определенную функцию S из $\mathbb{N}^2$ в $\mathbb{N}$ *изображением протокола* программы p длины $k+1$, если последовательность $s^0, \ldots, s^k$ функций от одного аргумента, определяемых по формуле $s^i(x) = S(i, x)$, является протоколом применения адресной программы p.

Лемма В.2. *Существует формула* β *с двуместным функциональным параметром* w_1, *двумя одноместными параметрами* v_1, v_2 *и с одним числовым параметром* x_1, *выражающая следующее свойство:*

«w_1 *является изображением протокола длины* x_1+1,

v_1 *является начальным состоянием этого протокола,*

v_2 *является заключительным состоянием этого протокола».*

Доказательство. Искомая формула имеет следующий вид:

$$(([v_1 = w_1^0] \wedge [v_2 = w_1^{x_1}]) \wedge (v_2(0) = 0)) \wedge$$
$$\wedge \forall x_2 \forall v_3 \forall v_4 (([x_2 + 1 \leqslant x_1] \wedge$$
$$\wedge ([v_3 = w_1^{x_2}] \wedge [v_4 = w_1^{x_2+1}])) \rightarrow$$

$\rightarrow$[v_1 является непосредственным результатом применения программы p к v_3]),

В этой записи $[v_1 = w_1^0]$ есть сокращение для $\forall x_3 (v_1(x_3) = w_1(0, x_3))$; сокращения $[v_2 = w_1^{x_1}], [v_3 = w_1^{x_2}]$ и $[v_4 = w_1^{x_2+1}]$ расшифровываются аналогично — последнее, например, обозначает формулу

$$\forall x_3 (v_4(x_3) = w_1((x_2 + 1), x_3));$$

[v_4 является непосредственным результатом применения программы p к v_3] обозначает формулу из леммы В.1, в которой переменные v_1 и v_2 заменены соответственно на v_3 и v_4.

Теперь все готово для доказательства слабой аналитичности графиков адресно вычислимых функций.

Теорема В.1. *График адресно вычислимой функции из* $\mathbb{N}^k$ *в* $\mathbb{N}$ *является слабо аналитическим множеством.*

Доказательство. Пусть f — адресно вычислимая функция из $\mathbb{N}^k$ в $\mathbb{N}$, p — адресная программа, ее вычисляющая. График функции f состоит из таких наборов $\langle x_1, \ldots, x_k, x_{k+1} \rangle$, для которых существует протокол w_1 некоторой длины с начальным состоянием v_1 и заключительным состоянием v_2, для которого

$$v_1(0) = 1,\ v_1(1) = x_1,\ \ldots,\ v_1(k) = x_k,$$
$$v_1(x) = 0 \text{ при } x \geqslant k + 1,\ v_2(1) = x_{k+1}.$$

Записывая предыдущую фразу в виде формулы расширенного арифметического языка, получаем искомую формулу — формулу, с которой слабо сопряжен график функции f. Теорема доказана.

В следующих пунктах В.4—В.6 мы докажем, что всякое слабо аналитическое множество и, следовательно, график всякой адресно вычислимой функции являются арифметическими множествами.

В.4. Сведение расширенного арифметического языка к обычному. В этом пункте мы будем доказывать, что всякое слабо аналитическое множество арифметично, т. е. что добавление к языку арифметики переменных, пробегающих множества всех *финитных* функций одного и двух натуральных аргументов, не увеличивает выразительных возможностей языка. (Как уже отмечалось, условие финитности существенно, добавляя переменные, пробегающие множество *всех* функций, мы получаем существенно более выразительный язык.) Попытаемся в самых общих чертах объяснить, почему добавление

переменных для обозначения финитных функций несущественно. Дело в том, что множество этих функций счетно, их можно закодировать натуральными числами (и это кодирование окажется арифметическим в уточняемом ниже смысле), и мы можем говорить о кодах функций вместо того, чтобы говорить о самих функциях. Тем самым мы ограничиваемся рассмотрением натуральных чисел.

Уточним сказанное. Пусть ν — отображение, сопоставляющее каждому элементу множества $\mathbb{N}^k$, т. е. каждому набору (кортежу) из k натуральных чисел, некоторою (всюду определенную) финитную функцию одного аргумента Назовем его *способом кодирования* (или, короче, просто *кодированием) финитных функции одного аргумента с помощью элементов* $\mathbb{N}^k$, если каждая финитная функция соответствует по крайней мере одному (но, возможно, и не только одному) элементу $\mathbb{N}^k$; если набору $\langle a_1, \ldots, a_k \rangle$ соответствует функция s, то будем называть этот набор *кодом* функции s (при данном способе кодирования). Кодирование назовем *арифметическим,* если множество

$$\{\langle a_1, \ldots, a_k, x, y\rangle \mid \text{значение финитной функции с кодом } \langle a_1, \ldots, a_k\rangle \text{ на числе } x \text{ равно } y\}$$

является арифметическим.

Ключевым пунктом нашего доказательства арифметичности слабо аналитических множеств является следующее утверждение

(*) *при некотором k существует арифметический способ кодирования финитных функций элементами* $\mathbb{N}^k$.

В пп В.5 и В.6 будут предложены два различных доказательства этого утверждения. А сейчас мы покажем, как из него вытекает арифметичность слабо аналитических множеств.

Аналогично данному выше определению арифметического кодирования финитных функций одного аргумента можно дать определение арифметического кодирования (всюду определенных) финитных функций двух аргументов. Оказывается, что существование такового вытекает из существования арифметического кодирования для финитных функций одного аргумента—желая закодировать функцию f из $\mathbb{N}^2$ в $\mathbb{N}$, мы сначала кодируем ее «сечения», т. е. функции $f_n(x)=f(n, x)$, а затем кодируем последовательность кодов сечений. Более точно, имеет место следующая

Лемма В.3. *Если ν — арифметическое кодирование финитных функций одного аргумента элементами $\mathbb{N}^k$, причем $\langle 0, 0, \ldots, 0\rangle$ есть код нулевой финитной функции, то функция μ, сопоставляющая набору $\langle a_1^1, \ldots, a_k^1, \ldots, a_1^k, \ldots, a_k^k\rangle$ функцию из $\mathbb{N}^2$ в $\mathbb{N}$, равную*

$$\nu\left(\nu\left(a_1^1, \ldots, a_k^1\right)(p), \ldots, \nu\left(a_1^k, \ldots, a_k^k\right)(p)\right)(q)$$

на паре $\langle p, q \rangle$, *является арифметическим кодированием финитных функций двух аргументов элементами* $\mathbb{N}^k$.

Легко видеть, что ограничение «$\langle 0, 0, \ldots, 0 \rangle$ есть код нулевой последовательности» несущественно: мы можем переделать любое арифметическое кодирование в кодирование с таким свойством, обменяв друг с другом два кодовых обозначения; при этом арифметичность сохранится.

