М. М. ГЛУХОВ, А. Б. ШИШКОВ

МАТЕМАТИЧЕСКАЯ ЛОГИКА ДИСКРЕТНЫЕ ФУНКЦИИ ТЕОРИЯ АЛГОРИТМОВ

РЕКОМЕНДОВАНО
УМО вузов России по образованию
в области информационной безопасности
в качестве учебного пособия для студентов вузов,
обучающихся по направлению подготовки (специальности)
090301 — «Компьютерная безопасность»
и 090303 — «Информационная безопасность
автоматизированных систем»

ЛАНЬ®

САНКТ-ПЕТЕРБУРГ · МОСКВА · КРАСНОДАР
2012

ББК 22.176я73
Г 55

Глухов М. М., Шишков А. Б.

Г 55 Математическая логика. Дискретные функции. Теория алгоритмов: Учебное пособие. — СПб.: Издательство «Лань», 2012. — 416 с.: ил. — (Учебники для вузов. Специальная литература).

ISBN 978-5-8114-1344-7

Учебное пособие содержит полное изложение материала учебных дисциплин «Математическая логика и теория алгоритмов» и «Дискретные функции» Государственного образовательного стандарта высшего профессионального образования по специальностям и направлениям «Компьютерная безопасность», «Информационная безопасность автоматизированных систем» и некоторым другим смежным специальностям.

Пособие состоит из трех взаимосвязанных частей, представляющих основы математической логики, теории дискретных функций и теории алгоритмов.

Предназначено для студентов вузов, обучающихся по специальностям и направлениям в области информационной безопасности, а также для аспирантов и студентов вузов других технических специальностей и направлений, изучающих дискретную математику.

ББК 22.176я73

Рецензенты:

В. Б. АЛЕКСЕЕВ — доктор физико-математических наук, профессор, зав. кафедрой математической кибернетики факультета ВМК Московского государственного университета им. М. И. Ломоносова;
В. П. ЗЯЗИН — кандидат физико-математических наук, профессор кафедры информационной безопасности Московского государственного технического университета радиотехники, электроники и автоматики.

Обложка
Е. А. ВЛАСОВА

ПРЕДИСЛОВИЕ

Данное учебное пособие состоит из трех взаимосвязанных частей, составляющих соответственно основы математической логики, дискретных функций и теории алгоритмов. Пособие содержит систематическое изложение учебного материала по математической логике, теории алгоритмов и дискретным функциям, изучаемого в цикле математических дисциплин по различным специальностям в области информационной безопасности.

Первая часть пособия содержит строгое изложение классического материала по алгебре логики, логике предикатов, исчислениям высказываний и предикатов 1-го порядка, называемого также узким исчислением предикатов, включая доказательства их непротиворечивости и полноты. Изложен также метод резолюций для распознавания истинности некоторых формул узкого исчисления предикатов. Вторая часть посвящена способам задания и изучению свойств булевых функций и функций k-значной логики при $k > 2$. Особое внимание уделяется свойствам функций, играющим важную роль в криптографии, в частности групповой классификации функций, вопросам сравнения произвольных функций с линейными и аффинными функциями, декомпозиции функций. В третьей части излагаются три основных подхода к определению понятия алгоритма, обсуждается связь между ними. Вводятся необходимые понятия о сложности алгоритмов и приводятся примеры нахождения нижних оценок сложности. Особое внимание уделяется вопросам сложности переборных задач. В частности, доказывается известная теорема Кука об NP-полноте проблемы выполнимости булевых функций. Рассматривается также вопрос о сложностной классификации систем булевых уравнений, приводится доказательство результата Шефера

о полиномиальности проблемы распознавания совместности систем булевых уравнений, составленных из функций любого одного класса Шефера.

Содержание пособия основано на реализации компетентностного подхода, положенного в основу федеральных государственных образовательных стандартов высшего профессионального образования (ФГОС ВПО) третьего поколения в области информационной безопасности.

Пособие предназначено для преподавания дисциплины «Математическая логика и теория алгоритмов» по специальностям:

090301 «Компьютерная безопасность»;

090303 «Информационная безопасность автоматизированных систем»;

090305 «Информационно-аналитические системы безопасности».

Материал пособия (часть, посвященная теории дискретных функций) может быть использован в процессе преподавания дисциплины «Дискретная математика» для тех же специальностей.

Несмотря на явную направленность пособия на подготовку специалистов по защите информации, оно может быть использовано также студентами других специальностей при изучении математической логики, теории алгоритмов и дискретной математики. В ходе изложения материала пособия некоторые сравнительно несложные фрагменты доказательств предлагаются в качестве упражнений для самостоятельной работы. Отдельно системы задач и упражнений в пособии не приводятся в связи с тем, что в 2008 г. был издан сборник [75], содержащий задачи и упражнения по всем разделам данного пособия. Кроме того, много задач и упражнений приводится в известных сборниках задач [73, 76]. Материал пособия опробован авторами в ходе преподавания соответствующих учебных дисциплин в Институте криптографии, связи и информатики и отражен во внутривузовских учебных пособиях [12, 13].

В пособии нумерация определений, теорем, утверждений и формул производится в каждой части по главам с указанием номера соответствующей главы.

Часть I

Математическая логика

Современную математическую логику называют наукой о математических доказательствах. Историю ее развития условно можно разделить на два этапа. На первом этапе, длившемся примерно до середины XIX в., она развивалась в недрах общей науки логики и имела своей целью формализацию и математический анализ отдельных фрагментов человеческого мышления. Наиболее полное и систематическое изложение основ общей логики впервые было дано древнегреческим философом Аристотелем (IV в. до н. э.). Аристотель в своих трактатах «Первая аналитика», «Вторая аналитика», «Категория», «Об истолковании», «Топика», «О софистических доказательствах», объединенных его комментаторами под общим названием «Органон» (означающем орудие или средство познания), выявляет общие закономерности рассуждений, формулирует ряд основных законов логики, разрабатывает теорию определений и доказательств, намечая в общих чертах дедуктивный путь построения научных дисциплин и, в частности, логики. В указанных работах Аристотеля содержатся зачатки алгебры высказываний, являющейся простейшей составной частью современной математической логики. Однако для становления и развития алгебры высказываний до современного состояния понадобилось более 2 тыс. лет. Важнейшие качественные сдвиги в развитии математической (или символической) логики связаны с работами немецкого математика и философа Г. В. Лейбница (1646–1716). В определенной степени они были предопределены общим развитием математики и, в частности, введением буквенных обозначений в алгебру и анализ.

Г. В. Лейбниц мечтал о создании универсального символического языка, который бы позволил заменить содержательные рассуждения формальными вычислениями и помог из данных опыта выводить все их логические следствия. В логике Г. В. Лейбниц, подобно Ф. Виету (1540–1603) и Р. Декарту (1596–1650) в алгебре, ввел буквенные обозначения для высказываний и сформулировал ряд законов алгебры высказываний. По Лейбницу каждое понятие должно быть сведено к фиксированному набору простых, т. е. не разложимых далее

понятий, а сложные понятия должны выводиться из простых с помощью логических операций.

В принципе неосуществимая и метафизическая идея, высказанная Г. В. Лейбницем, о сведении всего содержательного в человеческом мышлении к формальной вычислительной процедуре сыграла тем не менее определенную положительную роль в развитии математической логики.

Дальнейшее продвижение в развитии математической логики было осуществлено в работах ирландского математика Дж. Буля (1815–1864) «Математический анализ логики», (1847) и «Законы мысли» (1854). Дж. Буль ввел операции сложения, умножения и дополнения (отрицания) высказываний и распространил на них некоторые законы действий над числами. В построенной Дж. Булем алгебре все символы принимают лишь два значения 1 и 0 (соответствующие истине и лжи). Особое внимание Дж. Буль уделяет изучению логических функций и уравнений, выводу заключений из заданных посылок.

В дальнейшем алгебра Буля совершенствовалась и разрабатывалась многими математиками. Так, английский математик С. Джеванс (1835–1882) классифицировал булевы функции по группе преобразований переменных и ввел понятие типа функции. Немецкий математик Э. Шредер (1841–1902) видоизменил алгебру Буля, заменив в ней операцию сложения на дизъюнкцию (разделительное «или» на соединительное), и сформулировал в полученной алгебре принцип двойственности; немецкий математик Г. Фреге (1848–1925) применил аксиоматический подход к изучению алгебры высказываний, т. е. развил исчисление высказываний; ряд важных логических законов выявил американский математик Де Морган (1806–1871).

Большой вклад в развитие алгебры высказываний внесли русские математики: профессор Казанского университета П. С. Порецкий (1846–1907), профессор Московского университета И. И. Жегалкин (1869–1947) и др.

Так, П. С. Порецкий в 1880-х гг. построил вполне законченную теорию качественных умозаключений, позволившую

описать все заключения, выводимые из заданной системы посылок.

И. И. Жегалкин впервые обратил внимание на алгебру высказываний как на теорию функций над полем из двух элементов и дал подробное и доступное изложение этой теории.

В целом к концу XIX в. построение алгебры высказываний по существу было завершено. Вместе с тем к этому времени назрел новый этап развития математической логики как теории математических доказательств. Важнейшим стимулом к новому скачку в развитии математической логики стал все более распространяющийся в математике аксиоматический метод. В основу любой аксиоматически строящейся теории кладется некоторая легко обозримая система основных (неопределяемых) понятий и отношений и основных (недоказываемых) их свойств, называемых аксиомами теории. Далее все понятия определяются через основные, а все утверждения теории логически выводятся из аксиом. Одним из необходимых требований, предъявляемых к системе аксиом теории, является ее непротиворечивость. Требуется, чтобы, исходя из выбранной системы аксиом, нельзя было доказать никакого утверждения и его отрицания.

Особенно остро вопрос о непротиворечивости встал после создания Н. И. Лобачевским (1811–1860) неевклидовой геометрии, утверждения которой явно расходились с утверждениями привычной для всех геометрии Евклида. И хотя в весьма далеко развитой Н. И. Лобачевским геометрии внутренних противоречий не обнаруживалось, вопрос о строгом доказательстве ее непротиворечивости оставался открытым. Для решения этого вопроса необходимо было четко очертить средства доказательств и убедиться в их надежности, так как в противном случае даже полученное в теории противоречие не будет говорить о противоречивости системы аксиом, поскольку причины могут заключаться не в аксиомах, а в самих средствах вывода утверждений из них. Таким образом, перед математической логикой была поставлена задача создания теории математических доказательств и методов доказательства непротиворечивости аксиоматических теорий.

До этого основным способом убеждения в непротиворечивости теории был так называемый метод интерпретаций одной теории в рамках другой. Так, в 1871 г. немецкий математик Ф. Клейн (1849–1925) с использованием идей английского математика А. Кэли (1821–1895) построил модель геометрии Лобачевского в недрах геометрии Евклида и тем самым свел ее непротиворечивость к непротиворечивости геометрии Евклида. В 1899 г. немецкий математик Д. Гильберт (1862–1943) построил аналитическую интерпретацию геометрии Евклида, сведя тем самым вопрос о непротиворечивости и геометрии Евклида, и геометрии Лобачевского к вопросу о непротиворечивости теории действительных чисел, т. е. арифметики. Значение арифметики как фундамента математики осознавалось математиками и раньше, а потому на обоснование арифметики обращалось самое серьезное внимание.

К концу XIX в. усилиями таких математиков, как Р. Дедекинд (1831–1916), В. Гамильтон (1805–1856), К. Вейерштрасс (1815–1897), Г. Кантор (1845–1918), Г. Грассман (1809–1877), Д. Пеано (1853–1932), была построена аксиоматическая теория натуральных чисел и на ее основе теории целых, рациональных и действительных чисел. И когда уже казалось, что решение вопроса о непротиворечивости всей математики приближается к завершению, обнаружились парадоксы в созданной и развитой в 1870-х гг. немецким математиком Г. Кантором теории множеств. А так как теория множеств широко использовалась при построении числовых систем и, в частности, при построении теории действительных чисел, то появление парадоксов в теории множеств нанесло серьезный удар по основаниям математики.

В качестве примера приведем здесь один парадокс, опубликованный английским математиком Б. Расселом в 1903 г.

Обозначим буквой A множество всех таких множеств, которые не являются элементами самих себя. Тогда для любого множества M имеем

$$M \in A \Longleftrightarrow M \notin M.$$

Полагая теперь M равным A, приходим к явному противоречию:

$$A \in A \Longleftrightarrow A \notin A.$$

Различные математики по-разному реагировали на появление парадоксов в теории множеств. Некоторые считали их несерьезными и легко преодолимыми, другие глубоко переживали и искали выход из создавшегося положения.

Анализ парадоксов показывал, что все они основываются на том или ином свойстве «самоприменимости»: сущность, о которой идет речь, в каждом из них определяется посредством некоторой совокупности, к которой она сама принадлежит. Вместе с тем в математике, да и в других науках, идея самоприменимости часто используется и не приводит к противоречиям, а потому полное исключение образования самоприменимых понятий вряд ли было целесообразным. Задача же отделения «вредных» и «безвредных» самоприменимых понятий или условий их использования не представляется тривиальной.

Один из способов преодоления парадоксов в математике основывался на аксиоматических теориях множеств.

Аксиоматический подход характеризуется стремлением математиков ограничить слишком вольное обращение с понятием множества и особенно с бесконечными множествами. В связи с этим были предложены различные системы аксиом теории множеств. Первые две системы аксиом теории множеств были построены независимо и опубликованы в 1908 г. английским математиком Б. Расселом и немецким математиком Е. Цермело. В дальнейшем эти системы пересматривались и совершенствовались многими другими математиками (А. Френкелем, фон Нейманом, П. Бернайсом, В. Куайном, К. Геделем и др.). В настоящее время наиболее распространенной является система аксиом Цермело — Френкеля ZF. Все известные системы аксиом теории множеств устроены так, что в них невозможны рассуждения, ведущие к известным парадоксам. Но вопрос о возможности появления новых противоречий остается открытым.

Наиболее радикальную позицию в пересмотре оснований математики заняли так называемые интуиционисты. Они рассматривали парадоксы как симптомы неблагополучия во всей математике и главную их причину видели в свободном переносе методов математики с конечных множеств на бесконечные. Число интуиционистов было невелико. Однако интуиционистские взгляды разделяли такие крупные математики, как Л. Кронекер (1823–1891), А. Пуанкаре (1854–1913), Ф. Э. Борель (1871–1956), А. Лебег (1375-1941), Г. Вейль (1885–1955). Наиболее ярким представителем интуиционистов был голландский математик Л. Брауэр (1881–1966).

Движение интуиционизма неоднородно, однако можно выделить некоторые характерные черты, присущие различным школам интуиционизма и неоинтуиционизма. Укажем некоторые из них.

1. Отрицание актуальной бесконечности, т. е. существования бесконечных множеств как неких завершенных объектов. Признаются лишь бесконечные последовательности при условии, что указано правило построения их членов по предыдущим членам. Другими словами, признается лишь абстракция потенциальной бесконечности. Так, Л. Брауэр не признает множества всех натуральных чисел как математического объекта, а признает лишь правило построения последовательности натуральных чисел.

2. Отождествление понятия существования объекта с возможностью его эффективного построения. Для интуиционистов существует только то, что можно построить. Например, известное доказательство существования предельной точки в бесконечном замкнутом множестве действительных чисел для интуиционистов несостоятельно, поскольку в нем не указывается способ построения искомой предельной точки.

3. Отказ от использования закона исключенного третьего при доказательстве свойств объектов бесконечного множества. Закон исключенного третьего гласит, что «из двух высказываний, одно из которых является отрицанием другого, хотя бы одно — истинно».

Приведем один пример доказательства с использованием закона исключенного третьего, несостоятельного с точки зрения интуиционизма (см. [17]).

Теорема. *Существуют два иррациональных числа a и b, таких что число a^b — рационально.*

Для доказательства рассмотрим число $\sqrt{2}^{\sqrt{2}}$. Если оно рационально, то теорема верна при $a = b = \sqrt{2}$. Если же число $\sqrt{2}^{\sqrt{2}}$ иррационально, то теорема верна при $a = \sqrt{2}^{\sqrt{2}}$, $b = \sqrt{2}$. Это есть теорема чистого существования, поскольку в ней доказано существование искомых чисел a, b без указания эффективного способа их построения. При доказательстве мы использовали закон исключенного третьего («число α или рационально, или иррационально») в применении к объектам бесконечного множества действительных чисел. Для интуиционистов проведенные рассуждения не являются доказательством, и утверждение теоремы не считалось бы ими истинным, если бы не существовало другого, конструктивного доказательства этой теоремы. Однако в данном случае дело обстоит лучше, поскольку в 1934 г. советский математик А. О. Гельфонд решил 7-ю проблему Гильберта, доказав, что если α — алгебраическое число, $\alpha \neq 0, 1$, а β — алгебраическое иррациональное число, то α^β — трансцендентное (а потому и иррациональное) число (см. [11]). Следовательно, число $\sqrt{2}^{\sqrt{2}}$ — иррационально.

Термин «интуиционизм» происходит от основной доктрины соответствующего направления в математике — «изначальной интуиции» целого положительного числа или построения по математической индукции. Эта интуиция носит не чувственный и не эмпирический характер — она врожденна. Недаром первый из интуиционистов в математике нового времени Л. Кронекер провозгласил, что целые числа создал Бог, а все остальное в математике есть творение человека (см. [63], с. 293). Сам термин «эффективное построение объекта» интуиционистами также не определяется. Он считается интуитивно ясным. Вот что пишет по этому поводу ученик и последователь Л. Брауэра А. Гейтинг: «Математическое построение

должно быть таким непосредственным для разума, чтобы не нуждаться ни в каких обоснованиях» (см. [11], с. 14).

Философскую основу интуиционизма составляет субъективный идеализм, который проявляется в самых различных принципах интуиционизма, в понимании природы, сущности математики, ее объектов, в игнорировании объективной истины в математике. Так, в ответ на критику интуиционизма Л. Брауэр говорит, что, если даже кто-либо и верит в «объективную истину», характер ее метафизичен, а математические доказательства не могут опираться на метафизические соображения (см. [63], с. 267–271). О том же свидетельствует и высказывание А. Гейтинга: «Но только Брауэр открыл объект, который действительно требует иной формы логики. Этот объект — умственное математическое построение... Вы должны понять, в чем заключалась программа Брауэра. Она состояла в исследовании умственного математического построения как такового, безотносительно к таким вопросам о природе конструируемых объектов, как вопрос: существуют ли эти объекты независимо от нашего сознания?» (см. [11], с. 9–10).

С математической точки зрения признание принципов интуиционизма означало существенное ограничение исторически сложившихся областей математики, особенно математического анализа.

Основной вопрос о непротиворечивости в математике у интуиционистов остался нерешенным. Непротиворечивость их математики основывалась лишь на интуитивной надежности применяемых ими средств доказательств.

Вместе с тем нельзя не отметить и положительной роли интуиционизма в развитии математики. В частности, понимание существования как потенциальной возможности построения легло в основу интенсивно развивающегося в настоящее время конструктивного направления в математике (см. ниже).

Существенные ограничения математики, вытекающие из принципов интуиционистов, вызвали резкую критику интуиционизма со стороны многих крупных математиков того времени и, в частности, немецкого математика Д. Гильберта. Обнаруженные в теории множеств парадоксы не смогли поколебать глубокую веру Д. Гильберта в надежность основ клас-

сической математики. «Никто не может изгнать нас из рая, который создал нам Кантор», — заявил Д. Гильберт (см. [63], с. 15).

С именем Д. Гильберта и его школы связано третье направление в обосновании математики — формалистическое.

Д. Гильберт предложил широкую программу, получившую впоследствии название программы финитизма Гильберта. Программа состояла из двух частей. Первая часть заключалась в формализации арифметики, анализа и теории множеств. Предполагалось положить в основу математики некоторое множество аксиом-формул и некоторые четко определенные правила вывода одних формул из других, с тем чтобы в полученной формальной системе со строго детерминированным процессом вывода можно было получить все основные утверждения математики. Во второй части предполагалось доказать непротиворечивость построенного формального исчисления, т. е. невозможность вывода в нем какого-либо утверждения вместе с его отрицанием. При этом требовалось, чтобы рассуждения, используемые для доказательства непротиворечивости, носили настолько элементарный характер, что в их правильности нельзя было усомниться. Д. Гильберт настаивал, чтобы в теории доказательств (или в метаматематике) разрешалось пользоваться только финитными методами. Точного определения финитных методов Гильберт не дает. Их наиболее определенную характеристику можно извлечь из слов его ученика Ж. Эрбрана (1908–1931) (см. [63], с. 321): «... всегда рассматривается лишь конечное и определенное число предметов и функций; функции эти точно определены, причем определение позволяет произвести однозначное вычисление их значений; никогда не утверждается существование какого-либо объекта без указания способа построения этого объекта; никогда не рассматривается (как вполне определенное) множество всех предметов x какой-либо бесконечной совокупности; если же говорится, что какое-то рассуждение (или теорема) верно для всех этих x, то это означает, что это общее рассуждение можно повторить для каждого конкретного x, причем само это общее рассуждение следует при этом

рассматривать только как образец для проведения таких конкретных рассуждений».

Отсюда видно, что на рассуждения в метаматематике Д. Гильберт накладывал даже более жесткие ограничения, чем интуиционисты.

Программа Гильберта послужила толчком к дальнейшему развитию математической логики и к разработке формализованных логико-математических теорий. Сам Д. Гильберт и его ученики П. Бернайс, В. Аккерман и другие много усилий затратили на создание формализованной арифметики. При этом они исходили в основном из имеющейся арифметической системы аксиом Пеано и формализации логики предикатов, осуществленной ранее Г. Фреге, Б. Расселом и А. Уайтхэдом.

Для реализации программы Гильберта в части арифметики оставалось, с одной стороны, показать, что в рамках формализованной арифметики можно сформулировать и доказать основные утверждения классической арифметики, а с другой — доказать непротиворечивость созданной теории.

Обе эти задачи оказались невыполнимыми в принципе. В 1931 г. появилась работа молодого немецкого математика К. Гёделя, из результатов которой, грубо говоря, следовало, что каждая непротиворечивая логическая система, содержащая формализацию арифметики, содержит формулу, которую нельзя ни доказать, ни опровергнуть в данной системе. Этот вывод К. Гёделя обнаружил принципиальную ограниченность дедуктивных возможностей достаточно богатых формализованных теорий и показал неосуществимость первой части программы Гильберта. Более того, из результатов К. Гёделя следовало также, что никакое предложение, которое можно точным образом интерпретировать как выражающее непротиворечивость логической системы, содержащей арифметику, не может быть доказано в этой системе. Это утверждение, по существу, говорит о невыполнимости и второй части программы Гильберта о доказательстве непротиворечивости математики финитными методами.

Указанное заключение не умаляет, конечно, большого научного значения программы Гильберта и работ по ее реализации. Формалистический подход к обоснованию математики во

многом определил лицо современной математической логики, изучающей различные логические исчисления, являющиеся в большинстве своем формализованными теориями гильбертовского типа.

В целом, упомянутые нами аксиоматический, интуиционистский и формалистический подходы к обоснованию математики и различные их модификации не вывели математику из трудностей, связанных с парадоксами. Однако в ходе их реализации была проделана огромная работа по анализу оснований математики и ее дальнейшему развитию. В частности, была развита математическая логика, образовались такие богатые и интенсивно развивающиеся в настоящее время области науки, как теория моделей и теория алгоритмов. Развитые в этих теориях методы позволили прояснить и решить ряд классических задач математики (например, проблему континуума, 10-ю проблему Гильберта о разрешимости в целых числах диофантовых уравнений и др.).

Создание теории алгоритмов привело к появлению так называемого конструктивного направления в математике, исправившего один из существенных изъянов интуиционизма. Выше было указано, что интуиционисты понимали существование объекта только в смысле возможности его эффективного построения. Однако само понятие эффективного построения ими не определялось и понималось лишь интуитивно.

Конструктивисты, сохранив ту же точку зрения на существование объекта, дали определение эффективного построения объекта, использовав строго определенное понятие алгоритма. И хотя в конструктивной математике сохранились основные ограничения интуиционистской математики, конструктивизм оказал положительное влияние на развитие математики, ибо даже приверженцы классической (канторовской) математики в своих доказательствах стали более критически относиться к теоремам чистого существования, стремясь заменить их при возможности конструктивными доказательствами.

Большой вклад в развитие математической логики и ее приложений внесли советские математики. Укажем в качестве

примеров лишь некоторые работы советских математиков по математической логике и ее приложениям.

В 1925 г. А. Н. Колмогоров опубликовал работу (Матем. сб., т. 32, с. 646–667, 1924–1925), посвященную анализу закона исключенного третьего и соотношения между интуиционизмом и формализмом. Позднее (в 1932 г.) он дал математическое объяснение содержательного смысла конструктивной логики и интуиционистской логики как исчисления проблем.

В 1928 г. В. И. Гливенко доказал, что интуиционистская логика не эквивалентна никакому трехзначному логическому исчислению.

В 1936 г. А. И. Мальцев (1909–1967) опубликовал свою знаменитую локальную теорему, которая в дальнейшем нашла широкое применение в самых различных областях математики и легла в основу так называемого метода компактности в теории моделей. Сам А. И. Мальцев применил эту теорему к решению ряда проблемных вопросов теории групп (см. [39]).

Указанными работами и большой серией других работ А. И. Мальцев заложил основы новой науки «Теория моделей», и он по праву считается одним из ее создателей. В теории моделей язык и методы математической логики применяются к изучению конкретных математических структур.

Большой вклад в теорию моделей внесли многочисленные ученики А. И. Мальцева: А. Д. Тайманов, М. И. Каргополов, Ю. Л. Ершов и др.

В 1939 г. П. С. Новиков (1905–1975) предложил простое и остроумное доказательство непротиворечивости арифметики, в основе которого лежит построенное им логическое исчисление, допускающее, кроме операций логики высказываний, счетные логические суммы и произведения (см. [46]).

В 1940-х гг. А. А. Марков (1906–1979) разработал теорию нормальных алгоритмов и доказал алгоритмическую неразрешимость ряда массовых математических проблем, в частности проблемы тождества слов и проблемы изоморфизма для конечно определенных полугрупп [40]. В последующие годы А. А. Марков и его ученики использовали нормальные алгоритмы для развития конструктивного направления математи-

ки. А. А. Марков заслуженно считается одним из создателей конструктивной математики.

В 1952 г. И. С. Новиков доказал неразрешимость проблемы тождества слов, а также некоторых других проблем для конечно определенных групп [47]. В дальнейшем его учеником С. И. Адяном были разработаны некоторые общие подходы к доказательству отсутствия алгоритмов, решающих те или иные массовые проблемы [2].

В 1958 г. А. Н. Колмогоров и В. А. Успенский, пытаясь охарактеризовать абстрактно работу вычислителя, дали новое определение алгоритма как процесса переработки графов и доказали его эквивалентность известным определениям [28]. Тем самым было получено дополнительное обоснование правильности определения алгоритма.

В 1965 г. А. Н. Колмогоровым теория алгоритмов была использована для разработки подхода к определению понятия количества информации [27].

В 1972 г. молодой ленинградский математик Ю. В. Матиясевич получил завершающий результат в многолетнем пути математиков по решению 10-й проблемы Гильберта. Было доказано отсутствие алгоритма для распознавания разрешимости диофантовых уравнений [43].

Отметим еще, что в 1930–1940-х гг. алгебра высказываний нашла интересные приложения в электротехнике при построении различных типов релейно-контактных схем. Приоритет в этих приложениях математической логики принадлежит советскому ученому В. И. Шестакову, написавшему в 1935 г. работу «Алгебра релейно-контактных схем» (опубликованную лишь в 1941–1946 гг.).

В 1958–1960 гг. О. Б. Лупанов решил ряд трудных проблем по оценке сложности реализации булевых функций формулами алгебры логики, контактными и функциональными схемами, найдя асимптотические значения соответствующих функций Шеннона [36].

В заключение отметим, что отечественными математиками написано большое число оригинальных монографий по математической логике и ее приложениям. Некоторые из них указаны в списке литературы.