Итак, мы предполагаем, что зафиксировано некоторое арифметическое кодирование ν финитных функции одного аргумента элементами $\mathbb{N}^k$, а также некоторое арифметическое кодирование μ финитных функций двух аргументов элементами $\mathbb{N}^l$.

Мы построим для каждой формулы расширенного арифметического языка ее «перевод» — формулу языка арифметики, утверждающую то же самое, что исходная формула, но не о функциях, а об их кодах Для удобства мы добавим в язык арифметики новые переменные для чисел — по k новых переменных $V_i^1, \ldots, V_i^k$ для каждой одноместной функциональной переменной v_i, и по l новых переменных $W_i^1, \ldots, W_i^l$ для каждой двуместной функциональной переменной w_i. Ясно, что класс арифметических множеств не изменится от такого расширения — не все ли равно, как называются переменные! Дадим теперь точное определение перевода.

Пусть α— формула расширенного арифметического языка, имеющая числовые параметры $x_{p_1}, \ldots, x_{p_m}$, одноместные функциональные параметры $v_{q_1}, \ldots, v_{q_n}$ и двуместные функциональные параметры $w_{r_1} \ldots, w_{r_s}$. Пусть β — формула арифметического языка, параметры которой содержатся среди $x_{p_1}, \ldots, x_{p_m}, V_{q_1}^1, \ldots, V_{q_1}^k, \ldots$

$\ldots, V_{q_n}^1, \ldots, V_{q_n}^k, W_{r_1}^1, \ldots, W_{r_1}^l, \ldots, W_{r_s}^1, \ldots, W_{r_s}^l$.

Формула β называется *переводом* формулы α, если для любых натуральных чисел $x_{p_1}, \ldots, x_{p_m}, V_{q_1}^1, \ldots V_{q_n}^k, W_{r_1}^1, \ldots, W_{r_s}^l$ результат подстановки соответствующих им цифр вместо $x_{p_1}, \ldots, x_{p_m}, V_{q_1}^1, \ldots, V_{q_n}^k, W_{r_1}^1, \ldots, W_{r_s}^l$ в β является истинным суждением языка арифметики тогда и только тогда, когда результат подстановки $x_{p_1}, \ldots, x_{p_m}, \nu(V_{q_1}^1, \ldots, V_{q_1}^k) \ldots$

$\ldots, \nu(V_{q_n}^1, \ldots, V_{q_n}^k)$,

$\mu(W_{r_1}^1, \ldots, W_{r_1}^l), \ldots, \mu(W_{r_s}^1, \ldots, W_{r_s}^l)$ (точнее, не самих чисел и функций, а изображающих их констант) вместо

$x_{p_1}, \ldots, x_{p_m}, v_{q_1}, \ldots, v_{q_n}, w_{r_1}, \ldots, w_{r_s}$ в α является истинной оцененной формулой.

Теорема В.2. *Всякая формула расширенного языка арифметики имеет перевод.*

Прежде чем доказывать эту теорему, отметим, что из ее частного случая (случая формул без функциональных параметров), очевидно, следует интересующее нас утверждение об арифметичности слабо аналитических множеств: если множество слабо сопряжено с формулой расширенного языка арифметики, то оно сопряжено с ее переводом.

Доказательство. Предположим, что для элементарных формул расширенного языка переводы построены. Покажем, как построить их для остальных формул.

Лемма В.4. 1° *Если* β *— перевод* α, *то* $\neg\beta$ *— перевод* $\neg\alpha$;

2° *если* β_1 *и* β_2 *— переводы* α_1 *и* α_2, *то*

$$(\beta_1 \wedge \beta_2),\ (\beta_1 \vee \beta_2),\ (\beta_1 \to \beta_2),\ (\beta_1 \leftrightarrow \beta_2)$$

являются переводами формул

$$(\alpha_1 \wedge \alpha_2),\ (\alpha_1 \vee \alpha_2),\ (\alpha_1 \to \alpha_2),\ (\alpha_1 \leftrightarrow \alpha_2);$$

3° *если* β *— перевод* α, Q *— любой из знаков* $\forall$, $\exists$, *то*

$$Qx_i\beta \text{ — перевод } Qx_i\alpha,$$

$$QV_i^1 \ldots QV_i^k\beta \text{ — перевод } Qv_i\alpha,$$

$$QW_i^1 \ldots QW_i^l\beta \text{ — перевод } Qw_i\alpha.$$

Эта лемма непосредственно вытекает из определения перевода и определения истинности формул. Применяя ее, мы видим, что достаточно построить переводы элементарных формул, т. е. формул вида $(t=u)$. Заменяя их на $\exists\xi((t=\xi)\wedge(u=\xi))$ (ξ — числовая переменная, не входящая ни в t, ни в u) и применяя утверждения 2° и 3° доказанной леммы, мы видим, что достаточно перевести формулы вида $(t=\xi)$, где t — терм расширенного языка, а ξ — числовая переменная. Возможность перевода таких формул докажем индукцией по построению терма t:

1° если t — переменная или цифра, то в качестве перевода можно взять саму формулу;

2° если t есть (u_1+u_2), то, заменив формулу

$$((u_1+u_2)=\xi) \text{ на } \exists\eta_1\exists\eta_2(((u_1=\eta_1)\wedge (u_2=\eta_2))\wedge(\xi=(\eta_1+\eta_2))),$$

где η_1 и η_2 — числовые переменные, не входящие в t и отличные от ξ, мы сможем сослаться на предположение индукции и лемму В.4;

3° случай, в котором t есть $(u_1\cdot u_2)$, аналогичен предыдущему;

4° если t есть $p(u)$, где u— терм, а p — одноместная функциональная переменная, то, заменяя формулу

$$(p(u) = \xi)$$

На

$$\exists \eta ((u = \eta) \wedge (p(\eta) = \xi))$$

(где η переменная, не входящая в u и отличная от ξ), мы видим, что достаточно перевести формулу $(p(\eta) = \xi)$; это возможно в силу предположенной арифметичности кодирования;

5° случай, в котором $t = r(u_1, u_2)$, где u_1, u_2 — термы, r — двуместная функциональная переменная, аналогичен предыдущему.

Пример 1. В качестве перевода формулы

$$\forall x_1 (v_1(x_1) = v_2(x_1))$$

можно взять формулу

$$\forall x_1 \exists x_2 \text{ ([значение одноместной функции с кодом } \langle V_1^1, \ldots, V_1^k \rangle \text{ в } x_1 \text{ есть } x_2] \wedge$$
$$\wedge \text{ [значение одноместной функции с кодом } \langle V_2^1, \ldots, V_2^k \rangle \text{ в } x_1 \text{ есть } x_2]),$$

где записи в квадратных скобках обозначают формулы языка арифметики, выражающие записанные внутри скобок свойства и существующие в силу арифметичности кодирования.

Итак, для завершения доказательства арифметичности слабо аналитических множеств осталось лишь построить арифметическое кодирование финитных функций одного аргумента.

В.5. Первый способ построения арифметического кодирования — способ Гёделя. Мы начнем построение арифметического кодирования со следующего замечания: достаточно доказать, что существует всюду определенная арифметическая функция $\beta(x_1, ..., x_i, y)$ со следующим свойством:

(*) *для всякой конечной последовательности $n_0, ..., n_k$ натуральных чисел существуют такие $a_1, ..., a_i$, что* $\beta(a_1, \ldots, a_i, 0) = n_0$, $\beta(a_1, \ldots, a_i, 1) = n_1, \ldots, \beta(a_1, \ldots, a_i, k) = n_k$. (При этом значения $\beta(a_1, \ldots, a_i, y)$ при $y > k$ могут быть любыми.) В самом деле, пусть β обладает этим свойством. Тогда функция v, сопоставляющая набору $\langle x_1, \ldots, x_i, l \rangle$ финитную функцию $s(y)$, равную $\beta(x_1, \ldots, x_i, y)$ при $y \le l$ и равную 0 при $y > l$, является искомым арифметическим кодированием финитных функций одного аргумента элементами $\mathbb{N}^{i+1}$. То, что это кодирование, вытекает из свойства (*); то, что она арифметично, вытекает из арифметичностн функции β и арифметичностп свойства $y \le l$.

Итак, нам достаточно построить при некотором натуральном i функцию от $i + 1$ аргумента, обладающую свойством (*). В качестве такой функции мы, следуя Гёделю, возьмем функцию

$$\beta(x_1, x_2, y) = (\text{остаток от деления } x_1 \text{ на } x_2(y + 1) + 1).$$

Арифметичность этой функции следует из арифметичности свойства «x_3 есть остаток от деления x_1 на x_2», выражаемого формулой $[x_3 < x_2] \wedge \exists x_4(x_1 = ((x_2 \cdot x_4) + x_3))$, где $[x_3 < x_2]$ обозначает формулу $\exists x_5(((x_3 + x_5) + 1) = x_2)$. Чтобы доказать свойство (*), нам придется рассмотреть некоторые простые факты из теории чисел. Всюду дальше в этом пункте, говоря о числах, мы имеем в виду натуральные числа.

Число a называется делителем числа b, если $a = bc$ при некотором c. Если a является делителем чисел b и c, то оно является делителем их суммы и разности. Число $p > 1$, не имеющее делителей, отличных от 1 и p, называется простым. Всякое число разлагается на простые множители, причем однозначно: его разложения могут отличаться лишь порядком сомножителей. Если произведение нескольких чисел делится на простое число p, то один из сомножителей делится на p. Числа a и b называются взаимно простыми, если у них нет общих делителей, отличных от 1. Числа a и b взаимно просты тогда и только тогда, когда в их разложениях на простые множители нет общих множителей. Если числа $a_1, ..., a_n$, попарно взаимно просты, а число b делится на любое из них, то b делится на $a_1 \cdot a_2 \cdot \ldots \cdot a_n$.

Пусть $a_0, ..., a_k$ — попарно взаимно простые числа. Рассмотрим, какие наборы остатков $\langle r_0, \ldots, r_k \rangle$ возможны при делении некоторого числа x на числа $a_0, ..., a_k$. Остаток при делении на a_i — одно из чисел 0, 1, ..., a_i — 1; таким образом, существует $a_0 \cdot a_1 \cdot \ldots \cdot a_k$ возможных наборов остатков. Следующая лемма утверждает, что все возможности действительно реализуются.

Лемма В.5. (Китайская теорема об остатках.) *Пусть $a_0, ..., a_k$ — попарно взаимно простые числа, $r_0, \ldots, r_k$ — некоторые числа, причем $r_i < a_i$ при всех i. Тогда существует число x, дающее при делении на любое из чисел a_i остаток r_i.*

Доказательство. Назовем два числа эквивалентными, если они дают одинаковые остатки при делении на любое из a_i. Если два числа эквивалентны, то их разность делится на любое из чисел a_i и, следовательно, на $a_0, ..., a_k$ (в силу взаимной простоты). Поэтому никакие два из чисел 0, 1, ..., $a_0, ..., a_k - 1$ не эквивалентны и каждому соответствует свой набор остатков. Но этих чисел столько же, сколько возможных наборов. Поэтому любой набор $\langle r_0, \ldots, r_k \rangle$, у

которого $r_i < a_i$ при всех i, является набором остатков от деления некоторого числа x на a_a, ..., a_k.

Лемма В.6. *Для всякого числа n можно указать такое b, что числа* $b + 1, 2b + 1, \ldots, nb + 1$ *попарно взаимно просты. Число b можно выбрать большим любого заданного наперед натурального числа.*

Доказательство. Заметим сначала, что если p — общий простой делитель чисел $kb + 1$ и $lb + 1$, то p является делителем их разности — числа $(k — l)\ b$. Но делить число b он не может, так как иначе числа $kb + l$ и $lb + 1$ давали бы остаток 1 при делении на p. Поэтому $k — l$ делится на p. Из сказанного следует, что числа $b + 1$, ..., $nb + 1$ будут взаимно просты, если они не будут иметь общих делителей, меньших n. Этого можно достигнуть, взяв, например, b кратным $1 \cdot 2 \cdot \ldots \cdot n$; тогда числа $b + 1$, ..., $nb + 1$ будут давать остаток 1 при делении на любое число от 2 до n

Теперь мы легко можем доказать свойство (*) для построенной нами функции β. В самом деле, пусть n_0, ..., n_k — любые натуральные числа; нам надо найти такие x_1 и x_2, чтобы остаток от деления x_1 на $x_2(i + 1) + 1$ был равен n_i при $i \leq k$. Согласно лемме В.6 можно найти такое x_2, что числа $x_2 + 1, \ldots, x_2(k + 1) + 1$ попарно взаимно просты и x_2 больше любого из чисел n_0, ..., n_k; осталось выбрать x_1 с помощью леммы В.5.

Построение арифметического кодирования методом Гёделя закончено. В следующем пункте мы рассмотрим другой метод построения арифметического кодирования, не использующий теоретико-числовых соображений и предложенный Смальяном в книге Smul-1 y a n R. M. Theory of formal systems. — Princeton, 1961, русский перевод которой (Смальян Р. Теория формальных систем) вышел в издательстве «Наука» в 1981 г.

В 6. Второй способ построения арифметического кодирования — способ Смальяна. Введем понятие арифметического кодирования конечных подмножеств $\mathbb{N}$ натуральными числами, аналогичное понятию кодирования финитных функций элементами $\mathbb{N}^1$. А именно, функцию τ, сопоставляющую всем натуральным числам некоторые подмножества $\mathbb{N}$, назовем *кодированием,* если любое конечное подмножество $\mathbb{N}$ является значением τ на некотором числе; если $\tau(y) = A$, то будем называть число y *кодом* множества A (относительно τ). Кодирование назовем *арифметическим,* если множество

$$\{\langle x, y\rangle \mid x \in \tau(y)\}$$

является арифметическим подмножеством $\mathbb{N}^2$.