Глава 1

МНОЖЕСТВА С ОТНОШЕНИЯМИ И ОПЕРАЦИЯМИ

1.1. МНОЖЕСТВА И ОПЕРАЦИИ НАД НИМИ

Во всех областях современной математики, за исключением узко специальных ее разделов, связанных с аксиоматическим построением теории множеств, понятие множества принято считать основным, неопределяемым понятием. Создатель теории множеств немецкий математик Г. Кантор пояснял понятие множества следующим образом: «Множество, или совокупность, — это собрание определенных и различных объектов нашей интуиции или интеллекта, мыслимое в качестве единого». Говорят также, что множество — это совокупность, или собрание, или семейство каких-либо реально существующих или мыслимых объектов. Предполагается, что объекты, входящие в множество, попарно различимы. Объекты, из которых составлено множество, называются его элементами. Множества и элементы множеств обозначают различными буквами без индексов и с индексами. При этом, как правило, множества и элементы отождествляют с их обозначениями. Например, вместо фразы «элемент, обозначенный буквой a, содержится в множестве, обозначенном буквой A», говорят короче: «элемент a содержится в множестве A (или принадлежит множеству A)» и пишут $a \in A$. Запись $a \notin A$ означает, что a не является элементом множества A.

Если одно и то же множество обозначено разными буквами A, B, то говорят, что множества A, B равны, и пишут $A = B$. Если же буквами A, B обозначены разные множества, то пишут $A \neq B$ и говорят, что множества A, B не равны. Знаки

$=$ и $\neq$ будут использоваться и в более общей ситуации, когда множества обозначаются не буквами, а каким-либо другим способом.

Множество обычно задается или перечислением всех его элементов, или указанием правила перечисления, или указанием каких-либо характеристических свойств его элементов. В первом случае множество обозначается в виде заключенного в фигурные скобки списка его элементов, например

$$\{1, 2, 3, 4, 5\}\ ,\ \{a\}.$$

Во втором случае записывают в фигурных скобках несколько первых элементов с многоточием, например

$$\{0, 2, 4, 6, ...\}.$$

Если же множество A задается системой свойств $P_1, ..., P_k$ его элементов, то пишут

$$A = \{a : P_1, ..., P_k\} \text{ или } A = \{a | P_1, ..., P_k\},$$

и говорят, что A есть множество всех элементов a, обладающих свойствами $P_1, ..., P_k$. Например,

$$\{a : a \text{ — целое число, } a^3 + 2a - 3 = 0\}$$

есть множество целочисленных корней уравнения $x^3+2x-3 = 0$.

Для удобства в формулировках определений и утверждений теории множеств вводится множество, совсем не содержащее элементов. Такое множество называется пустым и обозначается символом $\emptyset$. Например, зачастую приходится говорить о множестве, про которое неизвестно заранее, содержит ли оно хотя бы один элемент. Так мы говорим о множестве решений системы уравнений, не решая ее и, значит, не зная еще, имеет ли она хотя бы одно решение.

Для некоторых часто используемых ниже и знакомых со средней школы числовых множеств примем стандартные обозначения:

$\mathbb{N} = \{1, 2, 3, ...\}$ — множество натуральных чисел;

$\mathbb{N}_0 = \{0, 1, 2, 3, ...\}$ — множество целых неотрицательных чисел;

$\mathbb{Z} = \{..., -2, -1, 0, 1, 2, ...\}$ — множество целых чисел;

$\mathbb{Q} = \{\frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0\}$ — множество рациональных чисел;

$\mathbb{R}$ — множество действительных (или вещественных) чисел;

$\mathbb{R}^+$ — множество положительных действительных чисел;

$\mathbb{C} = \{a + bi : a, b \in \mathbb{R}, i^2 = -1\}$ — множество комплексных чисел;

$\overline{m, n}$ (для $m, n \in \mathbb{Z}$) — множество $\{m, m + 1, ..., n\}$, если $m < n$; $\{m\}$, если $m = n$; $\emptyset$, если $m > n$.

Если каждый элемент множества A является элементом множества B, то говорят, что A есть подмножество множества B (или B включает A), и пишут $A \subset B$. В частности, любое множество является подмножеством самого себя, т. е. $A \subset A$, пустое множество $\emptyset$ считается подмножеством любого множества A, т. е. $\emptyset \subset A$. Если хотят подчеркнуть, что подмножество A множества B не совпадает с B, то пишут $A \subsetneq B$ и говорят, что B строго включает A или, что A является собственным подмножеством множества B.

Например, для указанных выше числовых множеств имеют место строгие включения:

$$\mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}\ ,\ \overline{m, n} \subset \mathbb{Z}.$$

В повседневной практике нам часто приходится получать из одних множеств другие, например объединяя заданные множества, выбирая их общие или, наоборот, необщие элементы и т. д. Для формализации таких способов получения множеств используются различные операции над множествами. Определим четыре операции над множествами: объединение, пересечение, вычитание и декартово произведение, обозначаемые соответственно $\cup, \cap, \setminus, \times$.

Определение 1.1. *Объединением множеств A, B называется множество $A \cup B$, состоящее из всех тех элементов, которые принадлежат хотя бы одному из множеств A, B:*

$$A \cup B = \{m : m \in A \text{ или } m \in B\}.$$

Определение 1.2. *Пересечением множеств A, B называется множество $A \cap B$, состоящее из всех тех элементов, которые содержатся в обоих множествах A, B:*

$$A \cap B = \{m : m \in A \ , \ m \in B\}.$$

Заметим, что пересечение двух множеств может оказаться пустым множеством. В этом случае исходные множества называются непересекающимися.

Определение 1.3. *Разностью множеств A, B называется множество $A \setminus B$, состоящее из всех элементов множества A, не содержащихся в B:*

$$A \setminus B = \{m : m \in A \ , \ m \notin B\}.$$

Определение 1.4. *Декартовым произведением множеств A, B называется множество $A \times B$, состоящее из всевозможных упорядоченных пар вида (a, b), где $a \in A$, $b \in B$:*

$$A \times B = \{(a, b) : a \in A \ , \ b \in B\}.$$

Пример 1.1. Для множеств $A = \{1, 2, 3\}$, $B = \{2, 4\}$ имеем

$$A \cup B = \{1, 2, 3, 4\} \ , \ A \cap B = \{2\} \ , \ A \setminus B = \{1, 3\} \ , \ B \setminus A = \{4\},$$

$$A \times B = \{(1, 2), (1, 4), (2, 2), (2, 4), (3, 2), (3, 4)\},$$

$$B \times A = \{(2, 1), (2, 2), (2, 3), (4, 1), (4, 2), (4, 3)\}.$$

Операции над множествами обладают многими интересными свойствами. Приведем некоторые из них, обозначая буквами A, B, C произвольные множества.

Коммутативность операций $\cup$, $\cap$:

1) $A \cup B = B \cup A$;

2) $A \cap B = B \cap A$.

Ассоциативность операций $\cup$, $\cap$:

3) $(A \cup B) \cup C = A \cup (B \cup C)$;

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

Идемпотентность операций $\cap$, $\cup$:

5) $A \cup A = A$;

6) $A \cap A = A$.

Законы поглощения:

7) $A \cup (A \cap B) = A$;

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

Законы правой дистрибутивности операций $\cup, \cap, \setminus, \times$ относительно операций $\cap, \cup$:

9) $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$;

10) $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$;

11) $(A \cup B) \setminus C = (A \setminus C) \cup (B \setminus C)$;

12) $(A \cap B) \setminus C = (A \setminus C) \cap (B \setminus C)$;

13) $(A \cup B) \times C = (A \times C) \cup (B \times C)$;

14) $(A \cap B) \times C = (A \times C) \cap (B \times C)$.

И другие законы:

15) $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C)$;

16) $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C)$;

17) $A \setminus (B \cup C) = (A \setminus B) \setminus C$;

18) $A \setminus (A \cup B) = \emptyset$;

19) $A \setminus (A \cap B) = A \setminus B$;

20) $A \cup \emptyset = A$;

21) $A \cap \emptyset = \emptyset$.

Справедливость выписанных равенств следует непосредственно из определений операций над множествами или легко проверяется. Так, проверку равенства 12 можно записать в виде последовательности утверждений, эквивалентных на основании определений 1.2, 1.3 (при этом эквивалентность утверждений обозначается знаком $\Longleftrightarrow$):

$$a \in (A \cap B) \setminus C \Longleftrightarrow a \in A \cap B,\ a \notin C \Longleftrightarrow$$
$$\Longleftrightarrow a \in A,\ a \in B,\ a \notin C \Longleftrightarrow a \in A \setminus C,\ a \in B \setminus C \Longleftrightarrow$$
$$\Longleftrightarrow a \in (A \setminus C) \cap (B \setminus C).$$

По аналогии с объединением и пересечением двух множеств можно ввести объединение и пересечение произвольного семейства множеств A_i, $i \in I$:

$$\cap_{i \in I} A_i = \{a : a \in A_i \text{ для всех } i \in I\},$$

$$\cup_{i \in I} A_i = \{a : a \in A_i \text{ хотя бы для одного } i \in I\}.$$

В частности, если $I = \{1, 2, ..., n\}$, то вместо

$$\cap_{i \in I} A_i \text{ и } \cup_{i \in I} A_i$$

пишут соответственно

$$\cap_{i=1}^{n} A_i \text{ и } \cup_{i=1}^{n} A_i$$

или

$$A_1 \cap ... \cap A_n \text{ и } A_1 \cup ... \cup A_n.$$

Если имеет место равенство $A = \cup_{i \in I} A_i$, то говорят, что множество A разложено в объединение своих подмножеств A_i, $i \in I$. Если при этом $A_i \neq \emptyset$ и $A_i \cap A_j = \emptyset$ при $i \neq j$, то говорят о разложении A в объединение непересекающихся подмножеств или о *разбиении* множества A.

Пусть A — произвольное множество и n — натуральное число. Декартовой n-й степенью множества A называется множество, обозначаемое через A^n и состоящее из всевозможных наборов длины n элементов из A:

$$A^n = \{(a_1, ..., a_n) : a_i \in A \ , \ i \in \overline{1, n}\}.$$

При этом два набора $(a_1, ..., a_n)$ и $(b_1, ..., b_n)$ элементов из A считаются равными, если $a_i = b_i$ для всех $i \in \overline{1, n}$. Иначе говоря, для набора существенным является не только состав элементов, но и порядок их расположения. Подчеркнем еще, что в отличие от множества, в котором все элементы всегда считаются различными, набор может содержать и одинаковые элементы.

1.2. ОТОБРАЖЕНИЯ МНОЖЕСТВ

Определение 1.5. *Пусть A, B — произвольные множества. Отображением множества A в множество B называется всякое правило f, по которому каждому элементу множества A сопоставляется вполне определенный (единственный) элемент множества B.*

Тот факт, что f есть отображение A в B, кратко записывают в виде

$$f : A \longrightarrow B.$$

Если при этом элементу a из A сопоставлен элемент b из B, то b называется образом элемента a, а a — прообразом элемента b при отображении f, что записывается в виде

$$f(a) = b \text{ или } af = b.$$

Из определения отображения f следует, что у каждого элемента a из A образ единственный, однако для элемента $b \in B$ прообразов может быть и много, а может и вообще не быть. Множество всех прообразов элемента b из B называется его полным прообразом и обозначается через $f^{-1}(b)$. Таким образом,

$$f^{-1}(b) = \{a : a \in A, f(a) = b\}$$

или, несколько короче,

$$f^{-1}(b) = \{a \in A : f(a) = b\}.$$

Естественным путем определяются образ $f(A_1)$ подмножества A_1 из A и полный прообраз $f^{-1}(B_1)$ подмножества B_1 из B при отображении f:

$$f(A_1) = \cup_{a \in A_1}\{f(a)\}\ ,\ f^{-1}(B_1) = \{\cup_{b \in B_1}\{f^{-1}(b)\}.$$

Пример 1.2. Пусть $A = \{0, 1, 2, 3, 4, 5, 6\}$ и f — отображение A в A, сопоставляющее каждому элементу a из A остаток от деления a на число 4. Тогда имеем

$$f(0) = 0\ ,\ f(1) = 1\ ,\ f(2) = 2\ ,\ f(3) = 3\ ,\ f(4) = 0\ ,$$

$$f(5) = 1\ ,\ f(6) = 2\ ,\ f^{-1}(0) = \{0, 4\}\ ,$$

$$f^{-1}(1) = \{1, 5\}\ ,\ f^{-1}(2) = \{2, 6\}\ ,\ f^{-1}(3) = \{3\}\ ,$$

$$f^{-1}(4) = f^{-1}(5) = \emptyset\ ,$$

$$f(\{0, 4, 5\}) = \{0, 1\},\ f^{-1}(\{0, 1\}) = \{0, 1, 4, 5\}.$$

Если f — отображение множества A в множество B, то множество всех пар вида (a, b), где $a \in A$, $b \in f(a)$, называется графиком отображения f. Так, графиком отображения из предыдущего примера является множество пар:

$$\{(0, 0), (1, 1), (2, 2), (3, 3), (4, 0), (5, 1), (6, 2)\}.$$

Заметим, что приведенное выше определение отображения не является математически строгим, поскольку в нем используется неопределенный термин «правило». Для строгого определения понятия отображения используется подход через график. Отображение A в B отождествляют с его графиком, а график уже определяется строго, как подмножество M декартова произведения $A \times B$, удовлетворяющее условию: для каждого элемента $a \in A$ в M существует единственная пара с первым элементом a. При таком определении равенство $f(a) = b$ означает наличие в множестве M пары (a, b).

Теорема 1.1. *Для любого отображения $f : A \longrightarrow B$ и любых подмножеств $A_1, A_2 \subset A$, $B_1, B_2 \subset B$ имеют место соотношения:*

1) $f(f^{-1}(B_1)) \subset B_1$;
2) $f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2)$;
3) $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$;
4) $f(A_1 \setminus A_2) \supset f(A_1) \setminus f(A_2)$;
5) $f^{-1}(f(A_1)) \supset A_1$;
6) $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$;
7) $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$;
8) $f^{-1}(B_1 \setminus B_2) = f^{-1}(B_1) \setminus f^{-1}(B_2)$.

□ Предварительно заметим, что по определению образа и полного прообраза подмножества X имеем

$$b \in f(X) \Longleftrightarrow b = f(x) \text{ для некоторого } x \in X, \tag{1.1}$$

$$a \in f^{-1}(X) \Longleftrightarrow f(a) \in X. \tag{1.2}$$

Теперь, пользуясь утверждениями (1.1), (1.2), докажем для примера утверждения 3 и 8.

1) Пусть b — любой элемент из B. Тогда:

$$b \in f(A_1 \cup A_2) \Longleftrightarrow b = f(a) \text{ для некоторого } a \in A_1 \cup A_2 \Longleftrightarrow$$

$$\Longleftrightarrow b = f(a) \text{ для некоторого } a \text{ из } A_1 \text{ или } A_2 \Longleftrightarrow$$

$$\Longleftrightarrow b \in f(A_1) \text{ или } b \in f(A_2) \Longleftrightarrow b \in f(A_1) \cup f(A_2).$$

Тем самым утверждение 3 доказано.

2) Пусть a — любой элемент из A. Тогда:

$$a \in f^{-1}(B_1 \setminus B_2) \Longleftrightarrow f(a) \in B_1 \setminus B_2 \Longleftrightarrow$$
$$\Longleftrightarrow f(a) \in B_1,\ f(a) \notin B_2 \Longleftrightarrow$$
$$\Longleftrightarrow a \in f^{-1}(B_1),\ a \notin f^{-1}(B_2) \Longleftrightarrow$$
$$\Longleftrightarrow a \in f^{-1}(B_1) \setminus f^{-1}(B_2).$$

Утверждение 8 доказано. Остальные утверждения доказываются аналогично. □

Покажите на примерах, что включения 1, 2, 4, 5 могут быть строгими.

В зависимости от свойств образов и прообразов различают отображения сюръективные, инъективные и биективные.

Определение 1.6. *Отображение $f : A \longrightarrow B$ называется сюръективным, если $f(a) = B$, т. е. в каждый элемент из B отображается хотя бы один элемент из A, или $f^{-1}(b) \neq \emptyset$ при любом $b \in B$.*

Определение 1.7. *Отображение $f : A \longrightarrow B$ называется инъективным, если оно разные элементы множества A отображает в разные элементы множества B, т. е.*

$$f(a_1) = f(a_2) \Longrightarrow a_1 = a_2,$$

или $f^{-1}(b)$ является либо пустым, либо одноэлементным множеством при любом $b \in B$. Инъективные отображения называются также вложениями.

Определение 1.8. *Отображение $f : A \longrightarrow B$ называется биективным, или взаимно однозначным, отображением A на B, если оно сюръективно и инъективно, т. е. если $f^{-1}(b)$ есть одноэлементное множество при любом $b \in B$.*

Пример 1.3. Определим отображение $f_1 : \mathbb{Z} \longrightarrow \mathbb{N}_0$, положив для $a \in \mathbb{Z}$

$$f_1(a) = |a|,$$

где $|a|$ — абсолютная величина числа a.

Очевидно, что f_1 — сюръективное, но не инъективное отображение.

Пример 1.4. Отображение $f_2 : \mathbb{Z} \longrightarrow \mathbb{Z}$, определенное правилом

$$f_2(a) = \begin{cases} 2a \text{ , если } a \geq 0 \\ 2|a| - 1 \text{ , если } a < 0 \end{cases},$$

является инъективным отображением. Оно не является биективным, поскольку $f_2(\mathbb{Z}) = \mathbb{N}_0 \neq \mathbb{Z}$. Однако если таким же образом определить отображение $\mathbb{Z}$ в $\mathbb{N}_0$, то получим биективное отображение.

Определение 1.9. *Отображение множества A в себя называется преобразованием множества A. Биективное преобразование множества A называется подстановкой множества A или подстановкой на A.*

Пример 1.5. Примером подстановки на множестве целых чисел может служить отображение $f : \mathbb{Z} \longrightarrow \mathbb{Z}$, определенное равенством

$$f(a) = a + 1 \text{ , } a \in \mathbb{Z}.$$

Заметим еще, что отображение f множества A в B называют также функцией, заданной на множестве A со значениями в множестве B. При этом элемент $f(a)$ называют значением функции f в точке a. Само множество A называют областью определения функции f, а множество $f(A) \subset B$ — областью значений функции f.

Функцию $f : A \longrightarrow B$ зачастую трактуют как переменную величину y, принимающую значения из B и так зависящую от переменной величины x, принимающей значения из A, что каждому значению a переменной величины x соответствует вполне определенное значение $f(a)$ величины y. При этом пишут $y = f(x)$ и вместо «функция f» говорят «функция $f(x)$».

В теории и на практике часто приходится осуществлять последовательно различные отображения множеств. В связи с этим дадим

Определение 1.10. *Композицией отображений ϕ и f, где $f : A \longrightarrow B$, $\phi : B \longrightarrow C$, называется отображение $\phi \circ f$:*

$A \longrightarrow C$, *определенное условием*

$$(\phi \circ f)(a) = \phi(f(a)) \ , \ a \in A. \tag{1.3}$$

То же самое отображение называют еще произведением отображений f и ϕ и обозначают в виде $f \cdot \phi$ или $f\phi$. Таким образом,

$$(f\phi)(a) = \phi(f(a)) \ , \ a \in A. \tag{1.4}$$

Каждая из записей (1.3), (1.4) имеет свои достоинства. В записи (1.3) первым применяется то преобразование, которое записано ближе к элементу. В (1.4) первым применяется то преобразование, которое написано первым. Если вместо $f(a)$ писать af, то равенства (1.3), (1.4) запишутся в виде

$$a(\phi \circ f) = (af)(\phi) \ , \ a(f\phi) = (af)\phi.$$

Отсюда видно, что при записи образа в виде $f(a)$ удобнее пользоваться композицией. Если же вместо $f(a)$ писать af, то удобнее пользоваться произведением отображений.

Отметим некоторые свойства композиции и произведения отображений.

1. Если

$$f_1 : A \longrightarrow B \ , \ f_2 : B \longrightarrow C \ , \ f_3 : C \longrightarrow D,$$

то

$$(f_3 \circ f_2) \circ f_1 = f_3 \circ (f_2 \circ f_1). \tag{1.5}$$

Для доказательства найдем образ элемента a из A при действии отображений, записанных в левой и правой частях равенства (1.5). Из (1.3) имеем

$$((f_3 \circ f_2) \circ f_1)(a) = (f_3 \circ f_2)(f_1(a)) = f_3(f_2(f_1(a))),$$

$$(f_3 \circ (f_2 \circ f_1))(a) = f_3((f_2 \circ f_1)(a)) = f_3(f_2(f_1(a))).$$

Отсюда и следует (1.5).

С использованием операции умножения равенство (1.5) запишется в виде

$$f_1(f_2 f_3) = (f_1 f_2) f_3.$$

2. Если отображения $f_1 : A \longrightarrow B$, $f_2 : B \longrightarrow C$ сюръективны, инъективны или биективны, то соответственно таким же будет и отображение $f_2 \circ f_1$, $f_1 f_2$.

Докажите это в качестве упражнения.

Заметим, что в общем случае из биективности $f_1 f_2$ не следует биективность f_1, f_2; из сюръективности $f_1 f_2$ следует сюръективность лишь f_2, а из инъективности $f_1 f_2$ следует инъективность лишь f_1. Приведите соответствующие примеры.

3. Если для отображений $f : A \longrightarrow B$, $\phi : B \longrightarrow C$, $\psi : B \longrightarrow C$ выполняется равенство

$$\phi \circ f = \psi \circ f \tag{1.6}$$

и отображение f сюръективно, то $\phi = \psi$.

Действительно, условие (1.6) означает, что для любого элемента $a \in A$

$$\phi(f(a)) = \psi(f(a)). \tag{1.7}$$

Так как f сюръективно, то для любого $b \in B$ существует $a \in \in A$, такое что $f(a) = b$. Отсюда и из равенства (1.7) следует, что $\phi(b) = \psi(b)$ для любого $b \in B$, т. е. $\phi = \psi$.

Приведите пример, показывающий, что условие сюръективности для f существенно.

4. Если для отображений $f : A \longrightarrow B$, $\phi : A \longrightarrow B$, $\psi : B \longrightarrow C$ выполняется равенство

$$\psi \circ f = \psi \circ \phi \tag{1.8}$$

и отображение ψ инъективно, то $\phi = f$.

Доказательство этого свойства и пример, показывающий, что условие инъективности ψ существенно, приведите в качестве упражнений.

5. Если f и ϕ — преобразования множества A, то их композиция $f \circ \phi$ (и произведение $f\phi$) также является преобразованием множества A.

1.3. ОТНОШЕНИЯ НА МНОЖЕСТВЕ. ОТНОШЕНИЯ ЭКВИВАЛЕНТНОСТИ И ПОРЯДКА

Человеку в своей деятельности чаще всего приходится иметь дело с такими множествами, между элементами которых существуют определенные связи, или отношения. Так, в коллективах людей имеются отношения родства, товарищества, соседства, старшинства и т. д. На множестве целых чисел можно рассматривать отношения «равенства», «делимости», «больше», «меньше», «между» и т. д.

Прежде чем дать определение понятия отношения, рассмотрим на примерах, чем определяются некоторые знакомые нам отношения на множестве чисел.

Пример 1.6. Пусть

$$M = \{1, 2, 3, 4, 5\}.$$

1. Отношение делимости на M можно задать перечислением всех пар чисел из M, в которых первое число делит второе. В данном случае такими парами будут

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

2. Отношение «меньше» также можно задать перечислением пар элементов, в которых первое число меньше второго. Получим множество пар:

$$\{(1,2), (1,3), (1,4), (1,5), (2,3), (2,4), (2,5), (3,4), (3,5), (4,5)\}.$$

3. Отношение равенства задается множеством пар

$$\{(1,1), (2,2), (3,3), (4,4), (5,5)\}.$$

4. Для задания отношения «меньше или равно» необходимо к только что перечисленным парам добавить пары из примера 2.

5. Отношение непосредственного следования (или «быть больше на единицу») задается множеством пар:

$$\{(1,2), (2,3), (3,4), (4,5)\}.$$

6. Отношение «a является квадратом b» задается множеством, состоящим из двух пар $(1,1)$, $(4,2)$.

7. Отношение «быть на 10 единиц больше» задается пустым множеством пар, поскольку в множестве M нет таких чисел a, b, что a на 10 единиц больше, чем b.

В пунктах 1–7 все отношения учитывали лишь связи между любыми двумя элементами множества. Такие отношения называются бинарными. Рассмотрим на том же множестве примеры отношений, в которых учитываются связи в тройках элементов.

8. Отношение «строго между» задается множеством троек:

$$\{(1,2,3),(1,2,4),(1,2,5),(1,3,4),(1,3,5),\\ (1,4,5),(2,3,4),(2,3,5),(2,4,5),(3,4,5)\}.$$

9. Отношение между тремя элементами a, b, c, означающее, что c есть сумма элементов a, b, задается на M следующим множеством троек:

$$\{(1,1,2),(1,2,3),(1,3,4),(1,4,5),(2,1,3),\\ (2,2,4),(2,3,5),(3,1,4),(3,2,5),(4,1,5)\}.$$

Отношения из пунктов 8, 9 называются тернарными.

После рассмотрения приведенных примеров становится естественным следующее

Определение 1.11. *Отношением арности n, или n-арным отношением, на множестве M называется произвольное подмножество R множества M^n:*

$$R \subset M^n.$$

При этом, если $(a_1,...,a_n) \in R$ или $(a_1,...,a_n) \notin R$, то говорят соответственно, что элементы a_1,..., a_n находятся или не находятся в отношении R.

Ниже мы ограничимся в основном рассмотрением лишь бинарных отношений (т. е. случаем $n = 2$). В этом случае вместо $(a,b) \in R$ чаще пишут aRb, например, $a \leq b$, $a > b$, $a = b$, $a|b$ (a делит b), $a \parallel b$ (a параллельна b), $a \perp b$ (a перпендикулярна

b) и т. д. Если же элементы a, b не находятся в отношении R, то будем писать $a\overline{R}b$ (при конкретных R используются разные формы записи, например, $a \not> b$ (a не больше b), $a \nmid b$ (a не делит b) и т. п.).

В зависимости от свойств множества пар, определяющего бинарное отношение, выделяют различные классы отношений.

Определение 1.12. *Отношение R на множестве M называется рефлексивным, симметричным, антисимметричным, транзитивным, связным, если выполняются соответственно условия (для любых элементов $a, b \in M$):*

aRa,
aRb влечет bRa,
aRb и bRa влечет $a = b$,
aRb и bRc влечет aRc,
aRb или bRa.

Например, отношение делимости на множестве $\mathbb{N}$ рефлексивно, антисимметрично, транзитивно и не связно; отношение параллельности на множестве прямых плоскости симметрично, но не рефлексивно, не транзитивно и не связно; отношение «a является квадратом b» на множестве $\mathbb{N}$ транзитивно, антисимметрично, не рефлексивно и не связно; отношение $\leq$ на $\mathbb{N}$ рефлексивно, антисимметрично, транзитивно и связно.

Можно говорить и о многих других свойствах бинарных отношений, однако мы ограничимся только указанными, поскольку через них определяются два наиболее важных для всей математики бинарных отношения — отношения эквивалентности и частичного порядка.

Определение 1.13. *Бинарное отношение R на множестве M называется отношением эквивалентности, если оно рефлексивно, симметрично и транзитивно.*

Значение отношений эквивалентности определяется следующей теоремой.

Теорема 1.2. *Пусть R — бинарное отношение на множестве M и $[a]_R$ — класс всех элементов из M, которые на-*

ходятся в отношении R с элементом $a \in M$, т. е.

$$[a]_R = \{x \in M : xRa\}.$$

Отношение R является отношением эквивалентности на множестве M тогда и только тогда, когда классы $[a]_R$ обладают следующими свойствами:

1) любые два элемента одного класса находятся в отношении R;

2) любые два элемента из разных классов не находятся в отношении R;

3) объединение всех классов $[a]_R$ равно M.

□ Пусть отношение R удовлетворяет условиям 1–3. Из условия 3 следует, что любой элемент a из M содержится хотя бы в одном классе, тогда по условию 1 aRa, т. е. R рефлексивно. Если $a, b \in M$ и aRb, то из условий 2–3 следует, что a и b содержатся в одном классе, тогда по условию 1 bRa, т. е. R — симметрично. Аналогично доказывается транзитивность отношения R.