Заметим, что (в нарушение аналогии с определением кодирования финитных функций) мы не требуем, чтобы подмножества, соответствующие всем натуральным числам, были конечными.

Арифметические кодирования конечных подмножеств $\mathbb{N}$ натуральными числами существуют. Прежде чем доказывать это, мы покажем, как отсюда вытекает существование арифметического кодирования финитных функций одного аргумента. Итак, пусть τ — арифметическое кодирование конечных подмножесгл натуральными числами.

Лемма В.7. *Существует арифметическая функция из* $\mathbb{N}^2$ *в* $\mathbb{N}$, *определенная на всем* $\mathbb{N}^2$ *и сопоставляющая различным элементам* $\mathbb{N}^2$ *различные элементы* $\mathbb{N}$.

Доказательство. Эта функция сопоставляет паре $\langle n_1, n_2\rangle$ натуральное число, являющееся наименьшим кодом множества $\{n_1, n_1 + n_2\}$: значение этой функции на паре $\langle n_1, n_2\rangle$ равно k тогда и только тогда, когда $n_1 \in \tau(k)$, $n_1 + n_2 \in \tau(k)$, любое число, принадлежащее $\tau(k)$, равно n_1 или $n_1 + n_2$ и все числа, меньшие k, не обладают такими свойствами. Записывая сказанное в виде формулы языка арифметики, устанавливаем арифметичность построенной функции. То, что эта функция сопоставляет разным парам разные числа, легко следует из ее определения.

Теперь мы построим арифметическое кодирование конечных подмножеств $\mathbb{N}^2$ натуральными числами. А именно, числу k мы сопоставим подмножество $\mathbb{N}^2$, состоящее из тех пар $\langle x, y\rangle$, для которых число $\nu(x, y)$ принадлежит множеству $\tau(k)$. (Здесь ν — функция из предыдущей леммы, τ — арифметическое кодирование конечных подмножеств $\mathbb{N}$ натуральными числами.) Легко видеть, что всякое конечное подмножество $\mathbb{N}^2$ поставлено в соответствие некоторому числу и мно жество

$$\{\langle k, x, y\rangle \mid \text{пара } \langle x, y\rangle \text{ принадлежит подмножеству } \mathbb{N}^2, \text{ соответствующему числу } k\}$$

является арифметическим подмножеством $\mathbb{N}^3$. (Эти требования мы и имели в виду, говоря о построении арифметического кодирования подмножеств $\mathbb{N}^2$.)

Теперь все готово для построения арифметического кодирования финитных функций одного аргумента элементами $\mathbb{N}^2$. Опишем, какая функцня s ставится в соответствие паре натуральных чисел $\langle k, l\rangle$. Пусть A — подмножество $\mathbb{N}^2$, соответствующее k при арифметическом кодировании подмножеств $\mathbb{N}^2$. Если для данного числа x, меньшего или равного l, существуют такие y, для которых $\langle x, y\rangle \in A$, то $s(x)$ равно наименьшему из таких y; если таких y нет, то

$s(x)$= 0; при x> l значение $s(x)$ равно 0. Всякой паре $\langle k, l \rangle$ соответствует финитная функция — функция, равная 0 на аргументах, бóльших l. Чтобы найти код данной финитной функции s, надо в качестве l взять наибольшее число, на котором она отлична от 0, а в качестве k взять код множества $\{\langle x, y \rangle \mid x \leqslant l \text{ и } y = s(x)\}$. Записывая определение построенного кодирования в виде формулы языка арифметики, убеждаемся в его арифметичностн.

Итак, для завершения доказательства существования арифметического кодирования финитных функций методом Смальяна осталось построить арифметическое кодирование конечных подмножеств $\mathbb{N}$. При этом построении мы будем использовать двоичную запись чисел, о которой можно прочесть, например, в уже упоминавшейся книжке О. Оре (Приглашение в теорию чисел. — М.: Наука, 1980).

Двоичная запись каждого натурального числа (кроме 0) начинается с 1; если мы условимся отбрасывать первую единицу, то получим взаимно однозначное соответствие между всеми положительными целыми числами и всеми словами в алфавите {0, 1}. Таким образом, инструкция «прибавь к числу 1, запиши его в двоичной системе и отбрось первую единицу» устанавливает взаимно однозначное соответствие между множеством натуральных чисел и множеством слов алфавита,{0, 1}:

0	пустое слово
1	0
2	1
3	00
4	01
5	10
6	11
7	000
8	001
. . . . .	

Слова в правой колонке расположены в порядке возрастания их длины, а слова одной длины расположены в словарном порядке. Число, стоящее в левой колонке, мы будем называть номером слова, стоящего в той же строке в правой колонке. При этой нумерации всякому множеству (двоичных) слов соответствует множество натуральных чисел. (Например, множеству слов, состоящие только из нулей, соответствует множество чисел, становящихся степенями двойки после прибавления 1.) Благодаря этому мы

можем говорить об арифметичности множеств слов, имея в виду арифметичность соответствующих множеств натуральных чисел. Мы будем говорить также об арифметичностн свойств слов, отождествляя свойство с множеством удовлетворяющих ему объектов. Аналогично определяются арифметические подмножества множества $(\{0, 1\}^{\infty})^n$, элементами которого являются последовательности из n слов; такие подмножества естественно отождествляются со свойствами последовательностей из n слов.

Установим теперь арифметнчность некоторых конкретных свойств.

1° Слово X предшествует слову Y в упомянутом порядке. В самом деле, это имеет место тогда и только тогда, когда номер слова X меньше номера слова Y.

2° Слово X состоит из одних нулей. В самом деле, в силу сказанного выше достаточно проверить арифметнчность свойства «быть степенью двойки», а она, как объяснялось в приложении Б, вытекает из того, что число x тогда и только тогда является степенью двойки, когда всякий его делитель либо равен 1, либо четен.

3° Слово X состоит из одних единиц.

В самом деле, слово тогда и только тогда состоит из одних единиц, когда следующее за ним слово состоит из одних нулей.

4° Слово Y состоит из одних нулей и имеет ту же длину, что и слово X.

В самом деле, это равносильно тому, что слово Y — наибольшее в смысле рассмотренного порядка слово, состоящее из одних нулей и не превосходящее слова X.

5° Слова X и Y имеют одну и ту же длину. В самом деле, это равносильно тому, что существует слово Z, состоящее из одних нулей, имеющее одинаковую длину со словом X и со словом Y.