Обратно, пусть R — отношение эквивалентности, т. е. R рефлексивно, симметрично и транзитивно. Если b, c — элементы какого-либо одного класса $[a]_R$, то имеем bRa, cRa. Тогда по свойству симметричности получим aRc. Применяя теперь свойство транзитивности к соотношениям bRa, aRc, получим bRc. Таким образом, любые два элемента одного класса находятся в отношении R.

Возьмем теперь элементы b, c из разных классов и докажем, что они не находятся в отношении R. Допустим, что

$$b \in [a_1]_R \ , \ c \in [a_2]_R \ , \ [a_1]_R \neq [a_2]_R \text{ и } bRc.$$

Так как $[a_1]_R \neq [a_2]_R$, то существует элемент a, такой что $a \in [a_1]_R \setminus [a_2]_R$ или $a \in [a_2]_R \setminus [a_1]_R$. В первом случае имеем

$$bRa_1 \ , \ cRa_2 \ , \ bRc \ , \ aRa_1 \ , \ a\overline{R}a_2.$$

В силу симметричности отношения R из bRa_1 следует a_1Rb. Теперь из соотношений

$$aRa_1\ ,\ a_1Rb\ ,\ bRc\ ,\ cRa_2$$

по свойству транзитивности получим соотношение aRa_2, противоречащее условию. Аналогично придем к противоречию и во втором случае. Таким образом, элементы из разных классов не находятся в отношении R. Осталось заметить, что

$$\cup_{a\in M}[a]_R = M,$$

поскольку $a \in [a]_R$ в силу рефлексивности отношения R. $\square$

Если R — отношение эквивалентности на множестве M, то классы $[a]_R$ называются классами эквивалентности, соответствующими отношению R. Из доказанной теоремы следует, что два класса или совпадают, или не пересекаются. Тем самым мы получаем разложение множества M в объединение непересекающихся непустых классов, т. е. разбиение множества M. Про это разбиение говорят, что оно индуцируется отношением R. Таким образом, каждое отношение эквивалентности на множестве M индуцирует разбиение множества M.

Обратно, если задано произвольное разбиение множества M, то, определив на M отношение R условием:

$$aRb \Longleftrightarrow a, b \text{ содержатся в одном классе разбиения,}$$

мы получим отношение эквивалентности, индуцирующее исходное разбиение.

Таким образом, между всеми отношениями эквивалентности на множестве M и всеми разбиениями множества M существует естественное взаимно однозначное соответствие (при котором отношению эквивалентности соответствует индуцируемое им разбиение).

Указанным соответствием и объясняется широкое использование отношений эквивалентности в научной и практической деятельности человека. Всякий раз, когда хотят какие-то объекты классифицировать по какому-либо признаку, то, по-существу, вводят отношение эквивалентности на рассматриваемом множестве. Классификация элементов множества облегчает задачу его изучения, поскольку в некоторых случаях

позволяет вместо всех элементов изучать лишь совокупность отдельных представителей из классов эквивалентности. Более того, эти представители могут быть выбраны наиболее простыми с точки зрения интересующих нас свойств. Например, из всех эквивалентных между собой систем уравнений достаточно исследовать и решить лишь одну, в некотором смысле наиболее простую, систему.

Определение 1.14. *Бинарное отношение R на множестве M называется отношением частичного порядка (или просто отношением порядка), если оно рефлексивно, транзитивно и антисимметрично. Связное отношение частичного порядка называется линейным порядком.*

Естественными примерами отношений частичного порядка являются отношение $\leq$ («меньше или равно») на произвольном множестве действительных чисел и отношение $\subset$ теоретико-множественного включения на множестве всех подмножеств любого фиксированного множества M. При этом в первом случае мы имеем отношение линейного порядка, во втором нет, если M не одноэлементно.

Отношение порядка на множестве обычно вводится в том случае, когда хотят установить определенную иерархию между его элементами. Оно обозначается чаще всего значком $\leq$. Если $a \leq b$ и $a \neq b$, то пишут $a < b$ и говорят «a меньше b».

Множество M с заданным на нем отношением частичного или линейного порядка $\leq$ называется соответственно частично или линейно упорядоченным множеством и обозначается в виде $(M, \leq)$.

Для каждого частично упорядоченного множества можно определить его геометрическое изображение, или его диаграмму. При построении диаграммы частично упорядоченного множества $(M, \leq)$ различные элементы из M отождествляются с различными точками плоскости так, что:

1) точка a лежит левее (или ниже) точки b, если $a < b$;

2) точка a соединяется отрезком с отличной от нее точкой b, если $a \leq b$ и не существует точки c, отличной от a, b, удовлетворяющей условию $a \leq c \leq b$ (в этом случае говорят,

что b непосредственно следует за a или a непосредственно предшествует b).

Начертите в качестве упражнения диаграмму множества всех подмножеств множества $A = \{1, 2, 3\}$, упорядоченного отношением включения.

В диаграмме линейно упорядоченного множества все точки расположены на одной прямой линии. Отсюда происходит и сам термин «отношение линейного порядка».

Для частично упорядоченного множества $(M, \leq)$ естественным образом определяются понятия минимального и максимального элементов. Элемент a называется минимальным (максимальным), если для любого элемента $b \in M$ из $b \leq a$ (соответственно из $a \leq b$) следует $a = b$.

Заметим, что частично упорядоченное множество может иметь один или несколько минимальных (максимальных) элементов и может совсем не иметь их. Так, множество $\mathbb{Z}$ с обычным отношением «меньше или равно» не имеет ни минимального, ни максимального элемента, множество $\mathbb{N}$, упорядоченное таким же отношением, имеет единственный минимальный элемент 1 и не имеет максимальных элементов. Множество $M = \overline{2,6}$, упорядоченное отношением делимости, имеет три минимальных элемента 2, 3, 5 и три максимальных элемента 4, 5, 6. Интересно отметить, что в данном примере число 5 является и минимальным, и максимальным элементом.

В математике особо важную роль играют частично упорядоченные множества, в которых любое подмножество имеет минимальный элемент. Про такие частично упорядоченные множества говорят, что они удовлетворяют условию минимальности. В частности, линейно упорядоченные множества с условием минимальности называются вполне упорядоченными.

Так, множество чисел M с обычным отношением порядка «меньше или равно» удовлетворяет условию минимальности при $M = \mathbb{N}$ и не удовлетворяет ему при $M = \mathbb{Z}$. Об особой роли частично упорядоченных множеств с условием минимальности будет сказано ниже (см. 1.6).

1.4. МНОЖЕСТВА С ОПЕРАЦИЯМИ

Определение 1.15. *Операцией арности n, или n-арной операцией, на множестве M при $n > 0$ называется произвольное отображение*

$$f : M^n \longrightarrow M.$$

При этом образ элемента $(a_1, ..., a_n)$ из M^n при отображении f называется результатом применения операции f к элементам $a_1, ..., a_n$ из M и обозначается в виде

$$f(a_1, ..., a_n).$$

Подчеркнем, что согласно определению операция применима к любому набору $(a_1, ..., a_n)$ из M^n и результат всегда принадлежит множеству M.

Для общности вводят также понятие и нуль-арной операции. Результатом нуль-арной операции по определению считается некоторый фиксированный элемент множества M. В связи с этим нуль-арную операцию и обозначают, как правило, тем же символом, что и элемент множества M, являющийся значением этой операции.

При $n = 1, 2, 3$ n-арные операции называются соответственно унарными, бинарными и тернарными.

Из сравнения определения 1.15 с определением 1.9 видно, что понятие унарной операции совпадает с понятием преобразования множества. Практически наиболее интересными являются бинарные операции. Если f — бинарная операция на множестве M, то вместо $f(a_1, a_2)$ чаще пишут $a_1 f a_2$. Во многих конкретных примерах вместо f используются символы $+, \cdot, -, \circ, \cup, \cap$ и др.

Приведем примеры бинарных операций.

Пример 1.7. Бинарными операциями являются хорошо знакомые со средней школы операции сложения и умножения на множествах $\mathbb{N}_0$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, а также операция вычитания на множествах $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.

Заметим, что вычитание не является операцией на множестве $\mathbb{N}$, поскольку разность натуральных чисел может и не

быть натуральным числом. То же можно сказать и о делении на любом из указанных множеств, поскольку невозможно деление, например, на нуль.

Пример 1.8. Пусть $(M, \leq)$ есть произвольное линейно упорядоченное множество, f — отображение M^2 в M, определенное условием:

$$f(a, b) = \max\{a, b\} \text{ для любых } a, b \in M,$$

где $\max\{a, b\}$ — максимальный элемент частично упорядоченного множества, состоящего из двух элементов a, b относительно порядка, индуцированного отношением $\leq$ на M. Так как $(M, \leq)$ — линейно упорядоченное множество, то при любых $a, b \in M$ $\max\{a, b\}$ есть вполне определенный элемент из M и потому f есть бинарная операция на M. Аналогично определяется операция взятия минимального элемента $\min\{a, b\}$ на M.

Пример 1.9. Рассмотрим множество $P(M)$ всех подмножеств фиксированного множества M. Так как пересечение и объединение любых двух подмножеств из M являются вполне определенными подмножествами из M, то операции пересечения и объединения пар подмножеств из M являются бинарными операциями на множестве $P(M)$. Они обозначаются соответственно $\cap$ и $\cup$.

Заметим, что на $P(M)$ обычно рассматривается еще унарная операция $'$ взятия дополнения: $A' = M \setminus A$.

Пример 1.10. Пусть $\Pi(M)$ есть множество всех преобразований непустого множества M. В 1.2 было определено понятие композиции $f \circ \phi$ любых двух преобразований множества M, причем $f \circ \phi$ есть также преобразование множества M. Следовательно, композиция преобразований есть бинарная операция на множестве $\Pi(M)$. То же самое можно сказать и о произведении преобразований множества M.

Пример 1.11. Обозначим через $\mathcal{B}(M)$ множество всех бинарных отношений на M. Определим понятие произведения бинарных отношений.

Определение 1.16. *Произведением бинарных отношений R_1, R_2 на множестве M называется отношение R на M, в ко-*

тором для любых элементов $a, b \in M$*:*

$$aRb \iff \text{существует } c \in M : aR_1c, cR_2b.$$

Произведение бинарных отношений есть бинарная операция на множестве $\mathcal{B}(M)$.

Даже из только что приведенных примеров видно, сколь разнообразными могут быть бинарные операции на множествах. В связи с этим для облегчения изучения множеств с операциями сами операции классифицируются по их свойствам.

Определение 1.17. *Бинарная операция* $*$ *на множестве* M *называется коммутативной или ассоциативной, если для любых элементов* $a, b, c \in M$ *выполняется соответственно равенство*

$$a * b = b * a,$$

$$(a * b) * c = a * (b * c).$$

Так, операции сложения и умножения чисел на $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ коммутативны и ассоциативны. Операция вычитания на множествах чисел $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ не коммутативна и не ассоциативна, операция композиции преобразований множества M ассоциативна, но не коммутативна, если M не одноэлементно (докажите это в качестве упражнения). Выясните, какими из указанных свойств обладает операция умножения бинарных отношений, определенная в примере 1.11.

Так как на одном и том же множестве может быть задано несколько бинарных операций, то могут существовать свойства, связывающие различные операции.

Определение 1.18. *Пусть* $*$ *и* $\circ$ *— две бинарные операции на множестве* M*. Операция* $*$ *называется лево- или праводистрибутивной относительно операции* $\circ$*, если для любых элементов* a, b, c *из* M *выполняется соответственно равенство*

$$a * (b \circ c) = (a * b) \circ (a * c),$$

$$(a \circ b) * c = (a * c) \circ (b * c).$$

Если выполняются оба эти равенства, то говорят просто о дистрибутивности операции $*$ относительно операции $\circ$.

Так, на числовых множествах $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ операция умножения дистрибутивна относительно операции сложения, но не наоборот. Обе бинарные операции, определенные в примере 1.7 (а также в примере 1.8), дистрибутивны одна относительно другой.

Определение 1.19. *Множество M с заданными на нем операциями и отношениями называется алгебраической системой. При этом M называют основным множеством системы, а множество символов, используемых для обозначения определенных на M операций и отношений, — ее сигнатурой.*

Алгебраическую систему с основным множеством M и с сигнатурой $\sigma = \{f_1, ..., f_k; R_1, ..., R_l\}$, состоящей из символов операций f_i арностей n_i и отношений R_j арностей m_j, обозначают в виде (M, σ) или, подробнее, $(M, f_1, ..., f_k; R_1, ..., R_l)$. При этом набор натуральных чисел

$$< n_1, ..., n_k; m_1, ..., m_l >$$

называют типом системы (M, σ).

Если на алгебраической системе определены только операции или только отношения, то она называется соответственно алгеброй или моделью.

Примером алгебраической системы может служить множество натуральных чисел $\mathbb{N}$ с бинарными операциями сложения и умножения и бинарными отношениями $=$, $<$. Примером модели является любое частично упорядоченное множество.

Широкое распространение в математике и ее приложениях получили такие алгебры, как полугруппы, квазигруппы, группы, кольца, поля, модули, решетки и др. Подробно они изучаются в алгебре. В математической логике особую роль играют так называемые булевы алгебры.

Определение 1.20. *Булевой алгеброй называется множество B с двумя бинарными операциями $\wedge$, $\vee$, одной унарной операцией $'$ и двумя нуль-арными операциями (т. е. выделенными элементами) 0, 1, удовлетворяющее условиям (при любых $a, b, c \in B$):*

1) $(a \wedge b) \wedge c = a \wedge (b \wedge c)$;
2) $(a \vee b) \vee c = a \vee (b \vee c)$;
3) $a \wedge b = b \wedge a$;
4) $a \vee b = b \vee a$;
5) $a \wedge a = a$;
6) $a \vee a = a$;
7) $a \wedge (a \vee b) = a$;
8) $a \vee (a \wedge b) = a$;
9) $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$;
10) $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$;
11) $a \wedge a' = 0$;
12) $a \vee a' = 1$.

Заметим, что система условий 1–12 не является независимой. Докажите, например, что равенства 5, 6 являются следствиями остальных равенств.

Покажите, что из условий 1–12 следуют равенства:

$$a \wedge 0 = 0\ ,\ a \wedge 1 = a\ ,\ a \vee 0 = a\ ,\ a \vee 1 = 1\ ,$$

$$(a \wedge b)' = a' \vee b'\ ,\ (a \vee b)' = a' \wedge b'.$$

Элементы 0 и 1 булевой алгебры B называют соответственно ее нулем и единицей. Иногда их обозначают в виде 0_B и 1_B.

Примером булевой алгебры является множество всех подмножеств произвольного множества M с бинарными операциями пересечения $\cap$ и объединения $\cup$, унарной операцией дополнения и нуль-арными операциями $\emptyset$, M, играющими соответственно роль 0 и 1.

Булеву алгебру образует также множество всех (положительных) делителей числа m, равного произведению различных простых чисел, с операциями

$$a \wedge b = \text{НОД}(a, b),\ a \vee b = \text{НОК}(a, b),\ a' = \frac{m}{a}$$

и с числами 1 и m в роли нуль-арных операций 0 и 1 соответственно. (Проверьте выполнение условий 1–12. Будут ли все они выполняться при любом m?) Ниже нам встретятся и другие примеры булевых алгебр.

Интересно отметить связь булевых алгебр с частично упорядоченными множествами. Положим для элементов a, b произвольной булевой алгебры B:

$$a \leq b \Longleftrightarrow a \vee b = b.$$

Из свойств 6, 4, 2 следует соответственно, что так определенное отношение $\leq$ на B рефлексивно, антисимметрично и транзитивно. В итоге имеем частично упорядоченное множество $(B, \leq)$. Его диаграмма называется диаграммой булевой алгебры B.

При изучении алгебраических систем их обычно классифицируют по типам и по свойствам. Так получаются классы полугрупп, групп, колец, полей, булевых алгебр и т. д. В каждом таком классе алгебраические системы обычно изучаются с точностью до изоморфизма.

Определение 1.21. *Алгебраические системы A, B одной и той же сигнатуры $\sigma = \{f_1, ..., f_k; R_1, ..., R_l\}$ типа $< n_1, ..., n_k; m_1, ..., m_l >$ называются изоморфными, если существует биективное отображение $\phi : A \longrightarrow B$, такое что:*

1) для любой операции $f_i \in \sigma$ и любых элементов $a_1, ..., a_{n_i} \in A$ выполняется равенство

$$\phi(f_i(a_1, ..., a_{n_i})) = f_i(\phi(a_1), ..., \phi(a_{n_i}));$$

2) для любого отношения $R_j \in \sigma$ и любых элементов $a_1, ..., a_{m_j} \in A$

$$R_j(a_1, ..., a_{m_j}) = R_j(\phi(a_1), ..., \phi(a_{m_j})).$$

При этом само отображение ϕ называется изоморфизмом системы A на систему B.

Пример 1.12. Пусть B — булева алгебра всех подмножеств множества $M = \{a_1, ..., a_n\}$, а B_2 — булева алгебра всех делителей числа $m = p_1 p_2 ... p_n$, где $p_1, ..., p_n$ — различные простые числа.

Определим отображение $\phi : B_1 \longrightarrow B_2$, положив

$$\phi(\{a_{i_1}, ..., a_{i_k}\}) = p_{i_1} ... p_{i_k} \text{ и } \phi(\emptyset) = 1.$$

Легко видеть, что

$$\phi(0_{B_1}) = 0_{B_2} \ , \ \phi(1_{B_1}) = 1_{B_2} \ , \ \phi(A') = \phi(A)',$$

а также

$$\phi(A \wedge B) = \phi(A) \wedge \phi(B) \ , \ \phi(A \vee B) = \phi(A) \vee \phi(B)$$

для любых подмножеств A, B множества M. Это и означает, что ϕ есть изоморфизм булевой алгебры A на булеву алгебру B.

Легко видеть, что отношение изоморфизма является отношением эквивалентности на любом множестве алгебраических систем одной сигнатуры, и потому все такие системы разбиваются на классы изоморфных систем. Из определения 1.21 видно, что изоморфные алгебраические системы сигнатуры σ с точки зрения свойств операций и отношений отличаются лишь обозначениями элементов. Отождествив в системах из определения 1.21 элементы a и $\phi(a)$, мы получим одну и ту же систему. Вот почему в алгебре и изучаются алгебраические системы с точностью до изоморфизма. Тем самым достигается существенная экономия наших сил и времени при изучении всего многообразия алгебраических систем.

Заметим, что понятие изоморфизма естественным образом распространяется на алгебраические системы различных, но однотипных сигнатур. При этом необходимо только предварительно установить между операциями, а также между отношениями систем взаимно однозначные соответствия, сохраняющие арности. Так, если операции f_i соответствует операция f'_i, то условие 1 определения 1.21 запишется в виде

$$\phi(f_i(a_1, ..., a_{n_i})) = f'_i(\phi(a_1), ..., \phi(a_{n_i})).$$

В частности, если f_i — бинарная операция $*$, а f'_i — бинарная операция $\circ$, то последнее равенство будет иметь вид

$$\phi(a_1 * a_2) = \phi(a_1) \circ \phi(a_2).$$

Примером изоморфизма алгебр с различными сигнатурами может служить отображение ϕ алгебры $(\mathbb{R}^+, \cdot)$ в алгебру $(\mathbb{R}, +)$, определяемое для любого $x \in R$ равенством $\phi(x) = \log_a x$. Условие 1 определения 1.21 гарантируется в этом случае известным свойством логарифмов:

$$\log_a(xy) = \log_a x + \log_a y.$$

Этот пример показывает, что в некоторых случаях переход к изоморфной алгебре позволяет существенно упростить вычисления. В этом проявляется еще одна положительная роль понятия изоморфизма.

1.5. АКСИОМАТИЧЕСКОЕ ПОСТРОЕНИЕ СИСТЕМЫ НАТУРАЛЬНЫХ ЧИСЕЛ

Арифметика как наука о числах является одной из древнейших наук и подобно другим наукам возникла из потребностей практической деятельности человека. «Как понятие числа, так и понятие фигуры заимствованы исключительно из внешнего мира, а не возникли в голове из чистого мышления» (Ф. Энгельс, «Анти-Дюринг», 1953, с. 37).

В силу потребностей практики происходило как накопление фактов о свойствах чисел, так и постепенное расширение самого понятия числа. Вслед за натуральными появляются дробные, отрицательные, рациональные, иррациональные и мнимые числа. Укоренившиеся исторически термины «иррациональный» и «мнимый» свидетельствуют о том первоначальном сопротивлении, которое встречало появление новых чисел на каждой стадии развития числовых систем.

Задача аксиоматического построения арифметики возникла значительно позже, чем в геометрии, поскольку арифметические предложения более тесно примыкают к непосредствен-

ному опыту человека и потому на первый взгляд не вызывают потребности в доказательствах.

Только бурное развитие всей математики в XIX в. потребовало строгого аксиоматического построения арифметики. К середине XIX в. появилась геометрия Лобачевского, не укладывающаяся в рамки привычной евклидовой геометрии. Встал вопрос обоснования как геометрии Лобачевского, так и геометрии Евклида. Оказалось, что непротиворечивость той и другой сводится к непротиворечивости арифметики. С другой стороны, логического обоснования арифметики требовало развитие математического анализа, полностью основанного на теории действительных чисел. Вот почему во второй половине XIX в. проблеме обоснования, или аксиоматического построения, арифметики уделили самое серьезное внимание многие крупные математики.

К. Вейерштрассом был указан метод построения целых чисел как классов пар (с содержательной точки зрения разностей) натуральных чисел, а также рациональных чисел, как классов пар (с содержательной точки зрения частных) целых чисел. До него В. Гамильтон построил модель системы комплексных чисел как пар действительных чисел. Наконец, Р. Дедекинд, Г. Кантор и К. Вейерштрасс независимо друг от друга построили различные, но эквивалентные теории действительных чисел, основываясь на теории рациональных чисел. Таким образом, для завершения аксиоматизации арифметики осталось построить систему аксиом для множества натуральных чисел. Эта работа, начатая немецким математиком Г. Грассманом еще в 1861 г., была завершена в 1891 г. итальянским математиком Д. Пеано.

Ниже мы приведем схему построения арифметики натуральных чисел, основанную на аксиоматике Пеано и на определении операций по Грассману. Основными понятиями при этом являются понятия множества, подмножества, элемента множества, а основными отношениями — бинарные отношения равенства, «следовать за», включения (подмножества в множество) и принадлежности (элемента множеству). В этих терминах множество натуральных чисел $\mathbb{N}$ определяется следующей системой аксиом Пеано:

1. Существует единственный элемент множества $\mathbb{N}$, не следующий ни за каким элементом из $\mathbb{N}$ (назовем его единицей и обозначим 1).

2. Для каждого элемента $a \in \mathbb{N}$ существует единственный элемент, следующий за a (будем обозначать его через a').

3. Для каждого элемента $a \in \mathbb{N}$ существует не более одного элемента, за которым следует a, т. е. из $b' = c'$ следует $b = c$.

4. (Аксиома полной математической индукции.) Если M есть подмножество множества $\mathbb{N}$, удовлетворяющее условиям:

а) $1 \in M$,

б) из $a \in M$ следует $a' \in M$ для любого $a \in M$, то $M = \mathbb{N}$.

В приведенном аксиоматическом определении множества $\mathbb{N}$ ничего не говорится о природе его элементов. Она может быть какой угодно, лишь бы совокупность удовлетворяла аксиомам 1–4. Выбирая в качестве $\mathbb{N}$ некоторое конкретное множество, в котором определено некоторое конкретное отношение «следовать за», удовлетворяющее аксиомам 1–4, мы получим модель, или интерпретацию, данной системы аксиом.

Можно доказать, что между любыми двумя моделями системы аксиом 1–4 можно установить взаимно однозначное соответствие, сохраняющее отношение «следовать за». Иначе говоря, любые две модели системы аксиом 1–4 изоморфны относительно отношения «следовать за». В качестве стандартной модели обычно выбирают выработанный в процессе исторического развития ряд символов (натуральных чисел):

$$1, 2, 3, 4, \ldots . \tag{1.9}$$

Используя аксиомы 1–4, можно доказать все известные свойства натуральных чисел. При этом главным средством доказательства является основанный на аксиоме 4 принцип полной математической индукции.

Теорема 1.3. *Пусть $T(n)$ — некоторое утверждение, зависящее от переменной величины n, принимающей все значения из $\mathbb{N}$, причем:*

1) $T(n)$ истинно при $n = 1$;

2) для любого $n \in \mathbb{N}$ из истинности $T(n)$ следует истинность $T(n')$.

Тогда $T(n)$ истинно для всех $n \in \mathbb{N}$.

□ Обозначим через M множество тех натуральных чисел, для которых утверждение $T(n)$ истинно. Тогда из условия теоремы следует, что:

1) $1 \in M$;

2) если $n \in M$, то $n' \in M$.

Отсюда по аксиоме 4 получаем $M = \mathbb{N}$, т. е. утверждение $T(n)$ истинно при любом $n \in \mathbb{N}$. □

На аксиоме индукции основываются не только доказательства, но и определения. Так же, как и в теореме 1.3, можно доказать, что функция $f(n)$ однозначно определена на множестве $\mathbb{N}$, если она определена при $n = 1$ и для любого n ее значение на n' однозначно определяется по ее значению на n.

Определим операции сложения и умножения натуральных чисел.

Определение 1.22. *Операцией сложения натуральных чисел называется отображение*

$$f : \mathbb{N}^2 \longrightarrow \mathbb{N},$$

удовлетворяющее условиям (при любых $a, b \in \mathbb{N}$):

1) $f(a, 1) = a'$;

2) $f(a, b') = f(a, b)'$.

Если вместо $f(a, b)$ условиться писать $a+b$, то условия 1–2 запишутся в следующем виде:

1) $a + 1 = a'$;

2) $a + b' = (a + b)'$.

Определение 1.23. *Операцией умножения натуральных чисел называется отображение*

$$\phi : \mathbb{N}^2 \longrightarrow \mathbb{N},$$

удовлетворяющее условиям (для любых $a, b \in \mathbb{N}$):

1) $\phi(a, 1) = a$;

2) $\phi(a, b') = \phi(a, b) + a$.

Используя вместо $\phi(a,b)$ общепринятое обозначение $a \cdot b$, или ab, условия 1–2 можно записать следующим образом:

1) $a \cdot 1 = a$;

2) $ab' = ab + a$.

Методом полной математической индукции нетрудно доказать, что указанные выше отображения f и ϕ существуют и определяются условиями 1–2 однозначно и что операции сложения и умножения ассоциативны, коммутативны и связаны дистрибутивным законом

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

Для примера докажем ассоциативность сложения

$$(a+b)+c = a+(b+c). \tag{1.10}$$

Пусть $c = 1$. Тогда, используя последовательно условия 1, 2, 1 для сложения, получим

$$(a+b)+1 = (a+b)' = a+b' = a+(b+1).$$

Следовательно, для $c = 1$ равенство (1.10) верно. Допустим, что оно верно для c, и докажем для c'. Используя последовательно условие 2, предположение индукции и снова условие 2, получим

$$\begin{aligned}(a+b)+c' &= ((a+b)+c)' = \\ &= (a+(b+c))' = a+(b+c)' = a+(b+c').\end{aligned}$$

Утверждение доказано.

Теперь на множестве $\mathbb{N}$ можно ввести отношение порядка.

Определение 1.24. *Будем говорить, что натуральное число a меньше натурального числа b, и писать $a < b$, если найдется такое $c \in \mathbb{N}$, что*

$$a + c = b.$$

Теорема 1.4. *Отношение $\leq$ на $\mathbb{N}$ является отношением полного порядка, т. е. множество $(\mathbb{N}, \leq)$ вполне упорядочено.*

□ 1) Сначала индукцией по b очевидным образом устанавливается, что для любых $a, b \in \mathbb{N}$ выполняется неравенство

$$a + b \neq b. \tag{1.11}$$

2) Рефлексивность, антисимметричность и транзитивность отношения $\leq$ следуют непосредственно из его определения, свойств сложения и неравенства (1.11).

3) Докажем условие минимальности. Пусть $M \subset \mathbb{N}$ и $M \neq \emptyset$. Обозначим через M_1 множество всех чисел из $\mathbb{N}$, не превосходящих каждого числа из M. Индукцией по a просто устанавливается, что $1 \leq a$ при любом $a \in \mathbb{N}$. Значит, $1 \in M_1$. Однако $M_1 \neq \mathbb{N}$, поскольку $M \neq \emptyset$, и $a \in M$ влечет $a+1 \notin M_1$. Следовательно, найдется такое $b \in M_1$, что $b + 1 \notin M_1$ (иначе было бы $M_1 = \mathbb{N}$). Покажем, что b — минимальный элемент в M. Действительно, $b \leq m$ для любого $m \in M$, поскольку $b \in$ $\in M_1$ и остается доказать, что $b \in M$. Если $b \notin M$, то $b < m$, и потому $b+1 \leq m$ для любого $m \in M$. Получили противоречие. □

Можно было бы доказать многие интересные свойства, связывающие отношение $<$ с операциями сложения и умножения натуральных чисел.

Например, монотонность сложения:

$$a < b \Longrightarrow a + c < b + c;$$

монотонность умножения:

$$a < b \Longrightarrow ac < bc;$$

свойство, называемое аксиомой Архимеда: для любых $a, b \in \mathbb{N}$ существует такое $c \in \mathbb{N}$, что

$$a < bc;$$

и т. д. Мы не будем на этом останавливаться, отослав интересующихся, к монографии [53].

В заключение сделаем отдельные замечания.

1. В начале данного параграфа говорилось о том, что проблема доказательства непротиворечивости геометрии и математического анализа требовала аксиоматического построения

арифметики натуральных чисел. Однако с построением аксиоматической теории натуральных чисел указанная проблема не была решена, поскольку противоречие в теории может быть следствием не только системы аксиом, но и логических средств, используемых в доказательствах. Вместе с тем проведенная математиками XIX в. работа по анализу и построению арифметики была необходимой для последующего развития всей математики. Выделив простейшие факты (аксиомы), лежащие в основе арифметики, она повысила нашу уверенность в прочности математического фундамента и направила дальнейшие усилия математиков на анализ средств доказательств (см. главу 5).

2. Система аксиом Пеано независима, т. е. для каждой из аксиом 1–4 существует множество с отношением «следовать за», на котором эта аксиома не выполняется при выполнении всех остальных аксиом. В качестве таких множеств можно выбрать соответственно:

$$M_1 = \{a, b, c\} \text{ , где } a' = b, b' = c, c' = a;$$

$$M_2 = \{a, b\} \text{ , где } a' = b;$$

$$M_3 = \{a, b, c, d\} \text{ , где } a' = b, b' = c, c' = d, d' = b;$$

$M_4 = \mathbb{N} \cup \widehat{\mathbb{N}}$, где $\mathbb{N}$ — множество (1.9) с естественным отношением «следовать за», а $\widehat{\mathbb{N}}$ — множество символов вида $\widehat{n}$ для всех $n \in \mathbb{N}$, в котором $(\widehat{n})' = \widehat{n'}$.

Следовательно, ни одна из аксиом 1–4 не является следствием остальных.

3. Используя условие минимальности для линейно упорядоченного множества, можно усилить принцип полной математической индукции.

Теорема 1.5. *Пусть $T(n)$ — утверждение, зависящее от натурального числа n, причем $T(n)$ истинно при $n = 1$ и из его истинности для чисел $1, 2, ..., n$ следует истинность для $n + 1$. Тогда $T(n)$ истинно для любого $n \in \mathbb{N}$.*

□ Допустим, что утверждение $T(n)$ истинно не для всех $n \in \mathbb{N}$, и обозначим через M множество всех тех натуральных чисел n, при которых $T(n)$ неверно. По теореме 1.4 в

M существует минимальное число, которое отлично от 1 и потому следует за некоторым числом n_0. Таким образом, имеем: утверждения $T(1), ..., T(n_0)$ истинны, а $T(n_0')$ — нет. Это противоречит условию теоремы. □

4. Утверждение, аналогичное теореме 1.5, имеет место и точно так же доказывается для любого частично упорядоченного множества с условием минимальности.

Теорема 1.6. *Пусть $(M, \leq)$ — произвольное частично упорядоченное множество с условием минимальности и $T(m)$ — утверждение, зависящее от параметра m, принимающего значения из M. Тогда утверждение $T(m)$ истинно для любого $m \in M$, если:*

1) $T(m)$ истинно для всех минимальных элементов из M;

2) из истинности $T(m)$ для всех элементов $m < n$ следует его истинность для $m = n$.

На этой теореме основан метод доказательства, называемый методом трансфинитной индукции.

1.6. МОЩНОСТЬ МНОЖЕСТВА. КОНЕЧНЫЕ И БЕСКОНЕЧНЫЕ МНОЖЕСТВА

В повседневной практике натуральные числа используются в двух различных смыслах — в порядковом и количественном. В соответствии с этим говорят «первый», «второй», «третий» и т. д., или «один», «два», «три» и т. д.

Построенная в предыдущем параграфе аксиоматическая теория арифметики натуральных чисел отражает лишь порядковый смысл натурального числа и потому называется порядковой теорией. В данном параграфе будет указана схема построения арифметики, отражающая количественный смысл натурального числа.

Количественная характеризация натуральных чисел основывается на сравнении множеств путем установления взаимно однозначного соответствия между их элементами или между элементами одного множества и частью другого. Фактически в данном параграфе будет строиться арифметика так

называемых кардинальных чисел, содержащая в себе как часть арифметику натуральных чисел.

Определение 1.25. *Непустое множество A называется равномощным множеству B, если существует биективное отображение A на B.*

Очевидно, что отношение равномощности множеств рефлексивно, симметрично и транзитивно, и потому все множества разбиваются на непересекающиеся классы равномощных множеств.

Сопоставим с каждым классом равномощных множеств некоторый символ так, чтобы разным классам соответствовали разные символы.

Символ, сопоставленный классу, называется кардинальным числом этого класса (или просто кардинальным числом), а также мощностью любого множества из этого класса. Мощность множества A обозначается в виде $|A|$.

Таким образом, мощность множества отражает лишь то общее, что есть у этого множества со всеми остальными множествами содержащего его класса. Из определения же видно, что эта общность заключается лишь в возможности установить взаимно однозначное соответствие между элементами любых двух множеств из этого класса.

Для некоторых кардинальных чисел введем общепринятые обозначения. Возьмем какой-либо элемент a и построим ряд попарно неравномощных множеств

$$\{a\}\ ,\ \{a,\{a\}\}\ ,\ \{a,\{a\},\{a,\{a\}\}\},\ldots\,. \tag{1.12}$$

Мощности этих множеств обозначаются соответственно символами

$$1, 2, 3, ...,$$

а множество всех этих символов — буквой $\mathbb{N}$. В итоге мы получим некоторое множество кардинальных чисел, называемых натуральными числами. Ниже будет показано, что кардинальные числа не исчерпываются множеством $\mathbb{N}$.

Для того чтобы кардинальные числа как символы несли смысловую нагрузку понятия мощности или количества, необходимо научиться сравнивать их между собой, складывать, умножать и т. п.

Определение 1.26. *Пусть α, β — любые кардинальные числа и $\alpha = |A|$, $\beta = |B|$. Говорят, что число α не превосходит числа β, и пишут $\alpha \leq \beta$, если существует инъективное отображение A в B.*

Легко видеть, что определение корректно, т. е. соотношение $\alpha \leq \beta$ не зависит от выбора представителей A, B из классов. Очевидно также, что отношение $\leq$ рефлексивно и транзитивно. Его антисимметричность следует из классической теоремы Шрёдера–Бернштейна. Мы приведем здесь ее наиболее прозрачное доказательство, принадлежащее Г. Биркгофу и С. Маклейну (см. [23]).

Теорема 1.7. *Если существуют инъективные отображения*

$$f : A \longrightarrow B \ , \ g : B \longrightarrow A,$$

то существует биективное отображение $h : A \longrightarrow B$.

□ Если хотя бы одно из преобразований f, g биективно, то утверждение теоремы верно. Пусть f, g не биективны. Определим отображения $f_1 : f(A) \longrightarrow A$, $g_1 : g(B) \longrightarrow B$, положив для $a \in A$ и $b \in B$:

$$f_1(f(a)) = a \ , \ g_1(g(b)) = b.$$

Заметим, что отображения f_1, g_1 не определены соответственно на множествах $B \setminus f(A)$, $A \setminus g(B)$.

Будем теперь классифицировать элементы из A на три подмножества A_0, A_1, A' следующим образом. Возьмем любой элемент $a \in A$ и будем применять попеременно преобразования g_1 и f_1 до тех пор, пока это можно. Если на некотором шаге получим элемент из $A \setminus g(B)$ или из $B \setminus f(A)$ (тогда процесс обрывается), то отнесем элемент a соответственно в

класс A_0 или A_1. В противном случае отнесем a к A'. В итоге получим разбиение множества A:

$$A = A_0 \cup A_1 \cup A'.$$

Аналогично, действуя на элементы из B попеременно преобразованиями f_1, g_1, получим разбиение множества B:

$$B = B_0 \cup B_1 \cup B'.$$

Очевидно, что

$$f(A_0) = B_1 \ , \ f(A') = B' \ , \ f(A_1) = B_0.$$

Теперь искомое биективное отображение $\phi : A \longrightarrow B$ определяется следующим образом:

$$\phi(a) = \begin{cases} f(a), & \text{если } a \in A_0 \cup A' \\ g_1(a), & \text{если } a \in A_1 \end{cases} .$$

□

Следствие 1. *Введенное определением 1.26 отношение на множестве кардинальных чисел является отношением частичного порядка.*

Известно также, что множество кардинальных чисел с отношением $\leq$ вполне упорядочено. Доказательство можно найти, например, в [37].

Докажем еще, что множество кардинальных чисел не ограничено. Этот факт является следствием следующей теоремы Кантора.

Теорема 1.8. *Если $P(M)$ — множество всех подмножеств произвольного множества M, то*

$$|M| < |P(M)|. \tag{1.13}$$

□ Определим отображение $f : M \longrightarrow P(M)$, положив для $m \in M$ $f(m) = \{m\}$. Очевидно, что f инъективно, и потому $|M| \leq |P(M)|$. Докажем еще, что $|M| \neq |P(M)|$. Допустим,

что $|M| = |P(M)|$, т. е. что существует биективное отображение $\phi : M \longrightarrow P(M)$. Определим множества:

$$A = \{m \in M : m \in \phi(m)\}\ ,\ B = M \setminus A.$$

Так как ϕ биективно, то в M найдется элемент m_0, удовлетворяющий условию

$$\phi(m_0) = B.$$

Выясним, какому из множеств A, B принадлежит m_0. Если $m_0 \in A$, то по определению A имеем $m_0 \in \phi(m_0)$, т. е. $m_0 \in$ $\in B$. Если допустить, что $m_0 \in B$, то $m_0 \in \phi(m_0)$ и потому $m_0 \in A$. А так как $A \cap B = \emptyset$, то в обоих случаях приходим к противоречию. Следовательно, $|M| \neq |P(M)|$, и неравенство (1.13) верно. □

Заметим, что доказательства теорем 1.7 и 1.8 не конструктивны.

Введем теперь некоторые операции над кардинальными числами.

Определение 1.27. *Пусть α, β — произвольные кардинальные числа, $\alpha = |A|$, $\beta = |B|$ и $A \cap B = \emptyset$. Суммой чисел α, β, произведением чисел α, β и β-й степенью числа α называются соответственно кардинальные числа*

$$\alpha + \beta = |A \cup B|\ ,\ \alpha \cdot \beta = |A \times B|\ ,\ \alpha^\beta = |A^B|,$$

где A^B — множество всевозможных отображений B в A.

Исходя из свойств операций над множествами, можно доказать следующие свойства операций над кардинальными числами:

$$(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)\ ;\ (\alpha\beta)\gamma = \alpha(\beta\gamma);$$

$$\alpha + \beta = \beta + \alpha\ ;\ \alpha\beta = \beta\alpha;$$

$$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma\ ;\ (\alpha\beta)^\gamma = \alpha^\gamma\beta^\gamma;$$

$$\alpha^{\beta+\gamma} = \alpha^\beta\alpha^\gamma\ ;\ (\alpha^\beta)^\gamma = \alpha^{\beta\gamma}.$$

Если M — произвольное множество, то, сопоставив каждому его подмножеству N характеристическую функцию

$$f_N(x) = \begin{cases} 1, & \text{если } x \in N \\ 0, & \text{если } x \in M \setminus N \end{cases},$$

получим биективное отображение $P(M)$ на $\{0,1\}^M$. Следовательно, $|P(M)| = 2^{|M|}$, и по теореме Кантора имеем: для любого кардинального числа α выполняется неравенство

$$\alpha < 2^\alpha.$$

Очевидно, что отношение полного порядка на множестве всех кардинальных чисел индуцирует отношение полного порядка на любом его непустом подмножестве и, в частности, на $\mathbb{N}$. Нетрудно показать, что на $\mathbb{N}$ оно совпадает с отношением порядка, установленным в предыдущем параграфе.

Определение 1.28. *Множество $\mathbb{N}$ с отношением $\leq$ называется натуральным рядом, а множество $\overline{1,n} = \{1,2,...,n\}$ с отношением $\leq$ — его начальным отрезком длины n.*

Очевидно, что мощность множества $\overline{1,n}$ равна n и не равна мощности множества $\mathbb{N}$ ни при одном значении n.

Определение 1.29. *Пустое множество и любое множество, равномощное какому-либо начальному отрезку натурального ряда, называются конечными множествами. Остальные множества называются бесконечными. Мощности конечных и бесконечных множеств называются соответственно конечными и бесконечными кардинальными числами.*

Таким образом, конечными кардинальными числами являются все натуральные числа $1,2,3,...$, а также мощность пустого множества, которая обозначается символом 0, называемым нулем.

Определение 1.30. *Все множества, равномощные множеству натуральных чисел $\mathbb{N}$, называются счетными. Счетной называется и их мощность.*

Счетная мощность обозначается символом $\aleph_0$ (алеф-нуль).

Теорема 1.9. *Кардинальное число $\aleph_0$ больше любого натурального числа и является наименьшим среди бесконечных кардинальных чисел.*

□ Первое утверждение теоремы следует из неограниченности натурального ряда. Докажем второе утверждение. Поставим в соответствие числу 1 произвольный элемент из бесконечного множества M. Такой найдется, поскольку $M \neq \emptyset$. Пусть мы уже сопоставили числам $1, 2, ..., n$ попарно различные элементы из M. Так как M бесконечно, то оно не равномощно отрезку $\overline{1, n}$ и потому в M найдется элемент, не сопоставленный ни одному из чисел $1, ..., n$. Сопоставим его числу $n + 1$. Таким образом, каждому натуральному числу мы сопоставим определенный элемент множества M так, что разным числам будут сопоставлены разные элементы. Следовательно, существует инъективное отображение $\mathbb{N}$ в M, т. е. $|\mathbb{N}| \leq |M|$ или $\aleph_0 \leq |M|$. □

Следствие 1. *Любое бесконечное множество равномощно некоторому собственному подмножеству.*

□ Действительно, по теореме 1.9 M содержит счетное подмножество $M_1 = \{m_1, m_2, ...\}$. Построим отображение f множества M в себя, положив: $f(m_i) = m_{i+1}$ для $m_i \in M_1$ и $f(a) = a$ для $a \in M \setminus M_1$. Очевидно, что f — биективное отображение M на $M \setminus \{m_1\}$. □

Таким образом, свойство множества не быть равномощным никакому собственному подмножеству является характеристическим для конечных множеств. Поэтому иногда (следуя Р. Дедекинду) его принимают за определение конечного множества.

Глава 2

АЛГЕБРА ВЫСКАЗЫВАНИЙ

2.1. ОСНОВНЫЕ ЛОГИЧЕСКИЕ ОПЕРАЦИИ И ИХ СВОЙСТВА

В математической логике изучаются высказывания и различные связи между ними. При этом понятие высказывания считается основным, неопределяемым понятием. В качестве пояснения говорят лишь, что высказывание — это утверждение, относительно которого известно, истинно оно или ложно. Предполагается также, что высказывание не может быть и истинным, и ложным. В данном параграфе высказывания будут обозначаться малыми латинскими буквами (возможно, с индексами).

Если высказывание a истинно или ложно, то будем говорить соответственно, что оно имеет значение «И» или «Л», и писать $a \equiv$ И или $a \equiv$ Л. Таким образом, по существу, мы предполагаем, что для любого множества высказываний задано его отображение в двухэлементное множество $\Omega = \{\text{И, Л}\}$.

Высказывания a, b, имеющие одинаковые значения, называются эквивалентными, что обозначается в виде $a \equiv b$.

Очевидно, что отношение $a \equiv b$ является отношением эквивалентности на любом заданном множестве высказываний M, и потому M разбивается на два класса высказываний — на класс истинных и класс ложных высказываний.

В обычной речи мы из определенных высказываний a, b с помощью различных связок можем образовывать новые высказывания, например,

$$<< a \text{ и } b >> \ , \ << a \text{ или } b >> \ ,$$

$<<$ если a, то $b>>$, $<<$ неверно, что $a>>$.

В математической логике эти высказывания обозначаются соответственно в виде

$$a\&b\ ,\ a\vee b\ ,\ a\to b\ ,\ \overline{a}\,(\text{или }\neg a) \tag{2.1}$$

и называются конъюнкцией, дизъюнкцией, импликацией высказываний a, b и отрицанием высказывания a. При этом во избежание двусмысленностей, которые нередко встречаются в обычном разговорном языке, в логике дается точное определение значений высказываний (2.1) в зависимости от значений заданных высказываний a, b.

Определение 2.1. *Для любых высказываний* a, b*:*

$$a\&b\equiv И\iff a\equiv И\ , b\equiv И;$$

$$a\vee b\equiv Л\iff a\equiv Л\ , b\equiv Л;$$

$$a\to b\equiv Л\iff a\equiv И\ , b\equiv Л;$$

$$\overline{a}\equiv И\iff a\equiv Л.$$

Все это можно записать также в виде таблицы 1.1.

Таблица 1.1

a	b	$a\&b$	$a\vee b$	$a\to b$	$\overline{a}$
И	И	И	И	И	Л
И	Л	Л	И	Л	Л
Л	И	Л	И	И	И
Л	Л	Л	Л	И	И

В высказываниях $a\&b$, $a\vee b$ исходные высказывания a и b называются членами, или компонентами, конъюнкции и дизъюнкции соответственно; в импликации $a\to b$ высказывание a называют посылкой, b — заключением. Иногда компоненты конъюнкции и дизъюнкции называют соответственно сомножителями и слагаемыми.

Сделаем два замечания.

1. В обычной речи союз «или» используется в двух смыслах: различают разделительное «или» и неразделительное «или». Дизъюнкция соответствует неразделительному «или», поскольку высказывание $a \vee b$ истинно во всех случаях, когда истинно хотя бы одно из высказываний a, b и когда истинны оба высказывания.

2. Из определения высказывания $a \to b$ видно, что оно истинно всегда, кроме случая, когда $a \equiv$ И, $b \equiv$ Л. Это означает, что в математической логике запрещено выводить лишь из истинных посылок ложные высказывания. Только в этом случае ложным будет сам вывод, т. е. сама импликация. Из ложного же высказывания даже правильными рассуждениями, т. е. путем истинной импликации, можно получить как истинное, так и ложное высказывание. В связи с этим говорят: «Из лжи следует все, что угодно».

Таблица 1.1 может служить также определением операций $\&, \vee, \to, \neg$ на двухэлементном множестве $\Omega = \{$И, Л$\}$. Они называются соответственно конъюнкцией, дизъюнкцией, импликацией и отрицанием. Их называют также основными логическими операциями. При этом прилагательное «основные» означает как наиболее широкое распространение в логике, так и тот факт, что через них естественным образом выражаются всевозможные другие операции на Ω.

Операции $\&$, $\vee$, $\neg$ обладают многими свойствами, сходными со свойствами операций $\cap$, $\cup$, $'$ на множестве всех подмножеств произвольного множества. А именно:

— операции $\&$, $\vee$ коммутативны, ассоциативны, идемпотентны, дистрибутивны одна относительно другой и связаны законами поглощения

$$a\&(a \vee b) \equiv a,$$

$$a \vee (a\&b) \equiv a;$$

— операция отрицания инволютивна, т. е. $\overline{\overline{a}} = a$, и связана с операциями $\&$, $\vee$ законами де Моргана

$$\overline{a\&b} \equiv \overline{a} \vee \overline{b}\ ,\ \overline{a \vee b} \equiv \overline{a}\&\overline{b}$$

и соотношениями

$$a\&\overline{a} \equiv \text{Л}\ ,\ a \vee \overline{a} \equiv \text{И}.$$

Отсюда следует, что множество $\Omega = \{\text{И}, \text{Л}\}$ относительно операций $\&$, $\vee$, $\neg$ является булевой алгеброй (см. определение 1.20). В ней роль 1 и 0 играют соответственно элементы И и Л. Эта простейшая, двухэлементная булева алгебра называется алгеброй высказываний (поскольку ее элементами являются значения высказываний).

Операция $\to$ на Ω также обладает рядом свойств, связывающих ее с другими операциями. Например, имеют место соотношения:

$$a \to b \equiv \overline{b} \to \overline{a} \text{ — закон контрапозиции,}$$

$$a \to (b \to a) \equiv \text{И},$$

$$a\&b \to b \equiv \text{И},$$

$$a \to b \equiv \overline{a} \vee b$$

и др. В частности, последнее из выписанных соотношений выражает импликацию через дизъюнкцию и отрицание.

Из соотношения $\overline{\overline{a}} \equiv a$ и законов де Моргана легко получаются равносильности

$$a\&b \equiv \overline{\overline{a} \vee \overline{b}}\ ,\ a \vee b \equiv \overline{\overline{a}\&\overline{b}},$$

показывающие, что операции $\&$ и $\vee$ могут быть выражены соответственно через $\vee$, $\neg$ и $\&$, $\neg$. Таким образом, при построении алгебры высказываний можно было бы взять всего лишь две основные операции $\&$, $\neg$. Однако в этом случае даже простейшие свойства алгебры высказываний имели бы достаточно громоздкие записи и были бы менее прозрачными с содержательной точки зрения. Так, свойство дистрибутивности операции $\&$ относительно $\vee$ записалось бы в виде

$$a\&\overline{(\overline{b}\&\overline{c})} \equiv \overline{(\overline{(a\&b)}\&\overline{(a\&c)})}.$$

2.2. ФОРМУЛЫ АЛГЕБРЫ ВЫСКАЗЫВАНИЙ

Прежде чем дать определение формулы алгебры высказываний, приведем некоторые замечания общего характера о формулах. Обычно под формулами понимают определенные строчки некоторых символов, или, как говорят, слова в некотором алфавите.

Алфавитом называют произвольное множество попарно различных символов, допускающих такую запись, по которой однозначно восстанавливаются сами символы. Обычно символ отождествляется с любой своей записью, в связи с чем символы алфавита называют также его буквами. В математике в качестве символов зачастую используются буквы латинского и других алфавитов, цифры, буквы с индексами, символы операций $+$, $\cdot$, $-$, $\oplus$ и др.

Если $A = \{a_1, a_2, ..., a_n\}$ — алфавит, то любая конечная последовательность

$$a_{i_1} a_{i_2} ... a_{i_m}$$

его букв называется словом в алфавите A. При этом число m называется длиной слова. Длина слова P обозначается в виде $l(P)$. Для удобства в рассуждениях вводится еще символ Λ для обозначения пустого слова, т. е. слова, не содержащего ни одной буквы. По определению $l(\Lambda) = 0$.

Так как непустое слово есть конечная последовательность букв, то можно все буквы в слове естественным образом занумеровать и говорить о 1-й, 2-й и т. д. буквах слова. Обычно слова записывают в строчки, а буквы в них нумеруют слева направо. Два слова называются равными, или графически равными, если они имеют одинаковую длину и соответствующие их буквы равны. Равенство слов будем обозначать знаком $=$. Множество всех слов и всех слов длины m в алфавите A обозначим соответственно через $W(A)$ и $W_m(A)$.

На множестве $W(A)$ можно ввести операцию умножения (или приписывания) слов, взяв в качестве произведения слов P, Q слово PQ, полученное приписыванием слова Q справа к слову P. Слово PQ называют также конкатенацией слов P, Q.

Говорят, что слово P является подсловом слова Q, если существуют такие слова L, R (возможно пустые), что

$$Q = LPR.$$

В этом случае говорят также, что слово P входит, или имеет вхождение, в слово Q. Может оказаться, что указанная выше пара слов L, R находится по словам P и Q неоднозначно. Выпишем все такие различные пары слов (L_i, R_i) и упорядочим их по возрастанию длины слова L_i. Получим

$$Q = L_1PR_1 = L_2PR_2 = ... = L_sPR_s.$$

В этом случае говорят, что имеется s вхождений слова P в слово Q и все эти вхождения упорядочивают по возрастанию длины слова L_i. В соответствии с этим говорят о 1-м, 2-м и т. д. вхождениях слова P в Q. Например, слово «арарат» содержит два вхождения слова «ара».

Под формулами в той или иной математической теории понимаются не произвольные слова в выбранном алфавите, а лишь слова, составленные по определенным правилам. Укажем основной алфавит и правила построения формул в алгебре высказываний.

Наряду с конкретными высказываниями, каждое из которых имеет значение «И» или «Л», будем рассматривать также переменные высказывания, значениями которых являются конкретные высказывания. Условимся обозначать конкретные высказывания малыми латинскими буквами a, b, c с индексами и без индексов, а переменные высказывания — буквами x, y, z, также возможно с индексами. Кроме указанных символов, включим в основной алфавит символы определенных выше логических операций и так называемые служебные символы — различные скобки и запятую.

Приводимое ниже определение формул будет конструктивным, в нем будет указано, каким образом каждая заданная формула построена из исходных постоянных и переменных высказываний. Сразу договоримся обозначать формулы большими латинскими буквами, возможно с индексами.

Определение 2.2. *1. Любое постоянное или переменное высказывание есть формула. Такие формулы называются элементарными.*

2. Если A и B — формулы, то слова

$$(A)\&(B)\ ,\ (A)\vee(B)\ ,\ (A)\to(B)$$

также являются формулами.

3.Если A — формула, то $\overline{(A)}$ тоже формула.

4. Других формул нет.

Все определенные таким образом формулы называются формулами алгебры высказываний.

Общее число всех логических операций, участвующих в формуле A, назовем рангом формулы A и обозначим через $r(A)$. Общее число всех постоянных и переменных высказываний, участвующих в формуле A, назовем длиной формулы A и обозначим через $l(A)$. Подчеркнем, что в определениях ранга и длины формулы каждый символ операции и высказывания считается столько раз, сколько раз он входит в формулу. Так, например, формула

$$\overline{(((a)\&(b))\to((x)\vee(y)))}\to\overline{(x)} \tag{2.2}$$

имеет ранг 6 и длину 5.

Для упрощения записи формул условимся опускать некоторые скобки в написании формул, но так, чтобы при необходимости мы по упрощенной записи формулы могли однозначно восстановить ее полную запись в соответствии с определением 2.2. Условимся:

1) не заключать в скобки элементарные формулы;

2) не заключать в скобки формулу, над которой находится знак отрицания;

3) считать, что операция $\&$ сильнее операции $\vee$ и обе эти операции сильнее операции $\to$;

4) не заключать в скобки большие латинские буквы, используемые для обозначения формул (например, вместо $(A)\&(B)$ писать $A\&B$).

Заметим, что, пользуясь указанными правилами, формулу (2.2) можно записать в виде

$$\overline{a\&b \to x \vee y} \to \overline{x}.$$

Приведем индуктивное определение подформул и главных подформул любой формулы алгебры высказываний.

Определение 2.3. *1. Подформулой элементарной формулы является лишь сама она.*

2. Подформулами любой формулы вида $A\&B$, $A \vee B$, $A \to B$ называются сама эта формула и все подформулы формул A и B. При этом формулы A и B называются главными подформулами.

3. Подформулами формулы $\overline{A}$ является сама она и все подформулы формулы A. При этом A называется главной подформулой.

Индуктивное определение формулы позволяет использовать метод полной математической индукции (по рангу или длине формулы) при доказательстве различных утверждений о формулах. Проиллюстрируем это на следующем простом, но постоянно используемом утверждении.