6° Слово X является соединением слов Y и Z, т. е. получается приписыванием слова Z к слову Y справа. $X = YZ$. Это, пожалуй, самый сложный пункт нашего рассуждения, в котором нам придется вспомнить о способе кодирования слов. Грубо говоря, арифметичность этого свойства вытекает из того, что для получения числа z, двоичная запись которого получается приписыванием друг к другу двоичных записей чисел x и y, необходимо умножить x на число $2^{(\text{длина } y)}$ и прибавить z. Следует, конечно, учесть, что при нашей нумерации слов мы прибавляем к числу единицу, а также не учитываем первую единицу двоичного разложения числа. Проделаем теперь все это подробнее. Пусть числа x, y и z являются номерами слов X, Y и Z. Это означает, что двоичная запись $x + 1$ есть $1X$, двоичная запись $y + 1$ есть $1Y$, а двоичная запись 2+1 есть 1Z. Пусть u — код слова, которое имеет ту же длину, что и Z, но состоит из одних нулей. Тогда двоичная

запись числа $u+1$ будет иметь вид $\underbrace{10\,000\,\ldots\,0}_{\text{длина } z}$. Если слово X является соединением слов Y и Z, то, умножая $y+1$ на $u+1$ и прибавляя (z+1) — (u + 1), мы получим число x +1. Из сказанного следует арифметичность интересующего нас свойства — искомой формулой будет запись на языке арифметики следующей фразы:

«существует u, являющееся кодом слова той же длины, что и Z, и состоящего из одних нулей, для которого $(y+1)\cdot(u+1)+(z-u)$ равно $x+1$».

Из арифметичности свойства «быть соединением» легко вывести арифметичность нескольких рассматриваемых дальше свойств.

7° Слово X является началом слова Y, т. е. существует такое слово Z, что Y является соединением X и Z.

8° Слово X является концом слова Y, т. е. существует такое слово Z, что Y есть соединение Z и X.

9° Слово X является подсловом (частью) слова Y. В самом деле, слово X тогда и только тогда является подсловом слова У, когда оно является началом некоторого конца слова Y.

10° Слово X есть соединение слов Y, Z и V. В самом деле, слово X является соединением слов Y, Z и V тогда и только тогда, когда существует такое слово W, что W есть соединение Y и Z, а X — соединение W и V.

Аналогично может быть доказана арифметичность свойства «Y есть соединение X_1, ..., X_n» при любом фиксированном n.

Теперь мы можем построить арифметическое кодирование конечных множеств натуральными числами. Пусть x, y — натуральные числа, X, Y — соответствующие им слова. Пусть — самое длинное слово из одних нулей, входящее в Y. Будем считать, что число x принадлежит множеству $\tau(y)$, если слово $1U1X1U1$ входит в Y. Докажем, что τ является кодированием конечных множеств. Пусть $\{x_1, \ldots, x_n\}$— конечное множество чисел, $\{X_1, \ldots, X_n\}$— множество соответствующих им слов. Пусть U — слово, состоящее из нулей и более длинное, чем любое из слов X_1, ..., X_n. Обозначим через Y слово

$$1U1X_11U1X_21U1\ \ldots\ 1U1X_n1U1.$$

Номер слова Y и будет кодом множества $\{x_1, \ldots, x_n\}$. Арифметичность этого кодирования легко следует из арифметичности рассмотренных нами свойств слов. Построение кодирования методом Смальяна закончено.

Приложение Г.

ЯЗЫКИ, СВЯЗАННЫЕ С АССОЦИАТИВНЫМИ ИСЧИСЛЕНИЯМИ

В этом приложении мы рассмотрим примеры языков с относительно просто устроенными множествами истинных утверждений. Эти примеры будут связаны с так называемыми ассоциативными исчислениями.

Ассоциативным исчислением в алфавите И называется произвольная конечная совокупность правил, разрешающих определенного вида преобразования слов в И. Эти правила называются двусторонними подстановками или (коль скоро мы не рассматриваем здесь односторонних подстановок) просто *подстановками* в алфавите И. Каждая подстановка в алфавите И записывается в виде

$$P \leftrightarrow Q,$$

где *P* и *Q* суть слова в И, а буква ↔ не принадлежит алфавиту И. (Например, ци↔цы есть подстановка в русском алфавите.) Подстановка $P \leftrightarrow Q$ означает разрешение заменять слово *P,* если оно встретится как часть другого слова, на слово *Q* и обратно. Сказанное оформляется более точно в виде следующих определении. Для каждого ассоциативного исчисления (т. е. для каждого списка подстановок) вводится понятие смежных слов и эквивалентных слов. Два слова *A* и *B* называются *смежными* (записывается $A \perp B$), коль скоро существуют такие слова *P, Q, X, Y,* что: 1) *A=XPY,* 2) *B=XQY n* 3) хотя бы одна из подстановок $P \leftrightarrow Q$ и $Q \leftrightarrow P$ есть подстановка рассматриваемого исчисления. Цепочку $\langle C_1, \ldots, C_n \rangle$ слов из И$^\infty$ назовем цепочкой смежности, если для каждого *i* имеет место $C_i \perp C_{i+1}$. Два слова *A* и *B* называются *эквивалентными,* коль скоро существует такая цепочка смежности $C_1, C_2, \ldots, C_n$, что $C_1 = A$, $C_n = B$.

Замечание 1. Если произвести факторизацию множества И$^\infty$ по так введенному отношению эквивалентности, получится алгебраическая система с ассоциативной операцией (возникающей при факторизации из операции приписывания друг к другу слов), отсюда и название — ассоциативное исчисление.

Пусть фиксировано некоторое ассоциативное исчисление в алфавите И. Существует алгоритм, позволяющий для каждых двух слов *A* и *B* из И$^\infty$ распознавать, смежны они или нет. Такой алгоритм состоит, например, в переборе всех четверок слов *P, Q, X, Y*, длина которых не превосходит длин *A* и *B,* и проверке условий 1), 2), 3). Таким образом, множество всех пар смежных слов разрешимо относительно

$И^{\infty} \times И^{\infty}$. Однако существование алгоритма, распознающего эквивалентность слов, очевидно лишь в простейших случаях.

Пример 1. Пусть $И = \{a, b, c\}$ и ассоциативное исчисление задано следующими подстановками:

$$ab \leftrightarrow ba,$$
$$ac \leftrightarrow ca,$$
$$bc \leftrightarrow cb.$$

Очевидно, что A и B эквивалентны тогда и только тогда, когда число букв a в слове A равно числу букв a в слове B, и то же самое выполняется для букв b и c. Такое исчисление естественно назвать *коммутативным.*

В общем случае не ясно, каким алгоритмом можно было бы обнаружить для произвольных слов, эквивалентны они или нет, т. е. имеется ли связывающая их цепочка смежных слов. И действительно, как показали А. А. Марков и Э. Л. Пост, возможно ассоциативное исчисление с неразрешимой проблемой распознавания эквивалентности (под проблемой распознавания эквивалентносги как раз и понимается проблема отыскания алгоритма, распознающего эквивалентность слов). Доказательство существования таких исчислений приводится, например, в монографии С. К. Клини (Клини С. К. Введение в математику. — М.: ИЛ, 1957). Мы здесь приведем без доказательства следующий пример, принадлежащий Г. С. Цейтину.