Теорема 2.1. *Если в формуле A некоторое вхождение подформулы B заменить на любую другую формулу C, то получим снова формулу.*

□ Докажем теорему индукцией по рангу формулы A. Если $\mathrm{rang}A = 0$, то A — элементарная формула, в которой единственной подформулой является сама она. Тогда при замене получим формулу C, и утверждение теоремы верно. Допустим, что теорема верна для всех формул A ранга $r \leq m$, и докажем для случая, когда $\mathrm{rang}A = m+1$. Так как $\mathrm{rang}A > 0$, то A есть формула по пункту 2 или 3 определения 2.2. Пусть $A = A_1\&A_2$. Тогда согласно определению 2.3 B либо совпадает с A, либо заменяемое вхождение B является подформулой одной из формул A_1, A_2. Если $B = A$, то после замены получим формулу C. Если B заменяется в A_1, то после замены

формула A_1 перейдет в формулу A'_1 по предположению индукции. Тогда A перейдет в слово $A'_1 \& A_2$, которое является формулой по пункту 2 определения 2.2. Аналогично доказывается утверждение теоремы и в остальных случаях. □

2.3. ЭКВИВАЛЕНТНЫЕ ФОРМУЛЫ

Все символы логических операций в формуле A можно считать занумерованными в порядке их использования при построении формулы A согласно определению формул. Порядок операций в формуле определяется скобками или указанными выше правилами сокращения числа скобок в записи формулы. Если в формулу A не входят никакие переменные высказывания, то формула сама является высказыванием, и его значение, называемое также значением формулы A, очевидным образом вычисляется по значениям входящих в формулу постоянных высказываний. Для этого достаточно произвести в нужном порядке все участвующие в формуле логические операции над значениями постоянных высказываний.

Приведите в качестве упражнения строгое индуктивное определение значения формулы, не содержащей переменных высказываний.

Если в формулу A не входят никакие переменные высказывания, кроме переменных из некоторой фиксированной системы $x_1, ..., x_n$, то формулу A обозначают также в виде

$$A(x_1, ..., x_n) \tag{2.3}$$

и говорят, что A есть формула от переменных $x_1, ..., x_n$. Заметим, что формула (2.3) может не содержать некоторых (и даже всех) переменных из $x_1, ..., x_n$, так что одну и ту же формулу можно рассматривать как формулу от разных систем переменных. Важно лишь, чтобы любая из таких систем переменных содержала все переменные, входящие в формулу A. Условимся еще считать, что все подформулы формулы $A(x_1, ..., x_n)$ также являются формулами от переменных $x_1, ..., x_n$.

Если в формуле (2.3) заменить переменные $x_1, ..., x_n$ соответственно постоянными высказываниями $a_1, ..., a_n$, то полу-

чится высказывание, которое обозначается в виде $A(a_1, ..., a_n)$. Значение этого высказывания называют значением формулы (2.3) при $x_1 = a_1, ..., x_n = a_n$.

Определение 2.4. *Формулы*

$$A(x_1, ..., x_n)\ ,\ B(x_1, ..., x_n)$$

алгебры высказываний называются эквивалентными, если они принимают одинаковые значения при любых значениях переменных высказываний $x_1, ..., x_n$.

Эквивалентность формул $A(x_1, ..., x_n)$, $B(x_1, ..., x_n)$ будем обозначать в виде

$$A(x_1, ..., x_n) \equiv B(x_1, ..., x_n). \tag{2.4}$$

Заметим, что определение 2.4 корректно, поскольку введенное отношение является отношением эквивалентности на множестве всех формул от любого фиксированного множества переменных высказываний.

Легко видеть также, что введенное отношение не зависит от выбора системы переменных $x_1, ..., x_n$, и потому вместо соотношения (2.4) можно писать короче: $A \equiv B$.

Так как значение формулы (2.3) при $x_1 = a_1, ..., x_n = a_n$ зависит не от содержательного смысла высказываний $a_1, ..., a_n$, а от их значений истинности, то для проверки соотношения (2.4) достаточно убедиться в том, что его левая и правая части принимают одинаковые значения при замене всех постоянных высказываний из формул A, B их значениями и при всевозможных заменах каждого из переменных $x_1, ..., x_n$ на «И», «Л».

Приведем примеры.

Пример 2.1. Выяснить, являются ли эквивалентными формулы:

$$A_1 = x_1 \to x_2\ ,\ A_2 = \overline{x}_1 \to \overline{x}_2\ ,$$

$$A_3 = \overline{x}_2 \to \overline{x}_1\ ,\ A_4 = x_2 \to x_1.$$

Составим таблицу значений истинности для указанных формул:

Таблица 1.2

x_1	x_2	A_1	A_2	A_3	A_4
И	И	И	И	И	И
И	Л	Л	И	Л	И
Л	И	И	Л	И	Л
Л	Л	И	И	И	И

Из таблицы 1.2 видно, что $A_1 \equiv A_3$, $A_2 \equiv A_4$ и A_1 не эквивалентна A_2.

Пример 2.2. Выяснить, являются ли эквивалентными формулы

$$A_1 = (\overline{x}_1 \vee x_2 \to x_3) \vee x_1 \& x_2 \ , \ A_2 = x_1 \vee x_2.$$

Здесь для решения вопроса можно было бы, как и в предыдущем примере, составить таблицы истинности для формул A_1, A_2 и сравнить их. Однако в данном случае задачу можно решить проще. Пользуясь определениями логических операций, нетрудно заметить, что формула A_1 принимает значение «Л» лишь при $x_1 = x_2 = x_3 =$ Л, а формула A_2 принимает значение «Л» еще при $x_1 =$ Л, $x_2 =$ И, $x_3 =$ Л. Следовательно, формулы A_1, A_2 не эквивалентны.

Укажем ряд основных соотношений эквивалентности для формул алгебры высказываний.

Теорема 2.2. *Для любых формул A, B, C алгебры высказываний имеют место следующие эквивалентности:*

1) $A\&B \equiv B\&A$;
2) $A \vee B \equiv B \vee A$;
3) $(A\&B)\&C \equiv A\&(B\&C)$;
4) $(A \vee B) \vee C \equiv A \vee (B \vee C)$;
5) $A\&(B \vee C) \equiv (A\&B) \vee (A\&C)$;
6) $A \vee B\&C \equiv (A \vee B)\&(A \vee C)$;
7) $A\&(A \vee B) \equiv A$;
8) $A \vee A\&B \equiv A$;

9) $A\&A \equiv A$;
10) $A \vee A \equiv A$;
11) $\overline{A\&B} \equiv \overline{A} \vee \overline{B}$;
12) $\overline{A \vee B} \equiv \overline{A}\&\overline{B}$;
13) $A \to B \equiv \overline{B} \to \overline{A}$;
14) $\overline{\overline{A}} \equiv A$;
15) $A \vee \overline{A} \equiv И$;
16) $A\&\overline{A} \equiv Л$;
17) $A \to B \equiv \overline{A} \vee B$.
18) а) $A\&И \equiv A$; *б)* $A \vee И \equiv И$; *в)* $A\&Л \equiv Л$; *г)* $A \vee И \equiv A$.

□ Справедливость свойств 1–8 следует непосредственно из определения эквивалентности формул и свойств основных логических операций. □

Эквивалентности 1–8 называют законами, или правилами логики. Многие из них имеют особые названия. В частности, законы 1–12 носят названия, сходные с названиями свойств соответствующих логических операций. Например, закон коммутативности конъюнкции и т. п. Эквивалентности 13–16 называются соответственно законом контрапозиции, законом двойного отрицания, законом исключенного третьего и законом противоречия. Эквивалентность 17 служит для выражения импликации через отрицание и дизъюнкцию.

Эквивалентности 1–8 часто используются для преобразования формул к нужному виду и при доказательствах теорем.

Так, зачастую вместо теоремы вида $A \to B$ доказывается эквивалентное ей утверждение $\overline{B} \to \overline{A}$. При этом используется закон контрапозиции 13. Закон исключенного третьего 15 обычно используется при доказательствах от противного, когда для доказательства теоремы A опровергают утверждение $\overline{A}$ и отсюда на основании эквивалентности $A \vee \overline{A} \equiv И$ делают вывод об истинности A. Эквивалентность 17 служит для выражения импликации через отрицание и дизъюнкцию.

Учитывая свойство ассоциативности операций $\&$ и $\vee$, можно условиться не расставлять скобки в конъюнкциях и дизъюнкциях многих компонент.

В преобразованиях формул часто используется также следующее правило замены.

Теорема 2.3. *Если в формуле $A = A(x_1, ..., x_n)$ некоторое вхождение подформулы $B = B(x_1, ..., x_n)$ заменить эквивалентной ей формулой $C = C(x_1, ..., x_n)$, то получим формулу, эквивалентную формуле A.*

Доказательство легко проводится индукцией по рангу формулы A. Схема рассуждений сходна с доказательством теоремы 2.1. Докажите теорему в качестве упражнения.

Определение 2.5. *Пусть A — формула, не содержащая символа $\to$. Тогда формула, полученная из A заменой каждого символа $\&$ на $\vee$, каждого символа $\vee$ на $\&$ и каждого постоянного высказывания на его отрицание, называется двойственной к A. Обозначим ее через A^*.*

Теорема 2.4. *Для любой формулы $A = A(x_1, ..., x_n)$, не содержащей символа $\to$, имеет место эквивалентность*

$$A^*(x_1, ..., x_n) \equiv \overline{A(\overline{x_1}, ..., \overline{x_n})}. \tag{2.5}$$

□ Докажем теорему индукцией по рангу r формулы A. Если $r(A) = 0$, то A либо постоянное высказывание a, либо переменное высказывание x_i. В первом случае $\overline{A(\overline{x_1}, ..., \overline{x_n})} = \overline{a}$ и $A^*(x_1, ..., x_n) = \overline{a}$, во втором — $\overline{A(\overline{x_1}, ..., \overline{x_n})} = \overline{\overline{x_i}}$ и $A^*(x_1, ..., x_n) = x_i$. В обоих случаях соотношение (2.5) верно. Допустим, что оно выполняется при $r \le m$, и пусть $r(A) = m + 1$. Возможны три варианта: $A = \overline{A_1}$, $A = A_1 \& A_2$, $A = A_1 \vee A_2$. Если $A = \overline{A_1}$, то используя предположение индукции, находим

$$A^*(x_1, ..., x_n) = (\overline{A_1(x_1, ..., x_n)})^* \equiv$$
$$\equiv \overline{\overline{A_1(\overline{x}_1, ..., \overline{x}_n)}} = \overline{A(\overline{x}_1, ..., \overline{x}_n)}.$$

Пусть теперь $A = A_1 \& A_2$. Воспользовавшись предположением индукции и законом де Моргана 11, получим

$$(A(x_1,...,x_n))^* = (A_1(x_1,...,x_n))^* \vee (A_2(x_1,...,x_n))^* \equiv$$
$$\equiv \overline{A_1(\overline{x}_1,...,\overline{x}_n)} \vee \overline{A_2(\overline{x}_1,...,\overline{x}_n)} \equiv$$
$$\equiv \overline{A_1(\overline{x}_1,...,\overline{x}_n)\&A_2(\overline{x}_1,...,\overline{x}_n)} = \overline{A(\overline{x}_1,...,\overline{x}_n)}.$$

Аналогично с использованием закона де Моргана 12 доказывается эквивалентность (2.5) и в случае, когда $A = A_1 \vee A_2$. □

Следствие 1 (Принцип двойственности). *Для любых формул A, B алгебры высказываний*

$$A \equiv B \Longleftrightarrow A^* \equiv B^*.$$

Принцип двойственности позволяет вместо двух равносильностей $A \equiv B$ и $A^* \equiv B^*$ доказывать любую одну из них.

2.4. ПРИВЕДЕННЫЕ ФОРМУЛЫ И НОРМАЛЬНЫЕ ФОРМЫ

В этом пункте и всюду далее для упрощения записей условимся вместо значений «И», «Л» высказываний использовать соответственно 1 и 0.

Так как отношение $\equiv$ является отношением эквивалентности на множестве всех формул от любой фиксированной системы переменных высказываний, то множество всех таких формул разбивается на непересекающиеся классы эквивалентных формул, и для нахождения значений любой заданной формулы можно воспользоваться другими (возможно, проще устроенными) формулами из того же класса. В качестве таких формул зачастую выступают так называемые приведенные формулы и конъюнктивные и дизъюнктивные нормальные формы.

Определение 2.6. *Формула A алгебры высказываний называется приведенной, если:*

1) A не содержит постоянных высказываний;

2) A не содержит операции $\rightarrow$;

3) операция отрицания в A относится лишь к элементарным подформулам.

Теорема 2.5. *Для любой формулы A существует эквивалентная ей приведенная формула.*

□ Для обеспечения условий 1–2 определения 2.3 достаточно заменить в A каждое истинное высказывание формулой $x \vee \overline{x}$, каждое ложное высказывание — формулой $x\&\overline{x}$ и любую подформулу вида $A_1 \to A_2$ эквивалентной ей формулой $\overline{A_1} \vee A_2$. Полученная в итоге формула в силу теоремы 2.3 будет эквивалентной A. Теперь для выполнения условия 3 нам, возможно, придется последовательно заменять подформулы вида

$$\overline{\overline{A_1}}\ ,\ \overline{A_1 \& A_2}\ ,\ \overline{A_1 \vee A_2}$$

соответственно формулами

$$A_1\ ,\ \overline{A_1} \vee \overline{A_2}\ ,\ \overline{A_1}\&\overline{A_2}$$

до тех пор, пока не получим приведенную формулу. □

Заметим, что приведенная формула, эквивалентная заданной формуле, находится неоднозначно.

Пример 2.3. Формула

$$A = \overline{(\overline{x_1 \& x_2} \to x_1 \& a) \vee x_1 \& x_3},$$

где $a \equiv$ И, эквивалентна каждой из приведенных формул

$$(\overline{x_1} \vee \overline{x_2})\&\overline{x_1}\&(\overline{x_1} \vee \overline{x_3})\ ,\ \overline{x_1}.$$

Для определения дизъюнктивных и конъюнктивных нормальных форм введем для высказывания x и числа $\alpha \in \{0, 1\}$ обозначение

$$x^\alpha = \begin{cases} x\ , \text{ если } \alpha = 1 \\ \overline{x}\ , \text{ если } \alpha = 0 \end{cases}.$$

Определение 2.7. *Формулы вида*

$$x_{i_1}^{\alpha_1} \& x_{i_2}^{\alpha_2} \& ... \& x_{i_k}^{\alpha_k} \text{ и } x_{i_1}^{\alpha_1} \vee x_{i_2}^{\alpha_2} \vee ... \vee x_{i_k}^{\alpha_k}, k > 0$$

называются соответственно элементарной конъюнкцией и элементарной дизъюнкцией. Элементарные конъюнкции (дизъюнкции), в которых каждое переменное имеет не более одного вхождения (с отрицанием или без отрицания), будем называть конъюнктами (дизъюнктами).

Определение 2.8. *Формула называется дизъюнктивной нормальной формой (ДНФ), если она является дизъюнкцией элементарных конъюнкций.*

Двойственным образом определяется конъюнктивная нормальная форма (КНФ).

Теорема 2.6. *Для всякой формулы A алгебры высказываний существуют эквивалентные ей дизъюнктивная и конъюнктивная нормальные формы (ДНФ и КНФ формулы A).*

□ Не теряя общности, будем считать, что формула A приведенная. Докажем теорему индукцией по длине $l(A)$ формулы A.

Если $l(A)$=1, то A — либо переменное высказывание x, либо результат применения (возможно, многократного) операции отрицания к x. В последнем случае, применяя закон двойного отрицания, мы получим формулу x или $\overline{x}$. А так как каждая из формул x, $\overline{x}$ является ДНФ и КНФ, то утверждение теоремы верно.

Пусть $l(A) > 1$. Так как A приведенная, то $A = A_1 \& A_2$ или $A = A_1 \vee A_2$. По предположению индукции имеем

$$A_1 \equiv \vee_{i=1}^{n} A_{1i} \equiv \&_{j=1}^{m} B_{1j}\ ,\ A_2 \equiv \vee_{i=1}^{r} A_{2i} \equiv \&_{j=1}^{t} B_{2j},$$

где A_{ij} — элементарные конъюнкции, а B_{ij} — элементарные дизъюнкции. Если $A = A_1 \& A_2$, то A равносильна КНФ

$$(\&_{i=1}^{m} B_{1i}) \& (\&_{j=1}^{s} B_{2j}),$$

а применяя закон дистрибутивности конъюнкции относительно дизъюнкции, получим и ДНФ:

$$A \equiv (\vee_{i=1}^{n} A_{1i}) \& (\vee_{j=1}^{r} A_{2j}) = \vee_{i=1}^{n} \vee_{j=1}^{r} (A_{1i} \& A_{2j}).$$

В случае $A = A_1 \vee A_2$ рассуждения двойственны. □

Заметим, что доказательство теорем 2.5, 2.6 указывает и сам алгоритм нахождения ДНФ и КНФ любой формулы A. Сначала нужно найти эквивалентную A приведенную формулу, а затем, пользуясь законами 1–6, приводить к ДНФ

(КНФ) все подформулы полученной формулы, начиная с подформул меньшей длины.

Пример 2.4. Найти КНФ и ДНФ, эквивалентные формуле

$$A = x_1 \& x_2 \to (x_1 \to x_2)(x_2 \to x_1).$$

Сначала, пользуясь законами 17 и 11, найдем приведенную формулу B, эквивалентную A:

$$B = \overline{x_1} \vee \overline{x_2} \vee (\overline{x_1} \vee x_2) \& (\overline{x_2} \vee x_1).$$

Из нее, воспользовавшись законом дистрибутивности 6, получим КНФ формулы A:

$$A \equiv (\overline{x_1} \vee \overline{x_2} \vee \overline{x_1} \vee x_2) \& (\overline{x_1} \vee \overline{x_2} \vee \overline{x_2} \vee x_1).$$

Применив к формуле B закон дистрибутивности 5, получим ДНФ формулы A:

$$A \equiv \overline{x_1} \vee \overline{x_2} \vee \overline{x_1} \& \overline{x_2} \vee \overline{x_1} \& x_1 \vee x_2 \& \overline{x_2} \vee x_2 \& x_1.$$

В принципе, нахождение КНФ формулы A можно свести к нахождению ДНФ для формулы A^*, используя очевидные свойства: $(A^*)^* = A$, и формула, двойственная к ДНФ, является КНФ.

Заметим, что ДНФ и КНФ для любой заданной формулы находятся неоднозначно. Так, в примере 2.3 для формулы A найдены две КНФ. В связи с этим представляют интерес так называемые совершенные ДНФ и КНФ.

Определение 2.9. *Совершенной дизъюнктивной нормальной формой (СДНФ) от переменных $x_1, ..., x_n$ называется такая ДНФ, в которой все элементарные конъюнкции различны и каждая из них имеет вид*

$$x_1^{\alpha_1} \& x_2^{\alpha_2} \& ... \& x_n^{\alpha_n}. \tag{2.6}$$

Двойственным образом определяется совершенная конъюнктивная нормальная форма (СКНФ).

Теорема 2.7. *Для всякой формулы* $A(x_1, ..., x_n) \not\equiv 0$ $(A(x_1, ..., x_n) \not\equiv 1)$ *существует единственная с точностью до перестановки конъюнктов (дизъюнктов) эквивалентная ей совершенная ДНФ (КНФ) от переменных* $x_1, ..., x_n$.

□ Так как

$$x^\alpha = 1 \Longleftrightarrow x = \alpha,$$

то

$$x_1^{\alpha_1} \& ... \& x_n^{\alpha_n} = 1 \Longleftrightarrow x_1 = \alpha_1, ..., x_n = \alpha_n. \qquad (2.7)$$

Пусть $A(x_1, ..., x_n) \not\equiv 0$. Обозначим

$$N(A) = \{(c_1, ..., c_n) : A(c_1, ..., c_n) = 1\}.$$

Тогда из (2.7) следует, что

$$A \equiv \vee_{(c_1,...,c_n)\in N(A)} (x_1^{c_1} \& ... \& x_n^{c_n}). \qquad (2.8)$$

Если для A существует другая СДНФ, то либо она содержит конъюнкт (2.6), не входящий в (2.8), либо в нее не входит некоторый конъюнкт из СДНФ (2.8). В первом случае мы имеем $A(c_1, ..., c_n) = 1$ для $(c_1, ..., c_n) \notin N(A)$, во втором — $A(c_1, ..., c_n) = 0$ для $(c_1, ..., c_n) \in N(A)$. В обоих случаях приходим к противоречию с определением $N(A)$.

Теперь утверждение теоремы о СКНФ следует из принципа двойственности. □

Заметим, что явное указание переменных в формулировке теоремы 2.7 существенно, так как одна и та же формула, рассматриваемая от разных переменных, имеет различные СДНФ. Например, СДНФ x_1 эквивалентна СДНФ $x_1 x_2 \vee x_1 \overline{x_2}$.

На теореме 2.7 основан следующий способ распознавания эквивалентности формул. Для заданных формул A, B находим СДНФ (или СКНФ) от одних и тех же переменных. Если найденные формулы совпадают с точностью до перестановки конъюнктов (дизъюнктов), то $A \equiv B$, в противном случае $A \not\equiv B$.

2.5. ВЫПОЛНИМЫЕ И ТОЖДЕСТВЕННО ИСТИННЫЕ ФОРМУЛЫ

Определение 2.10. *Формула A алгебры высказываний называется выполнимой, если она принимает значение 1 хотя бы на одном наборе значений, содержащихся в ней переменных высказываний. В противном случае она называется тождественно ложной, что обозначается в виде $A \equiv 0$.*

Определение 2.11. *Формула A алгебры высказываний называется тождественно истинной, или тавтологией, если она принимает значение 1, при любых значениях входящих в нее переменных высказываний. Обозначается это в виде $A \equiv 1$.*

В связи с приложениями алгебры высказываний естественно возникают задачи распознавания выполнимости, тождественной истинности и тождественной ложности любой заданной формулы.

Заметим, что, умея решать любую одну из этих трех задач, мы сможем решить две остальные задачи, поскольку:

$$A \equiv 1 \Longleftrightarrow \overline{A} \text{ невыполнима,}$$

$$A \equiv 0 \Longleftrightarrow A \text{ невыполнима.}$$

В принципе, каждая из указанных задач разрешима, поскольку в любую формулу входит лишь конечное число переменных, и все наборы их значений можно перебрать за конечное время. Поэтому речь идет не о принципиальной разрешимости задач, а о нахождении более простых алгоритмов по сравнению с перебором систем значений переменных высказываний. В некоторых случаях для этой цели могут использоваться следующие теоремы.

Теорема 2.8. *Формула A является тождественно истинной тогда и только тогда, когда в каждую элементарную дизъюнкцию любой ее КНФ входят какое-либо переменное и его отрицание.*

□ Пусть $\&_{i=1}^{k} A_i$ — какая-либо КНФ для формулы A и каждая элементарная дизъюнкция A_i содержит пару подформул вида x_{i_j} и $\overline{x_{i_j}}$. Так как операция $\vee$ коммутативна и $x \vee \overline{x} \equiv 1$, то $A_i \equiv 0$ при $i \in \{1, ..., k\}$, а потому и $A \equiv 1$.

Обратно, пусть $A \equiv 1$ и $\&_{i=1}^{k} A_i$ — КНФ для A. Тогда $\&_{i=1}^{k} A_i \equiv 1$, и по определению операции $\&$ имеем $A_i \equiv 1$ при $i \in \{1, ..., k\}$. Допустим, что в некоторой элементарной дизъюнкции A_i нет пары подформул вида x_j, $\overline{x_j}$. Тогда, придавая переменным, входящим в A_i без отрицания, значение 0, а остальным — значение 1, получим ложное высказывание. Следовательно, $A_i \not\equiv 1$, а потому и $A \not\equiv 1$, что противоречит условию. □

В силу принципа двойственности имеет место утверждение, двойственное к теореме 2.8.

Теорема 2.9. *Формула A тождественно ложна тогда и только тогда, когда в каждую элементарную конъюнкцию любой ее ДНФ входят какое-либо переменное и его отрицание.*

Например, формула A из примера 2.4 тождественно истинна, поскольку найденная в этом примере КНФ формулы A удовлетворяет условию теоремы 2.8.

2.6. СОКРАЩЕННЫЕ, ТУПИКОВЫЕ И МИНИМАЛЬНЫЕ ДНФ

Так как ДНФ (КНФ) формулы находится неоднозначно, то естественно возникает задача нахождения ее наиболее простой ДНФ (КНФ).

Определение 2.12. *ДНФ (КНФ) называется минимальной (сокращенно МДНФ (МКНФ)), если она имеет наименьшую длину среди всех эквивалентных ей ДНФ (КНФ).*

Существуют различные алгоритмы нахождения МДНФ (МКНФ). Учитывая принцип двойственности, мы рассмотрим вопрос о нахождении лишь МДНФ.

Для описания алгоритма нахождения МДНФ формулы A нам понадобится ряд новых понятий.

Заметим, что любая элементарная конъюнкция, не содержащая ни одного переменного вместе с его отрицанием, может быть приведена к эквивалентному ей конъюнкту с использованием лишь законов коммутативности 1 и идемпотентности 9.

Конъюнкты, отличающиеся лишь показателем одного из переменных, называются соседними. Если C_1, C_2 — соседние конъюнкты длины $k > 1$, отличающиеся показателем у переменного x_j, то

$$C_1 \vee C_2 \equiv C\&(x_j \vee \overline{x_j}) \equiv C,$$

где C — конъюнкт, полученный удалением из C_1 и C_2 переменного x_j. При этом говорят, что C получен склеиванием конъюнктов C_1, C_2 по переменному x_j, а C_1, C_2 получены из C расклеиванием по переменному x_j.

Определение 2.13. *Конъюнкт C называется импликантой формулы $A(x_1, ..., x_n)$, если он входит хотя бы в одну ДНФ формулы A.*

Из закона идемпотентности 10 следует, что конъюнкт

$$x_{i_1}^{\alpha_1}\&x_{i_2}^{\alpha_2}\&...\&x_{i_k}^{\alpha_k}$$

является импликантой формулы A тогда и только тогда, когда

$$A \vee x_{i_1}^{\alpha_1}\&x_{i_2}^{\alpha_2}\&...\&x_{i_k}^{\alpha_k} \equiv A.$$

В общем случае, если A, B — формулы и $A \vee B \equiv A$, то говорят, что формула A поглощает формулу B. Следовательно, можно сказать, что импликанта формулы A — это конъюнкт, поглощаемый формулой A.

Определение 2.14. *Импликанта формулы A называется простой, если удаление из нее любого сомножителя приводит к конъюнкту, не являющемуся импликантой формулы A.*

Согласно закону поглощения 8 любая импликанта формулы A поглощается некоторой простой импликантой. Следовательно, дизъюнкция всех простых импликант формулы A эквивалентна A. Поэтому корректно

Определение 2.15. *Дизъюнкция всех простых импликант формулы A называется сокращенной ДНФ формулы A.*

Непосредственно из определения сокращенной ДНФ видно, что для любой формулы $A \neq Л$ она существует и единственна с точностью до перестановок конъюнктов и сомножителей в конъюнктах.

Определение 2.16. *ДНФ*

$$C_1 \vee C_2 \vee ... \vee C_k \tag{2.9}$$

формулы A, составленная из ее простых импликант C_i, называется тупиковой ДНФ формулы A, если удаление из нее любого одного конъюнкта приводит к формуле, не эквивалентной A.

Сокращенная ДНФ формулы A играет важную роль при нахождении МДНФ формулы A в силу следующей теоремы.

Теорема 2.10. *Всякая импликанта, содержащаяся в какой-либо МДНФ формулы A, является простой, т. е. содержится в сокращенной ДНФ.*

□ Пусть (2.9) — любая МДНФ формулы A. Допустим, что импликанта C_1 не простая. Тогда найдется ее собственная часть B, являющаяся импликантой формулы A, т. е. $A \vee B \equiv \equiv A$. По закону поглощения 8 имеем также эквивалентность $C_1 \vee B \equiv B$. Из последних двух эквивалентностей, учитывая законы логики 2, 4, получим

$$A \equiv (C_1 \vee C_2 \vee ... \vee C_k) \vee B \equiv B \vee C_2 \vee ... \vee C_k.$$

Так как $l(B) < l(C_1)$, то (2.9) не является МДНФ формулы A, что противоречит условию. □

Из всего сказанного просматривается следующий путь нахождения МДНФ формулы A: сначала находим сокращенную ДНФ, затем из нее — все тупиковые ДНФ, из последних выбираем все ДНФ наименьшей длины.