Пример 2. Пусть $И = \{a, b, c, d, e\}$ и ассоциативное исчисление задано подстановками

$$ac \leftrightarrow ca,$$
$$ad \leftrightarrow da,$$
$$bc \leftrightarrow cb,$$
$$bd \leftrightarrow db$$
$$eca \leftrightarrow ce,$$
$$edb \leftrightarrow de,$$
$$cca \leftrightarrow ccae.$$

Как показал Г. С. Цейтьн, для этого исчисления не существует алгоритма, распознающего эквивалентность слов.

Ассоциативное исчисление будем называть *разрешимым,* если для него существует алгоритм распознавания эквивалентности; в противном случае будем называть ею *неразрешимым.* Очевидно, что разрешимость ассоциативного исчисления равносильна разрешимости множества всех пар эквивалентных слов (и разрешимости множества всех пар неэквивалентных слов) относительно $И^{\infty} \times И^{\infty}$.

Фиксируем некоторое ассоциативное исчисление $\mathfrak{Z}$ в алфавите И. Множество всех слов (в алфавите $И_*$) вида $A*B$, где

$A \in И^{\infty}$, $B \in И^{\infty}$ и A эквивалентно (соответственно, неэквивалентно) B, обозначим через T^{+} (соответственно через T^{-}), так что $T^{+} \cup T^{-} = И^{\infty} \times И^{\infty}$. Тогда разрешимость исчисления $\mathfrak{Z}$ означает разрешимость множества T^{+} (что равносильно разрешимости множества T^{-}) относительно $И^{\infty} \times И^{\infty}$

Замечание 2 Поэтому множество T^{+} (как и T^{-}), построенное для исчисления из примера 2, представляет собой индивидуальный пример неразрешимого подмножества множества $И^{\infty} \times И^{\infty}$. Характеристическая функция этого подмножества представляет собой в этом случае индивидуальный пример невычислимой функции.

С каждым ассоциативным исчислением в алфавите И мы свяжем теперь два языка — *позитивный язык,* утверждения которого будут утверждениями об эквивалентности произвольных двух слов в И, и *негативный язык,* утверждениями которого будут утверждения о неэквивалентности произвольных двух слов в И. В обоих случаях в качестве утверждений целесообразно рассматривать элементы множества $И^{\infty} \times И^{\infty}$ только в первом случае, для позитивного языка, слово $A*B$ будет интерпретироваться как утверждение об эквивалентности слов A и B, и потому множеством истинных утверждений будет служить T^{+}, тогда как во втором случае, для негативного языка, слово $A*B$ будет интерпретироваться как утверждение о неэквивалентности слов A и B, и потому множеством истинных утверждений будет служить T^{-}. Вспомним теперь, что в п. 1.3 р. 1.5.1 мы договорились считать язык заданным, коль скоро указана соответствующая фундаментальная пара. Итак, пусть фиксирован алфавит И и ассоциативное исчисление $\mathfrak{Z}$ в этом алфавите. Мы объявляем $\langle И_{*}, T^{+} \rangle$ фундаментальной парой позитивного языка, сопряженного с исчислением $\mathfrak{Z}$, а $\langle И_{*}, T^{-} \rangle$ — фундаментальной парой негативного языка, сопряженного с исчислением $\mathfrak{Z}$.

Нас будет занимать вопрос о возможности ввести полную непротиворечивую дедуктику для $\langle И_{*}, T^{+} \rangle$ и для $\langle И_{*}, T^{-} \rangle$. Мы увидим, что в первом случае этот вопрос решается всегда положительно, а во втором — в зависимости от разрешимости исчисления $\mathfrak{Z}$.

Лемма Г.1. *Множество Е всех цепочек смежности разрешимо относительно* $И^{\infty}_{*}$.

Доказательство вытекает из существования алгоритма, распознающего смежность любых двух слов из

Теорема Г.1. *Для любого ассоциативного исчисления множество всех пар эквивалентных слов перечислимо.*

Доказательство. Введем на $И_*^\infty$ функцию φ, полагая

$$\varphi(C_1 * C_2 * \ldots * C_n) = C_1 * C_n$$

для каждого слова $C_1 * C_2 * \ldots * C_n$, где все C_i суть слова из $И^\infty$. Очевидно, что A н и B эквивалентны тогда и только тогда, когда $A * B = \varphi(C)$ для некоторой цепочки смежности C. Поэтому $T^+ = \varphi(E)$, где E—множество всех цепочек смежности. Множество E разрешимо относительно $И_*^\infty$ (по лемме Г.1) и, следовательно, перечислимо (по лемме 2). Функция φ очевидным образом вычислима, а потому перечислимым будет и множество $\varphi(E)$. Но $\varphi(E) = T^+$, а перечислимость T^+ и надо было доказать.

Замечание 3. Таким образом, в случае неразрешимости исчисления T^+ будет служить примером перечислимого, но не разрешимого подмножества перечислимого множества $И^\infty \times И^\infty$. В силу леммы 3 всякий такой пример является одновременно примером перечислимого подмножества с неперечислимым дополнением. Ср. ниже замечание 5.

Следствие теоремы Г.1. *Для (фундаментальной пары) позитивного языка, сопряженного с произвольным ассоциативным исчислением, можно ввести полную непротиворечивую дедуктику.*

Замечание 4. Чтобы получить дедуктику, о которой говорится в этом следствии, нет нужды обращаться к теореме 1. Проще предъявить дедуктику $\langle И_*, E, \varphi \rangle$, где E и φ таковы, как в доказательстве теоремы Г.1; она и будет полной непротиворечивой дедуктикой для $\langle И_*, T^+ \rangle$. Эта дедуктика является совершенно естественной с содержательной точки зрения; в самом деле, лучшим доказательством эквивалентности слов A и B является предъявление связывающей их цепочки смежности.

Перейдем к вопросу о дедуктике для $\langle И_*, T^- \rangle$.

Теорема Г.2. *Пусть дано ассоциативное исчисление. Множество всех пар неэквивалентных слов тогда и только тогда перечислимо, когда это исчисление разрешимо.*

Доказательство. Заметим прежде всего, что $И^\infty \times И^\infty$ перечислимо (пример 6 из р.1.5.2). Пусть исчисление разрешимо; тогда T^- разрешимо относительно $И^\infty \times И^\infty$, а значит, само перечислимо (по лемме 2). Пусть теперь T^- перечислимо; поскольку его дополнение T^+ до перечислимого множества также перечислимо (по теореме Г.1), то в силу леммы 3 множество T^- разрешимо (относительно $И^\infty \times И^\infty$), а вместе с ним разрешимо и само рассматриваемое исчисление.

Замечание 5. Множество T^+, таким образом в случае неразрешимости исчисления служит примером перечислимого множества с неперечислимым дополнением (до некоторого объемлющего перечислимого множества); в силу леммы 3 всякий такой пример служит одновременно примером перечислимого множества, но являющегося разрешимым (опять-таки относительно некоторого перечислимого надмножества). Таким образом, существование неречислимого неразрешимого множества, доказанное нами в п.5.5, является следствием утверждения о существовании неразрешимых ассоциативных исчислений. Впрочем, обычные доказательства неразрешимости ассоциативных исчислений (в том числе исчисления из примера 2) как раз и опираются на существование перечислимого неразрешимого множества; этот факт, следовательно, подлежит сам доказательству, не опирающемуся на существование неразрешимых ассоциативных исчислений.

Следствие теоремы Г.2. *Для фундаментальной пары негативного языка, сопряженного с некоторым ассоциативным исчислением, тогда и только тогда можно ввести полную непротиворечивую дедуктику, когда это исчисление разрешимо.*

Введем теперь для произвольного ассоциативного исчисления $\mathfrak{I}$ в алфавите И *универсальный язык,* утверждениями которого будут служить как утверждения об эквивалентности слов, так и утверждения о неэквивалентности. Здесь нам придется отличать первые утверждения от вторых. С этой целью пополним алфавит И еще одной буквой, буквой $\neg$, в предположении, что она, как и $\leftrightarrow$ и *, не входит в И. Алфавит $И \cup \langle *, \neg \rangle$ обозначим через Л. Обозначим через $\neg T^-$ множество всех слов вида $\neg P$, где $P \in T^-$. Положим $T^0 = T^+ \cup \neg T^-$ и образуем фундаментальную пару $\langle Л, T^0 \rangle$. Элемент t из T^0 естественно интерпретировать как истинное утверждение об эквивалентности слов (если $t \in T^+$) или о неэквивалентности слов (если $t \in \neg T^-$).

Теорема Г.3. *Для любого ассоциативного исчисления $\mathfrak{I}$ соответствующее ему множество T^0 тогда и только тогда перечислимо, когда исчисление разрешимо.*

Доказательство. Если $\mathfrak{I}$ разрешимо, то T^- перечислимо (теорема Г.2), а потому перечислимо и $\neg T^-$ (пример 5 из п. 5.2). Тогда T^0 перечислимся по лемме 5. Пусть теперь перечислимо T^0. Образуем множество $\neg Л^\infty$ всех слов в алфавите Л, начинающихся с $\neg$; это множество перечислимо (примеры 2 и 5 из р.1.5.2). По лемме 5 перечислимся пересечение $T^0 \cap \neg Л^\infty$ Но $T^0 \cap \neg Л^\infty = \neg T^-$.

Поэтому перечислимо $\neg T\neg$, а вместе с ним и $T\neg$ (пример 5 из п.5.2); но тогда по теореме Г.2 разрешимо исчисление $\mathfrak{F}$.

Следствие. *Для фундаментальной пары универсального языка, сопряженного с некоторым ассоциативным исчислением, тогда и только тогда можно ввести полную непротиворечивую дедуктику, когда это исчисление разрешимо.*

ПриложениеД.

ИСТОРИЧЕСКИЕ ЗАМЕЧАНИЯ

Один из наиболее выдающихся математиков XX века (и, безусловно, самый выдающийся математический логик) Курт Гёдель (Kurt Godel) родился 28 апреля 1906 года в городе Брно (ныне Чехословакия, тогда Австро-Венгрия). С 40-х годов Гёдель работал в Принстоне (США), где и умер 14 января 1978 года. С именем Гёделя связаны важнейшие теоремы математической логики: теорема о полноте исчисления предикатов (1930 г.), теорема о неполноте арифметики (1930 г.), теорема о непротиворечивости аксиомы выбора и континуум-гипотезы (1938 г.).

Теорема о полноте исчисления предикатов утверждает, что можно предложить полную и непротиворечивую дедуктику для языка логики предикатов, а точнее —что некоторая конкретная (и ранее известная) дедуктика такова; таким образом, в этой дедуктике можно доказать все истины логики предикатов, т. е. всякую формулу, выражающую «закон логики» (и нельзя доказать никакие иные формулы). (Под «законом логики» понимается формула, истинность которой сохраняется при любом истолковании участвующих в ней имен.) Теорема о неполноте (ей и посвящена настоящая брошюра), напротив, утверждает, что подобная ситуация невозможна в случае арифметики: не только известные дедуктики не являются одновременно полными и непротиворечивыми, но такая дедуктика в принципе невозможна; как было разъяснено выше в основном тексте брошюры, ни при каком понятии формального доказательства не удается доказать все истины арифметики и только их. Ниже будет приведена формулировка теоремы о неполноте в той форме, как она была высказана самим Гёделем. Теорема о непротиворечивости аксиомы выбора и аксиомы, выражающей континуум-гипотезу, гласит, что теория множеств остается непротиворечивой после присоединения указанных двух аксиом, коль скоро она была непротиворечивой до такого присоединения. Этот результат Гёделя — первый фундаментальный

результат, относящийся к исследованию непротиворечивости теоретико-множественных утверждений, — в значительной степени изменил наши представления о смысле этих утверждений и положил начало новому направлению в математической логике.

О названных теоремах Гёделя можно прочесть в книгах, указанных в предисловии. Гёделю принадлежит и много других важных понятий и результатов [в частности первое (1934 г.) определение понятия рекурсивной функции: рскурсивность по Эрбрану — Гёделю]; всех их невозможно здесь перечислить. Мы сейчас остановимся подробнее на первоначальной гёделевой формулировке теоремы о неполноте.

Знаменитая работа Курта Гёделя «Ober formal unentscheidbare Satze der Principia Mathematica und verwandter Systemel» («О формально неразрешимых предложениях Principia Mathematica и родственных систем») была напечатана на с. 173—198 1-й тетради 38-го тома (за 1931 г.) лейпцигского журнала «Monatshefte fur Mathematik und Physik» (поступила 17/XI 1930). Предварительная сводка результатов была опубликована в венском журнале Anzeiger der Akademia der Wissenschaften in Wien, Mathematisch-naturwissenschaftliche Klasse, № 19 за 1930 г. (заседание от 23 октября 1930 г.).