Далее мы опишем два алгоритма нахождения сокращенной ДНФ и алгоритм нахождения МДНФ произвольной выполнимой формулы $A(x_1, ..., x_n)$.

Алгоритм нахождения сокращенной ДНФ по СДНФ

Пусть A_0 есть СДНФ формулы $A(x_1, ..., x_n)$. Алгоритм нахождения сокращенной ДНФ по СДНФ заключается в следующем.

Находим в A_0 все пары соседних конъюнктов и к каждой из них применяем операцию склеивания. В итоге получим

$$A_0 \equiv A_1 \vee B_1,$$

где A_1 — дизъюнкция всех попарно различных конъюнктов длины $n-1$, полученных в результате проведенных склеиваний, а B_1 — дизъюнкция всех конъюнктов, не имеющих соседних. Последние называют изолированными конъюнктами СДНФ A_0. Заметим, что каждая из формул B_1, A_1 может отсутствовать, если соответственно все конъюнкты изолированы или изолированных конъюнктов нет. В первом случае алгоритм закончен, во втором аналогичную процедуру применяем к формуле A_1. Получим

$$A_1 \equiv A_2 \vee B_2.$$

Продолжим этот процесс до тех пор, пока ни получим ДНФ A_m, в которой все конъюнкты изолированы, т. е. A_m не содержит пар соседних конъюнкций. Обозначим для общности записей $A_m = B_{m+1}$.

Теорема 2.11. *$B_1 \vee B_2 \vee ... \vee B_{m+1}$ есть сокращенная ДНФ формулы A.*

□ Из описания алгоритма видно, что A_i, B_i — являются дизъюнкциями импликант соответственно длин $n-i$, $n-i+1$ формулы A, и

$$A \equiv B_1 \vee B_2 \vee ... \vee B_{m+1}.$$

Теперь для доказательства теоремы достаточно доказать два утверждения:

1) A_i есть дизъюнкция всех импликант длины $n-i$ формулы A;

2) B_i есть дизъюнкция всех простых импликант длины $n-i+1$ формулы A.

Докажем их одновременной индукцией по i.

Пусть $i=1$ и C импликанта длины $n-1$ формулы A. Расклеив ее по не входящему в нее переменному, мы получим две различные импликанты длины n формулы A. Однако все такие импликанты входят в СДНФ формулы A. В описанном алгоритме они склеятся в импликанту C, которая войдет в A_1. Пусть теперь C простая импликанта длины n формулы A. Тогда она входит в СДНФ формулы A и не имеет в ней соседних конъюнкций, так как в противном случае при склеивании мы получили бы ее собственную часть, являющуюся импликантой формулы A. Это противоречит условию простоты импликанты C. Таким образом, при $i=1$ утверждения 1, 2 доказаны. Та же самая схема рассуждений применяется и при доказательстве перехода от $i=r$ к $i=r+1$. Разница лишь в том, что вместо свойства СДНФ содержать все импликанты длины n формулы A используется свойство ДНФ A_r содержать все импликанты длины $n-r$ формулы A, которое выполняется по предположению индукции. $\square$

Алгоритм нахождения сокращенной ДНФ по любой ДНФ

Не теряя общности, будем считать, что исходная ДНФ есть дизъюнкция конъюнктов.

Алгоритм состоит из преобразований двух типов. Сначала к исходной ДНФ формулы A добавляются ее новые импликанты так, чтобы в итоге получилась ДНФ, содержащая все простые импликанты формулы A. Затем из полученной ДНФ удаляются все конъюнкции, которые поглощаются простыми импликантами. Добавление импликант основано на следующей эквивалентности:

$$A\&x_i \vee B\&\overline{x_i} \equiv A\&x_i \vee B\&\overline{x_i} \vee A\&B \qquad (2.10)$$

имеющей место для любых формул A, B, не содержащих x_i. Эта эквивалентность проверяется непосредственно, отдельно

при $x = 0$ и $x = 1$. Удаление непростых импликант осуществляется с использованием закона поглощения 8. Пусть выполнимая формула $A(x_1, ..., x_n)$ задана произвольной ДНФ:

$$A(x_1, ..., x_n) = C_1 \vee C_2 \vee ... \vee C_m, \tag{2.11}$$

где, не теряя общности, можно считать, что $C_1, C_2, ..., C_m$ — импликанты формулы A.

Преобразование 1. Находим в ДНФ (2.11) импликанты, имеющие, с точностью до коммутативности, конъюнкции вид

$$C_i = B_i \& x_k,\ C_j = B_j \overline{x_k},$$

и такие, что в (2.11) нет конъюнкты C_r, эквивалентной конъюнкции $B_i \& B_j$. Тогда, заменив в (2.11) $C_i \vee C_j$ на $C_i \vee C_j \vee C_r$, получим новое представление формулы A в виде дизъюнкции ее импликант.

Теперь преобразование 1 применяем к полученной ДНФ, и т. д. до тех пор, пока не получим ДНФ, к которой не применимо преобразование 1.

Преобразование 2. Из полученной ДНФ удаляем импликанту, содержащую, с точностью до коммутативности конъюнкции, некоторую другую импликанту из этой же ДНФ. Далее применяем преобразование 2 к полученной ДНФ, и т. д. до тех пор, пока не получим ДНФ, к которой не применимо преобразование 2.

В итоге получится сокращенная ДНФ формулы A. Для доказательства этого факта достаточно показать, что если к ДНФ $B(x_1, ..., x_n)$ не применимы преобразования типа 1, то она содержит все свои простые импликанты.

Это утверждение легко доказывается индукцией по n.

При $n = 1$ оно очевидно. Пусть $n > 1$ и C — простая импликанта формулы B. Если длина $l(C) = n$, то C присутствует в ДНФ B, так как в противном случае C поглощалась бы некоторой конъюнкцией из B, что противоречит простоте импликанты C. Пусть $l(C) < n$. Тогда существует такое $i \in \in \{1, ..., n\}$, что C не содержит x_i. Если x_i не входит в B, то C входит в B по предположению индукции. В противном случае, учитывая, что к B не применимы преобразования 1,

B можно представить в виде $B \equiv x_i^\alpha \& B_1 \vee B_2$, где B_1, B_2 — ДНФ, не содержащие x_i, и все конъюнкты из B_2 входят в B. Так как C — импликанта формулы B, то

$$x_i^\alpha \& B_1 \vee B_2 \vee C \equiv x_i^\alpha \& B_1 \vee B_2,$$

откуда при $x_i^\alpha = 0$ получаем $B_2 \vee C \equiv B_2$. Отсюда по предположению индукции имеем: C входит в B_2, а потому и в B.

2.7. АЛГОРИТМ НАХОЖДЕНИЯ ТУПИКОВЫХ ДНФ ПО СОКРАЩЕННОЙ ДНФ

Из теоремы 2.9 следует, что для нахождения тупиковой ДНФ необходимо уметь находить, в каком случае один конъюнкт поглощается заданной ДНФ. Этой цели служит известный критерий поглощения. Для его формулировки полезно ввести понятие ортогональных формул.

Определение 2.17. *Формулы A и B называются ортогональными (обозначение: $A \perp B$), если $A \& B \equiv 0$.*

Легко видеть, что элементарные конъюнкции ортогональны тогда и только тогда, когда одна из них содержит некоторое переменное, а другая — его отрицание.

Теорема 2.12 (Критерий поглощения). *1. Если ДНФ A поглощает конъюнкта C и ДНФ B получена удалением из A всех элементарных конъюнкций, ортогональных C, то B поглощает C.*

2. Пусть C — конъюнкт, не ортогональный конъюнктам $C_1, ..., C_k$; C_{0i} — конъюнкция всех сомножителей, входящих в C и в C_i, причем $C_{0i} = 1$, если таких сомножителей нет; C_{1i} — конъюнкция остальных сомножителей из C_i, причем $C_{1i} = 1$, если $C_i = C_{0i}$. Тогда ДНФ $A = C_1 \vee ... \vee C_k$ поглощает C в том и только том случае, когда выполняется условие

$$B = C_{11} \vee ... \vee C_{1k} \equiv 1. \tag{2.12}$$

□ 1. Очевидно, что достаточно рассмотреть случай, когда B получена из A удалением лишь одного конъюнкта C_1.

Допустим, что утверждение 1 неверно, т. е.

$$A \vee C \equiv A, \text{ но } B \vee C \not\equiv B.$$

Тогда существует набор α значений переменных, при котором $B(\alpha) \equiv 0$ и $C(\alpha) \equiv 1$. А так как $C\&C_1 \equiv 0$, то $C_1(\alpha) \equiv 0$, что вместе с эквивалентностью $B(\alpha) \equiv 0$ приводит к условию $A(\alpha) \equiv 0$. Следовательно, $A \vee C\&A \not\equiv A$, что противоречит условию. В итоге утверждение 1 доказано.

2. Пусть выполнено условие (2.12) и $A \vee C\&A \not\equiv A$. Тогда существует набор значений переменных α такой, что $C(\alpha) \equiv \equiv 1$, $A(\alpha) \equiv 0$, а потому и $C_{0i}(\alpha) \equiv 1$ при $i = 1, ..., k$. Кроме того, из условия (2.12) следует, что $C_{1i}(\alpha) \equiv 1$ хотя бы для одного значения i. Тогда $A \vee C\&A \not\equiv A\ C_{0i}(\alpha)\&C_{1i}(\alpha) \equiv C_i(\alpha) \equiv \equiv 1$ и, следовательно, $A(\alpha) \equiv 1$. Получили противоречие.

Обратно, пусть $A \vee C \equiv A$, но условие (2.12) не выполнено. Тогда для некоторого набора переменных $\alpha = (a_1, ..., a_n)$ имеем: $C_{1i}(\alpha) \equiv 0$, а потому и $C_i(\alpha) \equiv 0$ для всех $i = 0, ..., k$. Так как C не ортогонально C_i, то в C_{1i} не может входить отрицание какого-либо сомножителя из C. Отсюда и из правила получения C_{1i} следует, что в формулу B могут входить лишь переменные, не входящие в C. Пусть, например, в B входят переменные $x_1, ..., x_m$, а в C — $x_{m+1}, ..., x_n$. Тогда имеем

$$B(a_1, ..., a_m, x_{m+1}, ..., x_n) \equiv A(a_1, ..., a_m, x_{m+1}, ..., x_n) \equiv 0$$

при любых значениях переменных $x_{m+1}, ..., x_n$. Так как C не ортогональна C_i, то C не может быть тождественно ложной формулой, и потому

$$C(a_1, ..., a_m, b_{m+1}, ..., b_n) \equiv 1$$

при некоторых $b_{m+1}, ..., b_n$. Следовательно, формулы A, C принимают на наборе $(a_1, ..., a_m, b_{m+1}, ..., b_n)$ различные значения, что противоречит условию $A \vee C \equiv A$. □

Теперь можно указать алгоритм получения всех тупиковых, а значит, и всех минимальных ДНФ любой выполнимой формулы A по ее сокращенной ДНФ B.

Сначала, пользуясь критерием поглощения, находим все конъюнкты из B, поглощаемые дизъюнкцией остальных конъюнктов. Если таких нет, то алгоритм закончен и ДНФ B является единственной тупиковой и минимальной ДНФ формулы A. В противном случае, удаляя из B конъюнкты, поглощаемые дизъюнкцией остальных конъюнктов, получим некоторую систему ДНФ, эквивалентных формуле A. Затем ту же самую процедуру применяем к каждой из найденных ДНФ и берем объединение полученных при этом ДНФ. Этот процесс продолжаем до тех пор, пока не получим систему ДНФ, к которым не применим критерий поглощения. Эта система и будет совпадать с множеством всех тупиковых ДНФ формулы A.

Пример 2.5. Найти все минимальные ДНФ формулы A по ее СДНФ

$$x_1\&x_2\&x_3\&x_4 \vee x_1\&x_2\&x_3\&\overline{x_4} \vee$$
$$\vee x_1\&x_2\&\overline{x_3}\&x_4 \vee x_1\&x_2\&\overline{x_3}\&\overline{x_4} \vee x_1\&\overline{x_2}\&\overline{x_3}\&\overline{x_4} \vee$$
$$\vee x_1\&\overline{x_2}\&\overline{x_3}\&x_4 \vee \overline{x_1}\&x_2\&x_3\&x_4 \vee$$
$$\vee \overline{x_1}\&x_2\&x_3\&\overline{x_4} \vee \overline{x_1}\&\overline{x_2}\&x_3\&x_4 \vee$$
$$\vee \overline{x_1}\&\overline{x_2}\&x_3\&\overline{x_4} \vee \overline{x}_1\&\overline{x}_2\&\overline{x_3}\&\overline{x_4}.$$

Применяя описанный выше алгоритм нахождения сокращенной ДНФ по СДНФ, получим сокращенную ДНФ B формулы A:

$$B = \overline{x_1}\&\overline{x_2}\&\overline{x_4} \vee x_1\&x_2 \vee x_2\&x_3 \vee$$
$$\vee x_1\&\overline{x_3}\& \vee \overline{x_1}\&x_3 \vee \overline{x_2}\&\overline{x_3}\&\overline{x_4}.$$

Теперь, применяя критерий поглощения, получим четыре тупиковых ДНФ:

$$\overline{x_1}\&\overline{x_2}\&\overline{x_4} \vee x_1\&x_2 \vee x_1\&\overline{x_3} \vee \overline{x_1}\&x_3,$$

$$\overline{x_1}\&\overline{x_2}\&\overline{x_4} \vee x_2\&x_3 \vee x_1\&\overline{x_3} \vee \overline{x_1}\&x_3,$$

$$\overline{x_2}\&\overline{x_3}\&\overline{x_4} \vee x_2\&x_3 \vee x_1\&\overline{x_3} \vee \overline{x_1}\&x_3,$$

$$\overline{x_2}\&\overline{x_3}\&\overline{x_4} \vee x_1\&x_2 \vee x_1\&\overline{x_3} \vee \overline{x_1}\&x_3.$$

Так как все они имеют одну и ту же длину, то все они являются и минимальными ДНФ формулы A.

В заключение данного параграфа укажем еще один метод получения МДНФ по СДНФ и сокращенной ДНФ, называемый методом Квайна. Мы проиллюстрируем его на предыдущем примере.

Занумеруем конъюнкты из сокращенной ДНФ числами 1–6 в том порядке, как они выписаны выше, и составим таблицу, у которой во входную строку выпишем указанные номера конъюнктов из сокращенной ДНФ, а во входной столбец — системы показателей переменных в конъюнктах из СДНФ. На пересечении столбца с конъюнктом i и строки с конъюнктом C из СДНФ поставим знак «+», если C поглощается конъюнктом с номером i, и знак «−» в противном случае. Теперь по таблице находим минимальные подсистемы конъюнктов из системы 1–6, поглощающие все конъюнкты из СДНФ. В нашем случае мы замечаем, что конъюнкт $x_1\&\overline{x}_2\&\overline{x}_3\&x_4$ поглощается только конъюнктом 4, поэтому последний обязан входить в любую тупиковую ДНФ формулы A. По тем же соображениям это верно и для конъюнкта 5. Теперь, учитывая отмеченные факты, замечаем, что для поглощения всех остальных конъюнктов из СДНФ можно взять лишь конъюнкты с номерами 1–2, или 2–6, или 1–3, или 3–6. В итоге приходим к указанным выше четырем минимальным ДНФ формулы A (табл. 1.3).

Замечание 2.1. В данной главе при определении формул алгебры высказываний в качестве исходной системы операций на множестве $\Omega = \{\text{И, Л}\}$ была использована система операций $\&, \vee, \to, \neg$. Однако на множестве Ω можно определить много других операций различных арностей, и потому в качестве исходной можно взять любую систему операций σ. Тогда по аналогии с определением 2.2 можно было бы определить понятие формулы в сигнатуре σ. В этом определении вместо пунктов 2, 3 нужно было бы взять следующее утверждение.

Если g — символ n-арной операции из σ и $A_1, A_2, ..., A_n$ — формулы, то $g(A_1, A_2, ..., A_n)$ — формула.

Для формул сигнатуры σ так же определяются понятия ранга и длины формулы, подформулы, значения формулы при

Таблица 1.3

ДНФ	1	2	3	4	5	6
1111	−	+	+	−	−	−
1110	−	+	+	−	−	−
1101	−	+	−	+	−	−
1100	−	+	−	+	−	−
1001	−	−	−	+	−	−
1000	−	−	−	+	−	+
0110	−	−	+	−	+	−
0011	−	−	−	−	+	−
0010	+	−	−	−	+	−
0000	+	−	−	−	−	+
0111	−	−	+	−	+	−

заданном наборе значений переменных высказываний, выполнимости, тождественной истинности, тождественной ложности формулы, эквивалентности формул и др. Ниже во второй части пособия формулы в различных сигнатурах будут использоваться при изучении дискретных функций.

Глава 3

ИСЧИСЛЕНИЕ ВЫСКАЗЫВАНИЙ

3.1. ОБЩЕЕ ПОНЯТИЕ О ЛОГИЧЕСКОМ ИСЧИСЛЕНИИ

В настоящее время большинство математических теорий строится дедуктивно. В основу теории кладется какое-либо достаточно хорошо обозримое множество основных понятий и утверждений, называемых аксиомами. Все остальные понятия определяются через основные или уже до этого определенные понятия, а все утверждения теории выводятся, как говорят, логически из аксиом или уже до этого доказанных утверждений. В любой такой теории естественно возникают вопросы: всякое ли утверждение, сформулированное в терминах данной теории, можно доказать или опровергнуть (вопрос о разрешимости), нельзя ли в ней доказать какое-либо утверждение и его отрицание (вопрос о непротиворечивости) и др. Такие вопросы можно считать корректными лишь в том случае, если будут точно определены понятие утверждения, сформулированного в терминах данной теории, и понятие доказательства. В ответ на такие потребности математики и возникли различные логические исчисления, призванные формализовать те или иные фрагменты математических теорий, а также доказательства в этих теориях.

Каждое логическое исчисление характеризуется:

1) набором используемых в нем символов, или алфавитом;

2) правилами построения из алфавита осмысленных утверждений, или формул;

3) некоторым фиксированным набором формул, называемым системой аксиом;

4) набором правил вывода, позволяющих выводить одни утверждения из других.

Алфавит, правила образования формул и само множество формул образуют язык исчисления. Язык логического исчисления должен выбираться так, чтобы с его помощью можно было записать или формализовать возможно большее число утверждений. Разные формальные языки отличаются друг от друга широтой охвата формализуемых в них утверждений или выразительностью, а также ориентацией на изучение той или иной теории.

Аксиомы и правила вывода логического исчисления позволяют выделить из множества всех формул так называемые доказуемые формулы, или теоремы. К ним относятся все аксиомы, а также формулы, которые могут быть получены из аксиом с помощью правил вывода. Если исчисление создается для обслуживания какой-либо математической теории, то естественно требовать, чтобы все доказуемые в нем формулы были формализацией истинных утверждений теории. Этот фактор должен накладывать определенные ограничения на выбор аксиом и правил вывода исчисления. В то же время набор аксиом и правил вывода должен быть достаточно богатым, чтобы с его помощью можно было доказать возможно большее число истинных утверждений теории.

Отметим еще один момент, связанный с выбором языка, аксиом и правил вывода логического исчисления. Логическое исчисление призвано, как правило, обслуживать не одну конкретную математическую теорию, а достаточно широкий класс теорий. В связи с этим в качестве аксиом такого исчисления должны выбираться формулы, отражающие утверждения, истинные в любой теории из обслуживаемого класса. Аналогично и правила вывода должны из истинных утверждений приводить к утверждениям, истинным во всех этих теориях.

Правила, определяющие содержательный смысл формул исчисления, и соответствие между понятиями доказуемости и истинности формул составляют предмет так называемой семантики логического исчисления.

Ниже мы рассмотрим одно из наиболее широко распространенных логических исчислений, называемое узким исчис-

лением предикатов, или исчислением предикатов 1-го порядка. Однако для облегчения его восприятия мы предварительно познакомимся с его простейшим фрагментом — с исчислением высказываний.

3.2. ЯЗЫК, АКСИОМЫ И ПРАВИЛА ВЫВОДА ИСЧИСЛЕНИЯ ВЫСКАЗЫВАНИЙ

Язык исчисления высказываний совпадает с языком алгебры высказываний, т. е. в нем используются те же самые символы и формулы, что и в алгебре высказываний. Разница только в том, что в исчислении высказываний формулы рассматриваются лишь как строчки символов, не имеющие каких-либо истинностных значений. Поэтому в исчислении высказываний можно обойтись лишь такими формулами из алгебры высказываний, которые не содержат постоянных высказываний. В их записи мы будем пользоваться теми же правилами сокращения скобок. Точно так же определяются подформулы и главные подформулы формул.

Аксиомы и правила вывода исчисления высказываний могут выбираться по-разному. Мы выберем систему аксиом из [46]. Аксиомы этой системы по используемым в них логическим операциям делятся на 4 подсистемы, которые мы занумеруем римскими цифрами. В них под буквами A, B, C понимаются произвольные формулы исчисления высказываний.

Аксиомы исчисления высказываний

I. 1) $A \to (B \to A)$;
   2) $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$.
II. 1) $A\&B \to A$;
   2) $A\&B \to B$;
   3) $(A \to B) \to ((A \to C) \to (A \to B\&C))$.
III. 1) $A \to A \vee B$;
   2) $B \to A \vee B$;
   3) $(A \to C) \to ((B \to C) \to (A \vee B \to C))$.
IV. 1) $A \to \overline{\overline{A}}$;

2) $\overline{\overline{A}} \to A$;
3) $(A \to B) \to (\overline{B} \to \overline{A})$.

Правило вывода исчисления высказываний

Правила вывода любого логического исчисления зачастую записывают в виде дроби, выписывая в числитель исходные формулы, а в знаменатель выводимые из них формулы.

В определении исчисления высказываний принимается лишь одно правило вывода, называемое правилом заключения. Оно имеет вид

$$\frac{A, A \to B}{B}.$$

3.3. ДОКАЗУЕМЫЕ ФОРМУЛЫ ИСЧИСЛЕНИЯ ВЫСКАЗЫВАНИЙ

Определение 3.1. *Формула A называется доказуемой, если существует конечная последовательность формул, в которой каждая формула — либо аксиома, либо получена по правилу вывода из некоторых предыдущих формул, и последняя формула совпадает с A. При этом указанная последовательность формул называется доказательством формулы A, а число ее членов — длиной доказательства.*

Тот факт, что формула A доказуема, обозначается в виде $\vdash A$.

Сделаем к определению 3.1 два замечания, полезных для упрощения записей доказательств формул.

1. Все формулы из доказательства любой формулы A — доказуемы, поскольку их доказательствами являются начальные отрезки доказательства формулы A.

2. При доказательстве формулы A, кроме аксиом и формул, получаемых по правилу заключения из предыдущих, можно использовать также любые доказуемые формулы, поскольку их доказательства можно было бы включить в качестве составных частей в доказательство формулы A.

Ниже нам придется строить доказательства ряда формул. При этом последовательности формул, являющиеся доказательствами, будем записывать (как правило) в столбик, указывая справа от формул их происхождение — номер аксиомы, правило заключения (п. з.) вместе с номерами формул, к которым оно применялось, и т. п.

Приведем простейшие, но полезные примеры на доказательство формул.

Пример 3.1. $\vdash A \to A$ при любой формуле A.

□ Построим нужную последовательность формул.

$A_1 = (A \to (B \to A)) \to ((A \to B) \to (A \to A))$ — I.2) при $C = A$;

$A_2 = A \to (B \to A)$ — I.1;

$A_3 = (A \to B) \to (A \to A)$ — п. з. A_1, A_2;

$A_4 = (A \to (B \to A)) \to (A \to A)$ — замена B на $B \to A$ в A_3;

$A_5 = A \to A$ — п. з. A_2, A_4. □

Пример 3.2. Если A — любая формула, а B — доказуемая формула, то $\vdash A \to B$.

□ Доказательство формулы $A \to B$ получается путем добавления к доказательству формулы B аксиомы I.1 $B \to (A \to \to B)$ и формулы $A \to B$, которая получается по правилу заключения из B и аксиомы I.1. □

Пример 3.3. $A \to A\&A$.

□ Достаточно взять аксиому II.3 при $B = C = A$ и дважды применить правило заключения с использованием примера 3.1. □

3.4. ВСПОМОГАТЕЛЬНЫЕ ПРАВИЛА ВЫВОДА

Определение 3.2. *Формула A называется выводимой из системы формул T, если существует конечная последовательность формул, в которой каждая формула — либо аксиома, либо формула из T, либо получена по правилу вывода из некоторых предыдущих формул, а последняя формула совпадает с A. При этом указанная последовательность формул называется выводом формулы A из системы*

T, число ее членов — длиной вывода, любая формула из T — посылкой вывода, формула A — заключением вывода.

Тот факт, что формула A выводима из T, обозначается в виде $T \vdash A$ или $B_1, ..., B_m \vdash A$, если $T = \{B_1, ..., B_m\}$.

Сравнивая определения 3.1 и 3.2, замечаем, что доказуемые формулы — это в точности те формулы, которые выводимы из пустой системы формул.

Задача построения доказательств и выводов формул в общем случае является сложной. Для облегчения этой работы доказывают сначала ряд вспомогательных правил вывода, которыми затем можно заменять целые куски доказательств или выводов. В некоторых случаях работу по выводу формул облегчает так называемая теорема дедукции.

Теорема 3.1. *Пусть T — произвольное множество формул и A, B — любые формулы. Тогда $T, A \vdash B$ в том и только том случае, когда*

$$T \vdash A \to B. \tag{3.1}$$

□ Пусть $T, A \vdash B$ и $B_1, ..., B_m = B$ — кратчайший вывод формулы B из T, A. Докажем утверждение (3.1) индукцией по m. Пусть $m = 1$. Если B — аксиома, то она доказуема, и в силу примера 3.2 имеем $\vdash A \to B$ для любой формулы A. Отсюда и следует (3.1). Если $B = A$, то (3.1) имеет место в силу примера 3.1. Если же $B \in T$, то можно указать следующий вывод формулы $A \to B$ из T:

1) $B \to (A \to B)$ — I.1;

2) B — посылка;

3) $A \to B$ — п. з. 1, 2.

Допустим, что утверждение верно при $m \leq n$ и докажем его при $m = n + 1$. Если B — аксиома или посылка, то 3.1 устанавливается точно так же, как и при $m = 1$. Пусть B получена по правилу заключения из формул B_i, B_j, где $i \leq n$, $j \leq n$. Тогда $B_j = A \to B_i$, и по предположению индукции имеем

$$T \vdash A \to B_i,$$

$$T \vdash A \to (B_i \to B).$$

Применяя теперь к формулам $A \to (B_i \to B)$, $A \to B_i$ и аксиоме I.2 дважды правило заключения, получим нужную нам формулу $A \to B$. Обратное утверждение следует непосредственно из правила заключения. $\square$

Докажем теперь ряд вспомогательных правил вывода. При этом в доказательстве правил будут использоваться и доказанные до этого вспомогательные правила вывода.

Правило умножения формул: $A, B \vdash A\&B$.
Вывод:

1) $(A \to A) \to ((A \to B) \to (A \to A\&B))$ — II.3;
2) $A \to A$ — пример 3.1;
3) $(A \to B) \to (A \to A\&B))$ — п. з. 2, 1;
4) $B \to (A \to B)$ — I.1;
5) B — посылка;
6) $A \to B$ — п. з. 5, 4;
7) $A \to A\&B$ — п. з. 6, 3;
8) A — посылка;
9) $A\&B$ — п. з. 8, 7.

Правило силлогизма: $A \to B, B \to C \vdash A \to C$.
Согласно теореме дедукции достаточно показать, что

$$A \to B, B \to C, A \vdash C.$$

Вывод:

1) A — посылка;
2) $A \to B$ — посылка;
3) B — п. з. 1, 2;
4) $B \to C$ — посылка;
5) C — п. з. 3, 4.