В этой статье для широкого класса формальных систем устанавливалось неизбежное существование в каждой из таких систем неразрешимого утверждения— неразрешимого в том смысле, что ни оно, ни его отрицание не могло быть выведено из аксиом системы. Именно, в статье Гёделя была сформулирована следующая теорема (теорема VI на с. 187):

Для каждого ω-непротиворечивого рекурсивного класса $\varkappa$ *формул* существует такая рекурсивная *классовая формула r,* что ни *v*Gen *r,* ни Neg(*v*Gen*r*) не принадлежит к Flg($\varkappa$) (где *v* есть *свободная переменная формулы r).*

Дадим некоторые пояснения к приведенным формулировкам. Эти пояснения предполагают наличие у читателя простейших сведений из математической логики. Речь здесь идет о формулах некоторой формальной системы *P,* которая строится на страницах 176—178 статьи Гёделя. Мы не будем приводить точных формулировок, а ограничимся следующей цитатой из Гёделя: «В сущности, *P* есть та система, которая получается, если надстроить пеановские аксиомы логикой Principia Mathematica (числа в качестве индивидов, отношение «следования за» в качестве неопределяемого понятия)» (с. 176). Курсив несет на себе определенный смысл. Он означает, что речь идет не непосредственно о знакосочетаниях рассматриваемой формальной системы (переменных, формулах и т. п.), а о номерах этих знакосочетаний в некоторой фиксированной нумерации (называемой

теперь гёделевской). Классовая формула — это формула с одной свободной переменной. Стало быть, *классовая формула*— это натуральное число, являющееся номером классовой формулы. Через *v* Gen *r* обозначается номер формулы, полученной навешиванием квантора общности по переменной с номером *v* на формулу с номером *r*; через Neg(*v* Gen *r*)—номер отрицания предыдущей формулы. Через Flg($\varkappa$) обозначается класс номеров всех тех и только тех формул, которые выводимы из класса формул, номера которых образуют класс $\varkappa$ (вывод из произвольного класса формул предполагает возможность использовать в процессе вывода также и аксиомы, так что в данном случае происходит присоединение класса $\varkappa$ к аксиомам исходной системы). Термины «рекурсивный класс» и «рекурсивная формула» мы оставим без объяснений; эти термины означают некоторую определимость рассматриваемых классов и формул с помощью примитивно рекурсивных функций (в статье Гёделя такие функции называются просто «рекурсивными»). Свойство ω-непротиворечивости, налагаемое на класс, означает условие более сильное, нежели простая непротиворечивость. Если непротиворечивость класса означает невозможность вывести из него как некоторую формулу, так и ее отрицание, то ω-непротиворечивость означает невозможность вывести как некоторую формулу вида «существует такое *x,* что $\mathfrak{A}(x)$», так и все формулы вида «не $\mathfrak{A}(0)$», «не $\mathfrak{A}(1)$», «не $\mathfrak{A}(2)$» и т. д.

В обозначениях обсуждаемой статьи Гёделя класс $\varkappa$ *формул* (т. е. номеров формул) называется ω-непротиворечивым, если не существует *классовой формулы a,* для которой: 1) $\mathrm{Neg}(v\ \mathrm{Gen}\ a) \in \mathrm{Flg}(\varkappa)$, 2) $\mathrm{Sb}\left(a^{v}_{Z(n)}\right) \in \mathrm{Flg}(\varkappa)$ при всех *n* (здесь $\mathrm{Sb}\left(a^{v}_{Z(n)}\right)$ означает номер результата подстановки в формулу с номером *a* формулы с номером *Z(n)* вместо переменной с номером *v,* причем *Z(n)*—номер выражения для числа *n).* Таким образом, теорема VI гласит, что для любого класса формул, подчиненного некоторым условиям, существует формула сравнительно простого вида такая, что ни она, ни ее отрицание невыводимы из этого класса. Поскольку в основе формальной системы *P,* подразумеваемой в названной теореме (ведь речь идет о формулах этой системы и о выводимости по правилам этой системы), лежат арифметические аксиомы Пеано, то сама эта теорема часто интерпретируется как теорема о неполноте формальной арифметики. Неполнота понимается здесь в синтаксическом смысле (см. приложение А).

Замечание 1. Если под формальной арифметикой понимать систему *P,* то неполнота формальной арифметики представляет собой весьма

частный случай теоремы VI, получающийся при $\varkappa = \varnothing$ (и справедливый в предположении, что сама *P* является ω-непротиворечивой, т. е. что ω-непротиворечив класс ее аксиом); в этом случае $\mathrm{Flg}(\varkappa)$ состоит просто из номеров всех формул, доказуемых в *P*.

Замечание 2. Правда, сама неразрешимая формула, указываемая в теореме VI, а именно, формула с номером *v* Gen *r*, еще не имеет арифметического характера, т. е. еще не записана на простейшем арифметическом языке. Однако на этот счет в статье Гёделя содержатся важные дальнейшие результаты. Именно, формула называется «арифметической», если она строится с помощью переменных, пробегающих натуральный ряд, отношения равенства и операций сложения и умножения (заметим, что знаки =, +, · не входят в исходный алфавит системы *P*. Поэтому «арифметическая формула» может существовать лишь в подходящем расширении системы *P*. В рамках же *P* эти знаки следует рассматривать как сокращающие. Так, выражение $x_1 = y_1$ понимается, согласно подстрочному примечанию 21 на с. 177, как сокращение для формулы $«x_2\Pi(x_2(x_1) \supset x_2(y_1))»$; здесь $x_2\Pi$ означает квантор общности по x_2. (Для $x+y$ и $x - y$ такие расшифровки в статье Гёделя не приводятся.)).

Далее, на с. 193 статьи Гёделя формулируется теорема VIII:

В каждой формальной системе, упоминаемой в теореме VI, существуют неразрешимые арифметические утверждения.

Замечание 3. Как указывает Гёдель (с. 190), его доказательство теоремы VI проходит не только для конкретной системы *P,* о которой идет речь в его статье, но для любой системы, обладающей следующими основными свойствами:

1) аксиомы и правила вывода системы рекурсивно определимы;
2) каждое рекурсивное отношение определимо внутри системы.

Как отмечает Гёдель, эти свойства выполняются для аксиоматических систем теории множеств Цермело — Френкеля и фон Неймана, а также для аксиоматической теории чисел, основанной на аксиомах Пеано и рекурсивных определениях. Во всех этих системах существуют, следовательно, неразрешимые предложения: чтобы обнаружить это, достаточно положить $\varkappa = \varnothing$ (ср. выше замечание 1). Правда, утверждение предыдущей фразы справедливо лишь в предположении ω-непротиворечивости рассматриваемой системы. Это предположение во всех конкретных случаях образует рабочую гипотезу, вытекающую из нашего убеждения в разумности рассматриваемой системы, т. е. в том, что она адекватно отражает некоторую реальность.

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.: Техніка, 1977.
6. Кузичев А. С. Диаграммы Венна. – М.: Наука, 1968.
7. Кононюк А. Ю. Вища математика. У 2 ч. Ч.1, – К: Кольори, 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: Вища школа, 1990.
13. Минский М. Фреймы для представления знаний. –М.: Энергия, 1979.
14. Вильсон А. Дж. Энтропийные методы моделирования сложных систем. – М.: Наука, 1978.
15. Юрков В. Ю., Лукина О. В. /Прикладная геометрия, вып. 8, N 18 (2006), стр. 9-36
16. Кузичев А.С. Дианраммы Венна. – М.: Наука, 1968.
17. Успенский В.А. Теорема Геделя о неполноте. – М.: Наука, 1982.

□