Правило умножения посылок: $A \to (B \to C) \vdash A\&B \to C$.
Согласно теореме дедукции достаточно доказать, что

$$A \to (B \to C), A\&B \vdash C.$$

Вывод:

1) $A\&B$ — посылка;
2) $A\&B \to A$ — II.1;

3) A — п. з. 1, 2;
4) $A \to (B \to C)$ — посылка;
5) $B \to C$ — п. з. 2, 3;
6) $A\&B \to B$ — II.2;
7) B — п. з. 1, 4;
8) C — п. з. 5, 3.

Аналогично доказываются следующие правила. Докажите их в качестве упражнений.

Правило разделения посылок: $A\&B \to C \vdash A \to (B \to C)$.

Правило перестановки посылок: $A \to (B \to C) \vdash B \to \to (A \to C)$.

Правило умножения заключений: $A \to B, A \to C \vdash A \to \to B\&C$.

Правило сложения посылок: $A \to C, B \to C \vdash A \vee B \to C$.

Правило введения посылки: $A \vdash B \to A$.

Правило контрапозиции: $A \to B \vdash \overline{B} \to \overline{A}$.

Несколько сложнее доказываются

Правила де Моргана:

а) $\overline{A} \vee \overline{B} \vdash \overline{A\&B}$;
б) $\overline{A\&B} \vdash \overline{A} \vee \overline{B}$;
в) $\overline{A}\&\overline{B} \vdash \overline{A \vee B}$;
г) $\overline{A \vee B} \vdash \overline{A}\&\overline{B}$.

Вывод:

а)
1) $A\&B \to A$ — II.1;
2) $A\&B \to B$ — II.2;
3) $\overline{A} \to \overline{A\&B}$ — пр. контрапозиции, 1;
4) $\overline{B} \to \overline{A\&B}$ — пр. контрапозиции, 2;
5) $\overline{A} \vee \overline{B} \to \overline{A\&B}$ — пр. сложения посылок, 3, 4;
6) $\overline{A} \vee \overline{B}$ — посылка;
7) $\overline{A\&B}$ — п. з., 6, 5.

б)
1) $\overline{A} \to \overline{A} \vee \overline{B}$ — III.1;
2) $\overline{B} \to \overline{A} \vee \overline{B}$ — III.2;
3) $\overline{\overline{A} \vee \overline{B}} \to \overline{\overline{A}}$ — пр. контрапозиции, 1;
4) $\overline{\overline{A} \vee \overline{B}} \to \overline{\overline{A}}$ — пр. контрапозиции, 2;

5) $\overline{\overline{A}} \to A$ — IV.2;
6) $\overline{\overline{B}} \to B$; — IV.2;
7) $\overline{\overline{A} \vee \overline{B}} \to A$ — пр. силлогизма, 3, 5;
8) $\overline{\overline{A} \vee \overline{B}} \to B$ — пр. силлогизма, 4, 6;
9) $\overline{\overline{A} \vee \overline{B}} \to A\&B$ — пр. умножения заключений, 7, 8;
10) $\overline{A\&B} \to \overline{\overline{\overline{A} \vee \overline{B}}}$ — пр. контрапозиции, 9;
11) $\overline{\overline{\overline{A} \vee \overline{B}}} \to \overline{A} \vee \overline{B}$ — IV.2;
12) $\overline{A\&B} \to \overline{A} \vee \overline{B}$ — пр. силлогизма, 10, 11.

Выводы в случаях *в*, *г* легко получаются путем замены в выводах *а*, *б* формул A, B, C на их отрицания с последующим применением закона контрапозиции и аксиом. Выпишите соответствующие выводы в качестве упражнения.

Правило заключения и указанные выше вспомогательные правила иногда называют естественными правилами вывода. В заключение данного параграфа приведем примеры на применение указанных правил.

Пример 3.4. $\vdash A\&\overline{A} \to B$ для любых формул A, B.

$\square$ $A \to (\overline{B} \to A)$ — I.1;
$(\overline{B} \to A) \to (\overline{A} \to \overline{\overline{B}})$ — IV.3;
$A \to (\overline{A} \to \overline{\overline{B}})$ — пр. силлогизма;
$A\&\overline{A} \to \overline{\overline{B}}$ — пр. умножения посылок;
$\overline{\overline{B}} \to B$ — IV.2;
$A\&\overline{A} \to B$ — пр. силлогизма. $\square$

Пример 3.5. $\vdash A\&\overline{A}$ для любой формулы A.

$\square$ 1)$\overline{A \vee \overline{A}} \to \overline{A}$ — пр. контрапозиции, III.1;
2) $\overline{A \vee \overline{A}} \to \overline{\overline{A}}$ — пр. контрапозиции, III.2;
3) $\overline{A \vee \overline{A}} \to A$ — пр. силлогизма, 2, IV.2);
4)$\overline{A \vee \overline{A}} \to A\&\overline{A}$ — пр. умножения заключений, 3,1;
5) $A\&\overline{A} \to \overline{R}$ — пример 3.4 при $B = \overline{R}$, где $\vdash R$;
6)$\overline{A \vee \overline{A}} \to \overline{R}$ — пр. силлогизма, 4,6;
7)$\overline{\overline{R}} \to \overline{\overline{A \vee \overline{A}}}$ — пр. контрапозиции, 6;
8) $R \to A \vee \overline{A}$,— пр. силлогизма, IV.1, 7, IV.2;
9)$A \vee \overline{A}$ — п. з., R, 8. $\square$

3.5. ПОЛНОТА И НЕПРОТИВОРЕЧИВОСТЬ ИСЧИСЛЕНИЯ ВЫСКАЗЫВАНИЙ

Рассмотрим вопрос о связи между тождественно истинными формулами в алгебре высказываний и доказуемыми формулами исчисления высказываний.

Теорема 3.2. *Всякая доказуемая формула исчисления высказываний является тождественно истинной в алгебре высказываний.*

□ Сначала непосредственной проверкой убеждаемся, что все аксиомы исчисления высказываний являются тождественно истинными в алгебре высказываний. Теперь докажем это для любой доказуемой формулы $A = A(x_1, ..., x_n)$ индукцией по длине $m = d(A)$ кратчайшего доказательства формулы A.

Если $m = 1$, то A — аксиома и утверждение верно. Пусть оно верно для любых формул при $m \leq k$ и $d(A) = k+1$. Относительно формулы A возможны лишь два варианта: либо она аксиома, либо получена по правилу заключения из формул B и $B \to A$ с меньшими длинами кратчайших доказательств. В связи с этим остается лишь доказать, что если $B \equiv 1$ и $B \to$ $\to A \equiv 1$, то и $A \equiv 1$. Допустим, что это не так, т. е. найдется набор α значений переменных $x_1, ..., x_n$, при которых $A(\alpha) \equiv$ $\equiv 0$. Тогда согласно определению операции импликации условие $B(\alpha) \to A(\alpha) \equiv 1$ может выполняться лишь при $B(\alpha) \equiv 0$, что противоречит условию тождественной истинности формулы B. □

Следствие 1. *Исчисление высказываний непротиворечиво, т. е. не существует доказуемой формулы A исчисления высказываний, для которой формула $\overline{A}$ также доказуема.*

□ Действительно, если $\vdash A$ и $\vdash \overline{A}$, то по теореме 3.2 $A \equiv$ И и $\overline{A} \equiv$ И, что противоречит определению операции отрицания в алгебре высказываний. □

Для теоремы 3.2 верно и обратное утверждение, которое называют теоремой о полноте исчисления высказываний относительно алгебры высказываний.

Теорема 3.3. *Исчисление высказываний является полным относительно алгебры высказываний, т. е. всякая тождественно истинная формула алгебры высказываний является доказуемой в исчислении высказываний.*

Эта теорема доказывается гораздо сложнее теоремы 3.2. Приведем общую схему доказательства.

1. Сначала сведем его к формулам, представленным в виде КНФ. Для этого нам понадобится аналог теоремы 2.3 об эквивалентности формул из алгебры высказываний и доказуемость всех равносильностей, соответствующих эквивалентностям, используемым для приведения к КНФ формул в алгебре высказываний.

2. Далее докажем, что если КНФ тождественно истинна в алгебре высказываний, то она доказуема в исчислении высказываний.

Для реализации указанной схемы доказательства нам понадобятся некоторые вспомогательные понятия и утверждения.

Лемма 3.4 (О монотонности логических операций). *Пусть A, B, C, D — любые формулы, для которых выполняются условия:*

$$\vdash A \to B\ ,\ \vdash C \to D.$$

Тогда

$$\vdash A\&C \to B\&D,\ \vdash A \vee C \to B \vee D,\ \vdash (B \to C) \to (A \to D).$$

В связи с этим говорят, что операции $\&$ и $\vee$ — монотонно возрастающие по обоим аргументам, а операция $\to$ — монотонно возрастающая по первому аргументу и монотонно убывающая по второму аргументу.

□ В зависимости от операции рассмотрим три случая.

1. Для операции $\&$. Здесь, учитывая теорему дедукции, достаточно доказать, что

$$A \to B\ ,\ C \to D\ ,\ A\&C \vdash B\&D.$$

Искомым выводом формулы $B\&D$ может служить последовательность формул

$$A\&C\ ,\ A\&C \to A\ ,\ A\&C \to C\ ,\ A\ ,\ C\ ,\ A \to \\ \to B\ ,\ B\ ,\ C \to D\ ,\ D\ ,\ B\&D\ ,$$

в которой кроме посылок используются лишь аксиомы II.1, II.2, правила заключения и умножения формул.

2. Для операции $\vee$. Здесь требуется доказать, что

$$A \to B\ ,\ C \to D \vdash A \vee C \to B \vee D.$$

Для построения искомого вывода воспользуемся аксиомой III.3, которую можно записать в виде

$$(A \to B \vee D) \to ((C \to B \vee D) \to (A \vee C \to B \vee D)).$$

Легко проверить, что выводом является последовательность формул:

$$A \to B, B \to B \vee D, A \to B \vee D\ ,$$

$$\text{III.3}),\ (C \to B \vee D) \to (A \vee C \to B \vee D),$$

$$C \to D\ ,\ D \to B \vee D\ ,\ C \to B \vee D\ ,\ A \vee C \to B \vee D.$$

3. Для операции $\to$. Здесь на основании теоремы дедукции достаточно доказать, что

$$A \to B\ ,\ C \to D\ ,\ B \to C\ ,\ A \vdash D.$$

Искомый вывод можно получить трехкратным применением правила заключения к посылкам. $\square$

Определение 3.3. *Формулы A, B называются эквивалентными, или равносильными, в исчислении высказываний, что обозначается в виде $A \sim B$, если*

$$\vdash A \to B \ и\ \vdash B \to A$$

или, что то же самое, $\vdash (A \to B)\&(B \to A)$ (в силу аксиом II.1, II.2 и правила умножения заключений).

Ниже, чтобы не путать с эквивалентностью формул в алгебре высказываний, мы здесь будем использовать термин «равносильность».

Нетрудно видеть, что введенное отношение $\sim$ является отношением эквивалентности на множестве всех формул исчисления высказываний. Действительно, рефлексивность доказывается примером 3.1, симметричность обеспечена самим определением, транзитивность следует из правила силлогизма.

Примеры равносильностей: закон двойного отрицания $\overline{\overline{A}} \sim A$, следующий из аксиом IV.2–3; законы де Моргана $\overline{A \vee B} \sim \overline{A}\&\overline{B}$ и $\overline{A\&B} \sim \overline{A} \vee \overline{B}$, следующие из доказанных выше правил де Моргана.

Лемма 3.5 (О равносильности формул). *Пусть A — формула, содержащая переменное x; B_1, B_2 — любые формулы и $A(B_1)$, $A(B_2)$ — формулы, полученные заменой в A каких-либо одних и тех же вхождений буквы x соответственно на B_1 и B_2. Тогда*

$$(B_1 \sim B_2) \vdash (A(B_1) \sim A(B_2)).$$

□ Докажем теорему индукцией по рангу r формулы A. Если $r = 0$, то $A = x$ и утверждение теоремы принимает вид

$$\vdash (B_1 \sim B_2) \to (B_1 \sim B_2).$$

Его справедливость установлена в примере 3.1. Допустим, что теорема верна при $r \leq m$, и пусть $r(A) = m+1$. Тогда формула A имеет одну из следующих форм:

$$A_1\&A_2,\ A_1 \vee A_2,\ A_1 \to A_2,\ \overline{A_1}.$$

Рассмотрим случай, когда $A = A_1\&A_2$. Пусть $B_1 \sim B_2$. Тогда по предположению индукции имеем четыре доказуемых формулы:

$$(A_1(B_1) \to A_1(B_2))\ ,\ (A_1(B_2) \to A_1(B_1)),$$

$$(A_2(B_1) \to A_2(B_2)),\ (A_2(B_2) \to A_2(B_1)).$$

Применяя лемму о монотонности операции $\&$ сначала к 1-й и 3-й, а затем к 2-й и 4-й формулам, получим две импликации, из которых по правилу умножения формул и получается нужная равносильность

$$(A_1(B_1)\&A_2(B_1) \sim (A_1(B_2) \sim A_2(B_2))),$$

т. е. $(A(B_1) \sim A(B_2))$.

Аналогично доказывается теорема и в случаях, когда $A = A_1 \vee A_2$ и $A = A_1 \to A_2$. В случае $A = \overline{A_1}$ утверждение теоремы следует непосредственно из правила контрапозиции. $\square$

Следствие 1. *Если в формуле A какие-либо вхождения подформулы B заменить сначала формулой C_1, а затем равносильной ей формулой C_2, то получатся равносильные формулы.*

$\square$ Достаточно заменить сначала в формуле A указанные вхождения формулы B на переменное x, а затем в полученную формулу вместо этих вхождений x подставить C_1, C_2 и применить лемму о равносильности формул. $\square$

Теперь установим равносильности, необходимые для приведения формул к равносильным им КНФ. Прослеживая доказательства теорем 2.5, 2.6 о приведении формул к КНФ в алгебре высказываний, замечаем, что в них использовались законы коммутативности, ассоциативности и дистрибутивности для операций $\&$, $\vee$, законы де Моргана, двойного отрицания, противоречия, исключенного третьего и эквивалентность 17: $A \to B \equiv \overline{A} \vee B$. Нужные нам равносильности соответствуют некоторым из эквивалентностей 1–17 алгебры высказываний, и мы сохраним для них те же самые номера.

1) $A\&B \sim B\&A$.

Применяя к аксиомам $A\&B \to B$, $A\&B \to A$ правило умножения заключений, получим $\vdash A\&B \to B\&A$. Аналогично доказывается и обратная импликация.

2) $A \vee B \sim B \vee A$.

Применяя к аксиомам $A \to B \vee A$, $B \to B \vee A$ правило сложения посылок, получим $\vdash A \vee B \to B \vee A$. Аналогично доказывается и обратная импликация.

Сходным образом доказываются равносильности 3, 4, соответствующие ассоциативности операций & и $\vee$. Проделайте это в качестве упражнения.

Всюду далее в этой главе свойствами коммутативности и ассоциативности операций мы будем пользоваться без ссылок.

17) $A \to B \sim \overline{A} \vee B$.

Выпишем сначала доказательство импликации

$$(A \to B) \to \overline{A} \vee B.$$

$\overline{A} \to \overline{A} \vee B$ — III.1;
$\overline{A} \vee B \to ((A \to B) \to \overline{A} \vee B)$ — I.1;
$\overline{A} \to ((A \to B) \to \overline{A} \vee B)$ — пр. силлогизма;
Аналогично, заменив аксиому III.1 на III.2, получим
$A \to ((A \to B) \to \overline{A} \vee B$;
$A \vee \overline{A} \to ((A \to B) \to \overline{A} \vee B)$ — пр. сложения посылок;
$A \vee \overline{A}$ — пример 3.5;
$(A \to B) \to \overline{A} \vee B$ — п. з.
Теперь докажем обратную импликацию.
$A\&\overline{A} \to B$ — пример 3.4;
$(A \to (\overline{A} \to B)$ — пр. разделения посылок;
$A \to (\overline{A} \vee B)$ — пр. перестановки посылок;
$B \to (A \to B)$ — I.1;
$(\overline{A} \vee B) \to (A \to B)$ — пр. сложения посылок. □

Для доказательства законов дистрибутивности 5, 6 нам необходимо доказать четыре импликации:

а) $A\&(B \vee C) \to (A\&B) \vee (A\&C)$;

б) $(A\&B) \vee (A\&C) \to A\&(B \vee C)$;

в) $A \vee B\&C \to (A \vee B)\&(A \vee C)$;

г) $(A \vee B)\&(A \vee C) \to A \vee B\&C$.

Докажем сначала формулу *б)*. Используя аксиомы II.1, II.2 правила силлогизма и умножения заключений, получим последовательность $A\&B \to A$, $A\&B \to B$, $B \to B \vee C$, $A\&B \to$ $\to B \vee C$, $A\&B \to A\&(B \vee C)$. Аналогично доказывается формула $A\&C \to A\&(B \vee C)$. Теперь, применяя к двум последним формулам правило сложения посылок, получим *б)*.

Так как формула *б)* доказана для любых формул A, B, C, то доказуема и формула, полученная из *б)* заменой A, B, C на их отрицания. Теперь, применив к полученной формуле закон контрапозиции и воспользовавшись леммой 3.5, правилами де Моргана и двойного отрицания, получим формулу *в)*.

Для доказательства формулы *г)* перейдем сначала от нее к равносильной ей формуле (воспользовавшись равносильностью 17)

$$(\overline{A} \to B)\&(\overline{A} \to C) \to \overline{A} \to B\&C.$$

Последняя формула доказуема, так как получается по правилу умножения посылок из аксиомы III.3.

Формула *а)* получается из *г)* по той же схеме, что и *в)* из *б)*.

□ Доказательство теоремы 3.3. Пусть A — формула исчисления высказываний, и $A \equiv 1$ в алгебре высказываний. Приведем ее по схеме доказательства теоремы 2.6 к эквивалентной ей КНФ B. Из теоремы о равносильности и доказанных выше равносильностей следует, что $B \sim A$, и потому $\vdash A$ тогда и только тогда, когда $\vdash B$. Так как B есть произведение некоторых элементарных дизъюнкций $B_1, ..., B_k$, то на основании правила умножения формул для доказательства формулы B достаточно доказать каждую из формул B_i, $i = 1, ..., k$. Из условия $A \equiv 1$ и теоремы 2.8 следует, что в каждую дизъюнкцию B_i входит некоторое переменное и его отрицание. Следовательно, B_i равносильна формуле вида $x \vee \overline{x} \vee C$. Эта формула имеет следующее доказательство:

$x \vee \overline{x}$ — пример 3.5;

$x \vee \overline{x} \to x \vee \overline{x} \vee C$ — III.1;

$x \vee \overline{x} \vee C$ — пр. заключения. □

Определение 3.4. *Непротиворечивое логическое исчисление называется полным в узком смысле, если присоединение к его аксиомам любой одной не доказуемой в нем формулы приводит к противоречивому логическому исчислению.*

Теорема 3.6. *Исчисление высказываний является полным в узком смысле.*

□ Пусть A — любая недоказуемая формула исчисления высказываний. Не теряя общности, можно считать, что она является КНФ. Если каждая из составляющих ее элементарных дизъюнкций доказуема, то согласно правилу умножения формул доказуема и формула A. Следовательно, найдется недоказуемая элементарная дизъюнкция C. Согласно теореме 3.2 она не является тождественно истинной в алгебре высказываний. Следовательно, найдется такой набор $\alpha = (a_1, ..., a_n)$ значений входящих в нее переменных $x_1, ..., x_n$, при котором каждое слагаемое в C и сама формула C принимают значение «Л». Выберем любую доказуемую в исчислении высказываний формулу R и заменим в C переменное x_i на R, если $a_i \equiv$ И, и на $\overline{R}$, если $a_i \equiv$ Л. Легко видеть, что получится формула B, равносильная $\overline{R}$ в исчислении высказываний. Это следует непосредственно из равносильности $\vdash A \vee A \sim A$. Добавим теперь формулу A к аксиомам исчисления высказываний. Получим новое логическое исчисление. Ясно, что в нем доказуемы все фомулы, доказуемые в исчислении высказываний. Поэтому в нем $\vdash B \equiv \overline{R}$, а значит и $\vdash B \to \overline{R}$. А так как в нем A — аксиома, то B получена подстановкой в доказуемую формулу C, и следовательно, доказуема. Отсюда по правилу заключения получим $\vdash \overline{R}$. Таким образом, в новом исчислении доказуемы формулы R и $\overline{R}$, что и означает, что оно противоречиво. □

Глава 4

АЛГЕБРА ПРЕДИКАТОВ

4.1. ПРЕДИКАТЫ И ОПЕРАЦИИ НАД НИМИ

В математике при изучении того или иного множества M зачастую используют предложения, содержащие символы переменных со значениями из M и превращающиеся в истинные или ложные высказывания при замене символов переменных элементами из M. Приведем примеры, взяв в качестве M множество натуральных чисел $\mathbb{N}$ и буквы x, y, z в качестве символов переменных.

Рассмотрим предложение

$$<< x \text{ есть простое число} >>,$$

обозначив его ради краткости через $p(x)$. Ясно, что это предложение не является высказыванием, поскольку о нем нельзя сказать, истинно оно или ложно. При $x = 1$ оно ложно, при $x = 2, 3$ — истинно, при $x = 4$ — ложно и т. д. Предложение $p(x)$ естественнее считать функцией от x, которая при каждом конкретном значении $x \in \mathbb{N}$ становится высказыванием. Такие функции принято называть предикатами. Другими примерами предикатов на множестве $\mathbb{N}$ могут служить предложения:

1) «x есть число, кратное 5»;
2) «Число x не превосходит 7»;
3) «Число x имеет хотя бы один делитель»;
4) «Число x удовлетворяет уравнению $x^2 + x - 1 = 0$»;
5) «Произведение чисел x, y больше ста»;

6) «Число x делит число y»;

7) «Число z является суммой чисел x, y»;

8) «Число z заключено между числами x, y»

и т. д.

По числу переменных, участвующих в предложении, различают 1-местные (или унарные) предикаты, 2-местные (или бинарные) предикаты, 3-местные (или тернарные) предикаты и т. д.

В общем случае под n-местным (или n-арным) предикатом на произвольном множестве M понимают всякое предложение, которое содержит n различных переменных, принимающих значения из M, и превращается в высказывание при замене переменных произвольными элементами из M.

В дальнейшем мы не будем исключать и случай $n = 0$, понимая под 0-местным предикатом произвольное фиксированное высказывание.

n-местные предикаты, содержащие символы переменных $x_1, ..., x_n$, будем обозначать в виде

$$p(x_1, ..., x_n), q(x_1, ..., x_n), ...,$$

или, короче, $p, q, ...$, если их арности и переменные определены контекстом.

Значение α (1 или 0) высказывания, получающегося при замене в предикате $p(x_1, ..., x_n)$ переменных $x_1, ..., x_n$ соответственно элементами $a_1, ..., a_n$, называют значением предиката в точке $(a_1, ..., a_n)$ или при $x_1 = a_1, ..., x_n = a_n$, записывая это в виде

$$p(a_1, ..., a_n) \equiv \alpha.$$

Таким образом, с каждым n-местным предикатом p на M сопоставляется отображение множества M^n в двухэлементное множество $\Omega = \{1,0\}$, или, как говорят, n-местная логическая функция. Предикат p зачастую отождествляют с сопоставляемой ему логической функцией, в связи с чем появляется возможность дать более строгое определение предиката как отображения M^n в Ω. Нам в данном параграфе будет удобнее придерживаться менее строгой точки зрения, понимая под предикатом предложение с переменными.

Укажем еще на связь между понятиями n-арного предиката и n-арного отношения на множестве M. По n-арному предикату p естественным образом определяется n-арное отношение R:

$$(a_1, ..., a_n) \in R \Longleftrightarrow p(a_1, ..., a_n) \equiv 1.$$

Тем самым устанавливается взаимно однозначное соответствие между множествами всех n-арных отношений и всех n-арных предикатов на множестве M.

Множество M с системой определенных на нем предикатов σ называется моделью сигнатуры σ и обозначается в виде $M(\sigma)$.

Опишем некоторые способы, позволяющие получать из одних предикатов на множестве M другие предикаты на том же множестве.

1. Пусть $p(x_1, ..., x_n)$ — произвольный предикат на M. Заменив в нем x_1 некоторым элементом $a \in M$, мы получим новый, $(n-1)$-местный предикат на M, который будем обозначать в виде

$$p(a, x_2, ..., x_n).$$

Так, если $p(x, y, z)$ есть трехместный предикат «$x + y = z$» на $\mathbb{N}$, то $p(2, y, z)$ есть двухместный предикат «$2 + y = z$», который можно записать также в виде предложения «Число z на две единицы больше числа y».

Аналогично новые предикаты можно получать из предиката $p(x_1, ..., x_n)$, заменяя в нем какую-либо другую переменную или даже несколько переменных элементами из M. Ясно, что заменив k переменных, мы получим $(n-k)$-местный предикат.

2. Заменив в предикате $p(x_1, ..., x_n)$ при $n \geq 2$ x_1 на x_2 (или, как говорят, отождествив переменные x_1 и x_2), мы получим новый предикат, который обозначим через

$$p(x_2, x_2, ..., x_n).$$

Ясно, что этот предикат уже содержит $n-1$ различных переменных, а потому является $(n-1)$-местным предикатом.

Например, отождествляя в предикате «$x + y = z$» на $\mathbb{N}$ переменные x, y, получим двухместный предикат «$y + y = z$»

или «$2y = z$», который можно записать также предложением «Число z в два раза больше числа y».

Аналогично можно получать из предиката $p(x_1, ..., x_n)$ новые предикаты, отождествляя какие-либо другие переменные.

3. Учитывая связь понятия предиката с понятием высказывания, можно определить логические операции для предикатов. Так, соединяя два предиката (предложения) p, q союзом «или», мы получим новый предикат, который обозначим через $p \vee q$ и назовем дизъюнкцией предикатов p, q. Если p — n-местный, q — m-местный предикаты, а переменные, входящие в p, не входят в q, то $p \vee q$ будет $(m+n)$-местным предикатом. Его значение при конкретных значениях переменных равно дизъюнкции соответствующих значений предикатов p, q.

Аналогично определяются конъюнкция и импликация предикатов, а также отрицание предиката.

4. Кроме операций $\&$, $\vee$, $\rightarrow$, $\neg$, для предикатов на множестве можно определить еще логические операции навешивания кванторов всеобщности и существования. Пусть $p(x)$ — одноместный предикат на M, т. е. предложение, содержащее переменную x, принимающую значения из M. Построим из него новое предложение, добавив перед ним фразу «Для всех x» и обозначив новое предложение через

$$\forall x \; p(x). \tag{4.1}$$

Из смысла полученного предложения видно, что оно является высказыванием, которое истинно, если предикат $p(x)$ принимает значение 1 при любом значении переменной x из M, и ложно, если $p(x)$ принимает значение 0 хотя бы при одном значении x из M. Символ $\forall$ называют квантором всеобщности и говорят, что высказывание (4.1) получено из предиката $p(x)$ навешиванием квантора всеобщности по переменной x.

Рассмотрим теперь n-местный предикат $p(x_1, ..., x_n)$ на множестве M. Добавив к нему фразу «Для всех x_1» или «Для всякого x_1», получим новое предложение, которое обозначим в виде

$$\forall x_1 \; p(x_1, ..., x_n). \tag{4.2}$$

Из построения этого предложения видно, что при замене в нем переменных $x_2, ..., x_n$ соответственно элементами $a_2, ..., a_n \in \in M$ получится высказывание

$$\forall x_1 \; p(x_1, a_2, ..., a_n),$$

которое истинно в том и только том случае, когда высказывание $p(a, a_2, ..., a_n)$ истинно при любом $a \in M$. Таким образом, (4.2) является $(n-1)$-местным предикатом. Говорят, что он получен из предиката p навешиванием квантора всеобщности по переменной x_1. Понятно, что квантор $\forall$ можно навешивать и по другим переменным.

Добавляя перед предикатом $p(x_1, ..., x_n)$, как и перед предложением с переменными $x_1, ..., x_n$, фразу «Существует x_1, такое, что», мы получим новое предложение, которое обозначают в виде

$$\exists x_1 \; p(x_1, ..., x_n). \tag{4.3}$$

Подставив в него элементы $a_2, ..., a_n$ вместо $x_2, ..., x_n$, получим высказывание

$$\exists x_1 \; p(x_1, a_2, ..., a_n),$$

которое истинно в том и только том случае, когда высказывание $p(a, a_2, .., a_n)$ истинно хотя бы при одном a из M. Следовательно, предложение (4.3) есть $(n-1)$-местный предикат на M.

Символ $\exists$ называется квантором существования, а о предложении (4.3) говорят, что оно получено из предиката $p(x_1, ..., x_n)$ навешиванием квантора существования по переменной x_1. Квантор существования можно навешивать и по другим переменным.

Приведем примеры. Пусть $p(x, y)$ есть предикат на множестве $\mathbb{N}$: «x делит y». Тогда предложение $\forall y \; p(x, y)$ зависит только от переменной x. При $x = 1$ оно принимает значение 1, поскольку 1 делит любое натуральное число. При любом другом значении x из $\mathbb{N}$ оно принимает значение 0, поскольку неверно, что x делит все натуральные числа.

Предложение $\forall x \; p(x, y)$ зависит только от переменной y и принимает значение 0 при любом значении y, поскольку в $\mathbb{N}$ не существует чисел, делящихся на все натуральные числа.

С помощью рассмотренных выше способов можно из некоторой исходной системы предикатов на множестве M получать все новые и новые предикаты, записываемые в виде различных выражений через исходные предикаты, символы переменных, элементов из M, логических операций $\&$, $\vee$, $\to$, $\neg$, $\forall$, $\exists$ и служебных символов — скобок и запятых. Такие выражения называются формулами. Ниже будет дано точное (индуктивное) определение формулы.

4.2. ФОРМУЛЫ АЛГЕБРЫ ПРЕДИКАТОВ

Зафиксируем произвольную алгебраическую систему (M, Σ). Ее сигнатуру Σ представим в виде объединения $\Sigma = F \cup P$, где F — конечное множество символов операций, а P — конечное непустое множество символов предикатов на M. В частности, множество F может быть и пустым. Обозначим через F_0 подмножество из F обозначений всех нуль-арных операций, т. е. выделенных элементов множества M. Оно может быть любым подмножеством из M. Не исключается также случай, когда P содержит символы нуль-арных предикатов. Под нуль-арным предикатом понимается постоянное (истинное или ложное) высказывание. Множество таких предикатов обозначим через P_0.

При определении формул алгебры предикатов в качестве символов алфавита будут использоваться различные буквы (возможно, с индексами): a, b, c, d для элементов из F_0; f, ϕ, ψ для элементов из F; p, q, r для элементов из P. Кроме того, будет использоваться счетное множество X символов предметных переменных со значениями из M, обозначаемых буквами x, y, z, u, v (возможно, с индексами); логические операции $\&, \vee, \to, \neg, \forall, \exists$ и служебные символы — скобки и запятые. Весь алфавит алгебры предикатов обозначим буквой $\mathcal{A}$.

В конкретных примерах для операций и предикатов будут использоваться также общепринятые обозначения: $+$, $\oplus$, $\cdot$, $=$, $\leq$ и т. д. При этом знак равенства $=$ в соотношениях вида $A = B$ всегда будет пониматься в смысле совпадения A и B.

Определим предварительно понятие терма на системе (M, Σ).

Определение 4.1. *1. Каждый символ предметного переменного из X или константы из F_0 есть терм.*

2. Если f — символ n-арной операции из Σ и $t_1, ..., t_n$ — термы, то слово $f(t_1, ..., t_n)$ есть терм.

3. Других термов нет.

Если вместо переменных из X в терм t подставить элементы из M и произвести все указанные в t операции, то мы получим элемент из M, называемый значением терма t. Следовательно, терм t определяет функцию на M, зависящую от тех переменных, которые входят в t. Так, если $M = \mathbb{N}$ и $\Sigma = \{0, 1, +, \cdot\}$, то термами могут быть представлены все многочлены над $\mathbb{N}$.

Теперь определим понятие формулы. В формулах нам необходимо будет различать свободные и связанные вхождения переменных. Эти понятия удобнее определить параллельно с определением формулы.

Определение 4.2. *1. Любой символ нуль-местного предиката из P_0 есть формула.*

2. Если p — символ n-местного предиката из P, а $t_1, ..., t_n$ — термы, то слово $p(t_1..., t_n)$ есть формула.

Формулы, определенные в пунктах 1–2, называются элементарными. Все вхождения предметных переменных в элементарные формулы называются свободными.

3. Если A и B — формулы, то слова

$$(A)\&(B),\ (A) \vee (B),\ (A) \to (B),\ \neg(A)$$

также являются формулами. В них каждое вхождение предметной переменной является вхождением в формулу A или B и считается таким же (свободным или связанным), каким оно было соответственно в A или B.

4. Если A — формула, в которой есть свободное вхождение предметной переменной x_i, то слова

$$\forall x_i(A) \text{ и } \exists x_i(A)$$

являются формулами, в которых все вхождения предметных переменных, отличных от x_i, называются так же, как

и в A, а все вхождения переменной x_i называются связанными. При этом формула A называется областью действия записанного перед ней квантора $\forall$ или $\exists$ по переменной x_i.

5. Других формул нет.

Замечание. В приведенном определении формулы на алгебраической системе (M, Σ) от M, по существу, использовалось лишь множество выделенных элементов P_0. Поэтому формулу на (M, Σ) можно рассматривать также и как формулу на любой другой алгебраической системе сигнатуры Σ. В связи с этим формулы на алгебраической системе (M, Σ) называют просто формулами алгебры предикатов сигнатуры Σ.

Число всех логических операций, участвующих в записи формулы A, назовем рангом формулы A и обозначим через $r(A)$. В частности, $r(A) = 0$ в том и только том случае, когда A — элементарная формула.

Договоримся далее вместо $\neg(A)$ писать $\overline{A}$.

Рассмотрим пример. Пусть p_1, p_2 — символы двухместных предикатов. Тогда выражение

$$\forall x_1((\exists x_1(p_1(x_2, x_1))) \vee (\forall x_2(\overline{(p_2(x_1, x_2))}))) \qquad (4.4)$$

есть формула, поскольку $p_1(x_2, x_1)$, $p_2(x_1, x_2)$ являются формулами по пункту 2 определения 4.2, это элементарные формулы. Тогда $\overline{(p_2(x_1, x_2))}$ есть формула по пункту 3, $\exists x_1(p_1(x_2, x_1))$ и $\forall x_2(\overline{(p_2(x_1, x_2))})$ есть формулы по пункту 4, далее

$$(\exists x_1(p_1(x_2, x_1))) \vee (\forall x_2(\overline{(p_2(x_1, x_2))}))$$

есть формула по пункту 3. В ней последнее вхождение x_1 свободно, а потому (4.4) есть формула по пункту 4.

Заметим, что в формуле (4.4) все вхождения переменной x_1, связанные, первое вхождение переменной x_2 свободно, остальные — связанные.

Число скобок при записи формулы можно уменьшить, если принять все правила сокращения скобок из алгебры высказываний и кроме того условиться:

1) не заключать в скобки элементарные формулы;

2) считать, что операции навешивания кванторов сильнее всех других операций;

3) записывать формулу $\delta_1 x_1(\delta_2 x_2(...\delta_k x_k A)...))$ с кванторами $\delta_1, ..., \delta_k$ в виде $\delta_1 x_1 \delta_2 x_2 ... \delta_k x_k A$ и в виде $\delta x_1, x_2, ..., x_k A$ при $\delta_1 = ... = \delta_k = \delta$.

При указанных соглашениях формулу (4.4) можно записать в виде

$$\forall x_1(\exists x_1 p_1(x_2, x_1) \vee \forall x_2 \overline{p_2(x_1, x_2)}).$$

В дальнейшем для сокращения записей иногда будем опускать знак &, т. е. вместо $A\&B$ писать AB.

Если A — формула сигнатуры Σ, то любое ее подслово, являющееся формулой, называют подформулой формулы A. Подробнее понятие подформулы можно определить индукцией по рангу формулы.

Определение 4.3. *1. Подформулой элементарной формулы A называется лишь сама формула A.*

2. Подформулами любой формулы вида

$$A\&B\ ,\ A \vee B\ ,\ A \to B$$

называются сама эта формула и все подформулы формул A и B.

3. Подформулами любой из формул

$$\overline{A}\ ,\ \forall x A\ ,\ \exists x A$$

называются сама эта формула и все подформулы формулы A.

Пользуясь определением 4.3, нетрудно выписать все подформулы любой заданной формулы. Так, подформулами формулы (4.4) будут она сама и формулы:

$$p_1(x_2, x_1),\ p_2(x_1, x_2),\ \exists x_1 p_1(x_2, x_1),\ \overline{p_2(x_1, x_2)},$$

$$\forall x_2 \overline{p_2(x_1, x_2)},\ \exists x_1 p_1(x_2, x_1) \vee \forall x_2 \overline{p_2(x_1, x_2)}.$$

Из рассмотренных в 4.1 правил получения предикатов на множестве M из заданных предикатов видно, что любая формула A алгебры предикатов сигнатуры Σ является предикатом на алгебраической системе (M, Σ), зависящим лишь от тех предметных переменных, которые имеют свободные вхождения в A. Если такими переменными являются $x_1, ..., x_n$, то формулу A иногда записывают в виде $A(x_1, ..., x_n)$. Заменив свободные вхождения переменных в формуле A элементами из M (одинаковыми для всех вхождений одной переменной), мы получим высказывание, значение которого можно вычислить, исходя из значений элементарных высказываний и используя определения логических операций. Значения этих высказываний называются значениями формулы A на (M, Σ) при соответствующих значениях переменных.

Определение 4.4. *Формула A алгебры предикатов сигнатуры Σ называется выполнимой на алгебраической системе (M, Σ), если она принимает значение 1 хотя бы при одном наборе значений из M для переменных, имеющих свободные вхождения в A. В противном случае формула A называется ложной на (M, Σ). Формула A называется истинной на (M, Σ), если она принимает значение 1 при любых наборах значений из M для переменных, имеющих свободные вхождения в A.*

Определение 4.5. *Формула алгебры предикатов сигнатуры Σ называется выполнимой, тождественно истинной или тождественно ложной, если она соответственно выполнима хотя бы на одной алгебраической системе, истинна на всех системах или ложна на всех системах сигнатуры Σ.*

Тождественную истинность (ложность) формулы A обозначают в виде $A \equiv 1$ ($A \equiv 0$).

Примеры.

1. $\overline{A \& B \to A} \equiv$ Л для любых формул A, B.
2. Формула

$$\overline{x_1 x_1 = x_1} \& \forall x_2, x_3((x_1 = x_2 x_3) \to \\ \to (x_1 = x_2) \vee (x_1 = x_3)) \quad (4.5)$$

выполнима, но не истинна на алгебраической системе $(\mathbb{N}, \cdot; =)$. Легко видеть, что она зависит лишь от x_1 и принимает истинное значение в том и только том случае, когда значением переменной x_1 является простое число.

3. Формула

$$\forall x_1, x_2 \, (x_1 x_2 = x_2 x_1) \tag{4.6}$$

выполнима, так как она выполнима, например, на системе $(\mathbb{N}, \cdot; =)$. Однако она не тождественно истинна, поскольку существуют алгебраические системы с некоммутативной операцией умножения.

Примеры 2–3 показывает, что формулами алгебры предикатов можно записывать те или иные свойства элементов алгебраических систем или характеризовать те или иные классы алгебраических систем. Так, формула (4.5) характеризует подмножество всех простых чисел из множества $\mathbb{N}$, формула (4.6) выделяет класс алгебраических систем с коммутативной операцией умножения.

Зачастую запись свойства с помощью формулы является более короткой, а главное, более прозрачной по сравнению со словесной. Этим и объясняется широкое использование языка алгебры предикатов в самых различных областях математики. При этом, кроме указанных здесь, допускаются различные другие упрощения в записях формул. Так, условие непрерывности функции $f(x)$ в точке x_0 записывают в виде формулы

$$\forall \epsilon > 0 \, \exists \delta > 0 : \, (|x - x_0| < \delta \to |f(x) - f(x_0)| < \epsilon).$$

4.3. ЭКВИВАЛЕНТНОСТЬ ФОРМУЛ. ОСНОВНЫЕ СООТНОШЕНИЯ ЭКВИВАЛЕНТНОСТИ

Определение 4.6. *Формулы A, B алгебры предикатов сигнатуры Σ называются эквивалентными на алгебраической системе $M(\sigma)$, если они принимают на (M, Σ) одинаковые значения при любом наборе значений предметных переменных, имеющих свободные вхождения в A или B.*

Из определения 4.6 видно, что эквивалентность тех или иных формул сигнатуры Σ зависит от свойств алгебраической

системы (M, Σ). Например, формулы $\forall x A$ и $\exists x A$ эквивалентны на любой одноэлементной системе, однако не эквивалентны в общем случае. Можно привести и менее тривиальные примеры. Изучение эквивалентностей, имеющих место для отдельных конкретных алгебраических систем, не отвечает целям и задачам математической логики как науки об общих закономерностях в рассуждениях. В связи с этим более ценным является следующее понятие эквивалентности.

Определение 4.7. *Формулы алгебры предикатов сигнатуры* Σ *называются эквивалентными, если они эквивалентны на любой алгебраической системе сигнатуры* Σ.

Эквивалентность формул A, B, как и эквивалентность высказываний, будем обозначать в виде $A \equiv B$.

Отметим следующие очевидные свойства эквивалентности формул алгебры предикатов.

1. Отношение $\sim$ является отношением эквивалентности на множестве всех формул сигнатуры Σ, и, следовательно, все указанные формулы разбиваются на классы эквивалентных формул.

2. Если формула A' получена из A заменой некоторой подформулы B эквивалентной ей формулой B', то $A' \equiv A$.

Легко видеть, что для формул алгебры предикатов выполняются эквивалентности 1–18, приведенные в параграфе 2.3. Дополним этот список эквивалентностями, связанными с операциями навешивания кванторов.

Правила перестановки одноименных кванторов:

19) $\forall x A(x) \equiv \forall y A(y)$;

20) $\forall x \forall y A \equiv \forall y \forall x A$;

21) $\exists x \exists y A \equiv \exists y \exists x A$.

Правила отрицания кванторов:

22) $\overline{\forall x A} \equiv \exists x \overline{A}$;

23) $\overline{\exists x A} \equiv \forall x \overline{A}$.

Законы дистрибутивности кванторов $\forall$, $\exists$ относительно операций $\&$ и $\vee$ (соответственно):

24) $\forall x (A \& B) \equiv \forall x A \,\&\, \forall x B$;

25) $\exists x (A \vee B) \equiv \exists x A \vee \exists x B$.

Правило расширения области действия кванторов:

26) $\delta x A * B \equiv \delta x(A * B)$, где δ — квантор $\forall$ или $\exists$, $*$ — операция $\&$ или $\vee$ и формула B не содержит свободных вхождений x.

Правило переименования свободных переменных:

27) $\delta x A(x) \equiv \delta y A(y)$, где δ — квантор $\forall$ или $\exists$, $A(x)$ — формула, не содержащая буквы y, $A(y)$ — формула, полученная из $A(x)$ заменой всех свободных вхождений x на y.

Эквивалентности 1–27 являются простыми следствиями свойств логических операций. Они выполняются для любых систем. Докажите это в порядке упражнения. Их называют основными законами логики предикатов. Они постоянно (явно или неявно) используются при доказательствах утверждений во всех разделах математики.

Приведем два замечания, предостерегающие от ошибок в преобразованиях формул.

1. В общем случае нельзя переставлять разноименные кванторы. Например, на системе $(\mathbb{N}, <)$ формула $\forall x \exists y(x < y)$ истинна, а формула $\exists y \forall x(x < y)$ ложна.

2. В общем случае нельзя использовать правила дистрибутивности квантора $\forall$ относительно операции $\vee$ и квантора $\exists$ относительно операции $\&$. Покажите в качестве упражнения, что формулы

$$\forall x(A \vee B) \text{ и } \forall x A \vee \forall x B,$$

а также формулы

$$\exists x(A \& B) \text{ и } \exists x A \& \exists x B$$

в общем случае не эквивалентны.

В алгебре предикатов, как и в алгебре высказываний, вводится понятие двойственных формул.

Определение 4.8. *Пусть A — формула алгебры предикатов, не содержащая операции «$\to$». Формула, полученная из A заменой всех вхождений $\vee$ на $\&$, $\&$ на $\vee$, $\forall$ на $\exists$ и $\exists$ на $\forall$, называется двойственной к A и обозначается через A^*.*

Очевидно, что $(A^*)^* = A$, и потому формулы A и A^* называются двойственными. Двойственными называются и взаимозаменяемые операции $\vee$ и $\&$, а также кванторы $\forall$ и $\exists$.

Теорема 4.1. *Пусть A — формула алгебры предикатов и $p_1, ..., p_s$ суть все различные элементарные подформулы в A, короче:*

$$A = A(p_1, ..., p_s).$$

Тогда имеет место эквивалентность формул:

$$A^*(p_1, ..., p_s) \equiv \overline{A}(\overline{p_1}, ..., \overline{p_s}).$$

□ Докажем теорему индукцией по рангу r формулы A. При $r = 0$ ее утверждение верно. Допустим, что оно верно для любой формулы ранга $r \leq n$ и пусть $r(A) = n + 1$.

Возможны следующие пять случаев:

$$A = A_1 \& A_2\ ,\ A = A_1 \vee A_2\ ,\ A = \forall x A_1\ ,\ A = \exists x A_1\ ,\ A = \overline{A_1}.$$

1. $A = A_1 \& A_2$. Тогда, используя предположение индукции, закон де Моргана 11 и общие свойства эквивалентности, получим

$$\begin{aligned} A^*(p_1, ..., p_s) &= A_1^*(p_1, ..., p_s) \vee A_2^*(p_1, ..., p_s) \equiv \\ &\equiv \overline{A_1}(\overline{p_1}, ..., \overline{p_s}) \vee \overline{A_2}(\overline{p_1}, ..., \overline{p_s}) \equiv \overline{A}(\overline{p_1}, ..., \overline{p_s}). \end{aligned}$$

В трех следующих случаях рассуждения аналогичны. Вместо эквивалентности 11 в них используются соответственно эквивалентности 12, 21, 22, 14. □

Следствие 1 (Принцип двойственности). *Для любых формул A, B алгебры предикатов:*

$$A \equiv B \Longleftrightarrow A^* \equiv B^*.$$

Принцип двойственности позволяет вместо двух эквивалентностей $A \equiv B$ и $A^* \equiv B^*$ доказывать любую одну из них.

4.4. ПРИВЕДЕННЫЕ И ПРЕДВАРЕННЫЕ ФОРМУЛЫ

Эквивалентности 1–27 зачастую используются также для преобразования формул к эквивалентным им формулам нужного вида.

В качестве примеров рассмотрим преобразование формул алгебры предикатов к так называемым приведенным и предваренным формулам.

Определение 4.9. *Формула алгебры предикатов называется приведенной, если в ней не используется операция* $\to$*, а отрицание или не используется совсем, или относится лишь к элементарным формулам.*

Определение 4.10. *Предваренной (или нормальной, или пренексной) формулой алгебры предикатов называется любая формула вида*

$$\delta_1 x_{i_1} \delta_2 x_{i_2} ... \delta_k x_{i_k} A, \tag{4.7}$$

где $\delta_1, ..., \delta_k$ *— кванторы, а* A *— приведенная формула, не содержащая кванторов.*

Теорема 4.2. *Для всякой формулы* A *алгебры предикатов существует эквивалентная ей приведенная формула.*

□ Докажем теорему индукцией по рангу $r(A)$ формулы A. Если $r(A) = 0$, то утверждение очевидно. Пусть $r(A) > 0$. Тогда по определению формулы A совпадает с одной из формул вида

$$A_1 \& A_2 \ , \ A_1 \vee A_2 \ , \ \delta x A_1 \ , \ A_1 \to A_2 \ , \ \overline{A_1}. \tag{4.8}$$

По предположению индукции формулы A_1, A_2 эквивалентны приведенным формулам. Заменив ими в (4.8) формулы A_1, A_2, мы в трех первых случаях сразу получим приведенные формулы. А так как $A_1 \to A_2 \equiv \overline{A_1} \vee A_2$ (см. эквивалентность 17), то остается рассмотреть случай, когда $A = \overline{A_1}$. Если A_1 — элементарна, то A — приведенная формула. Если же A_1 не элементарна, то она может иметь вид

$$B_1 \& B_2 \ , \ B_1 \vee B_2 \ , \ \delta x B_1 \ , \ B_1 \to B_2 \ , \ \overline{B_1},$$

где $r(B_i) \leq r(A) - 2$. Тогда по свойствам 11, 12, 22, 23, 19, 14 формула A эквивалентна одной из формул:

$$\overline{B_1} \vee \overline{B_2} \ , \ \overline{B_1} \& \overline{B_2} \ , \ \delta^* x \overline{B_1} \ , \ B_1 \& \overline{B_2} \ , \ B_1, \tag{4.9}$$

где δ^* — квантор, двойственный к δ.

Остается применить предположение индукции к формулам B_1, $\overline{B_1}$, $\overline{B_2}$ и заменить их в (4.9) приведенными формулами. □

Теорема 4.3. *Для всякой формулы A алгебры предикатов существует эквивалентная ей предваренная формула.*

□ По теореме 4.2, не теряя общности, можно считать, что A — приведенная формула. Снова применим индукцию по $r(A)$. Для $r(A) = 0$ утверждение верно. Пусть $r(A) > 0$. Тогда формула A может иметь вид

$$1)\ A_1 \& A_2\ ,\ 2)\ A_1 \vee A_2\ ,\ 3)\ \delta x A_1\ ,\ 4)\ \overline{A_1},$$

причем в случае 4 A_1 — элементарная формула и для нее утверждение теоремы верно. Рассмотрим случаи 1–3.

1. $A = A_1 \& A_2$. По предположению индукции A_i эквивалентна предваренной формуле B_i, $i = 1, 2$, причем согласно эквивалентности 27 связанные переменные любой из формул B_1, B_2 можно считать отличными от всех переменных другой формулы. Таким образом, $A \equiv B_1 \& B_2$, где можно считать

$$B_1 = \delta_1 x_1 ... \delta_k x_k C_1\ ,\ B_2 = \delta_{k+1} x_{k+1} ... \delta_n x_n C_2,$$

$x_1, ..., x_k$ не входят в C_2, $x_{k+1}, ..., x_n$ не входят в C_1 и формулы C_1, C_2 не содержат кванторов. Отсюда, используя эквивалентность 26, получим

$$A \equiv \delta_1 x_1 ... \delta_k x_k \delta_{k+1} x_1 ... \delta_n x_n (C_1 \& C_2).$$

2. $A = A_1 \vee A_2$, рассуждения двойственны.

3. $A = \delta x A_1$. По предположению индукции A_1 эквивалентна предваренной формуле

$$A_1 \equiv \delta_1 x_1 ... \delta_k x_k B, \tag{4.10}$$

причем можно считать, что $x \neq x_1, ..., x_k$. Возможны два подслучая:

а) B содержит свободные вхождения x. Тогда A эквивалентна предваренной формуле

$$\delta x \delta_1 x_1 ... \delta_k x_k B;$$

б) B не содержит свободных вхождений x. Тогда из эквивалентности (4.10) следует, что значения формулы A_1 не зависят от значений переменной x, а потому $A \equiv A_1$, и теорема доказана. □

Глава 5

ИСЧИСЛЕНИЕ ПРЕДИКАТОВ

5.1. ЯЗЫК, АКСИОМЫ И ПРАВИЛА ВЫВОДА ИСЧИСЛЕНИЯ ПРЕДИКАТОВ

В качестве алфавита исчисления предикатов возьмем то же самое множество $\mathcal{A}$, которое служило алфавитом при определении формул алгебры предикатов сигнатуры Σ. За его элементами сохраним те же самые обозначения, хотя здесь на все буквы мы должны пока смотреть просто как на символы, не имеющие какого-либо содержательного смысла. Например, символ операции f здесь не обозначает какую-либо конкретную операцию, определенную на каком-либо конкретном множестве. То же относится и к символам предикатов. Термины же «символ операции» и «символ предиката» объясняются тем, что в приложениях исчисления предикатов к конкретным математическим теориям они будут трактоваться (интерпретироваться) как операции и предикаты на конкретных множествах. Аналогично предметным переменным будут придаваться значения из этих множеств.

Таким образом, в формулах исчисления предикатов могут участвовать функциональные и предикатные символы, символы предметных переменных и логических операций, скобки и запятые.

Понятия терма, формулы сигнатуры Σ, свободных и связанных вхождений предметных переменных в формулу и области действия квантора определяются в исчислении предикатов

буквально так же, как в алгебре предикатов (см. определения 4.1, 4.2).

Равенство формул, как и в алгебре предикатов, будем обозначать знаком «=». При записи формул будут использоваться те же правила сокращения числа скобок, что и в алгебре предикатов.

Множество всех термов и всех формул исчисления предикатов сигнатуры Σ обозначим соответственно через $T(\Sigma)$ и $\Phi(\Sigma)$. Всюду далее в общих рассуждениях данной главы сигнатура будет считаться произвольной, и потому вместо $T(\Sigma)$ и $\Phi(\Sigma)$ будем писать соответственно T и Φ. Множество всех предметных переменных обозначим через X.

Из множества всех формул ниже особую роль будут играть формулы, не содержащие свободных вхождений предметных переменных. Они называются замкнутыми формулами, или предложениями.

В качестве аксиом исчисления предикатов принимаются все аксиомы исчисления высказываний, в которых теперь буквами A, B, C обозначаются любые формулы исчисления предикатов, и еще две аксиомы, содержащие операции навешивания кванторов:

V. 1) $\forall x A(x) \to A(t)$,

2) $A(t) \to \exists x A(x)$,

где $A(x)$ — формула, содержащая свободные вхождения предметной переменной x, а $A(t)$ — формула, полученная заменой в $A(x)$ всех свободных вхождений x термом t, удовлетворяющим условию: ни одно свободное вхождение x в $A(x)$ не находится в области действия квантора по какой-либо переменной, содержащейся в t (при этом условии говорят, что терм t свободен для x в формуле $A(x)$).

В качестве правил вывода формул исчисления предикатов выбирают:

I. Правило заключения

$$\frac{A,\ A \to B}{B},$$

где A, B — любые формулы.

II. Правило $\forall$-введения:

$$\frac{B \to A}{B \to \forall x A},$$

где A содержит свободные вхождения переменной x, а B не содержит свободных вхождений x.

III. Правило $\exists$-удаления:

$$\frac{A \to B}{\exists x A \to B},$$

где A, B — формулы, удовлетворяющие тем же условиям, что и в правиле II.

Заметим, что, подставляя в аксиомы вместо A, B, C произвольные формулы, мы получим бесконечное множество формул. Таким образом, записи аксиом являются, по существу, схемами, по которым можно получать все новые и новые формулы. То же можно сказать и о формулах из правил вывода.

Определив формулы, аксиомы и правила вывода, мы определили тем самым логическое исчисление, называемое исчислением предикатов 1-й ступени сигнатуры Σ. Меняя Σ, т. е. меняя множество формул, и сохраняя схемы аксиом и правил вывода, мы будем получать другие исчисления предикатов. В дальнейшем алфавит будем считать фиксированным и соответствующее логическое исчисление будем обозначать буквой $\mathcal{L}$.

Заметим, что, кроме исчисления предикатов 1-й ступени, в математической логике и ее приложениях рассматривается также исчисление предикатов 2-й ступени. В нем кванторы $\forall$, $\exists$ могут навешиваться не только на предметные переменные, но и на функциональные переменные и предикатные переменные.

5.2. ВЫВОДИМОСТЬ И ДОКАЗУЕМОСТЬ ФОРМУЛ

Определение 5.1. *Выводом формулы A из конечного или бесконечного множества формул T в исчислении $\mathcal{L}$ называется конечная последовательность формул*

$$A_1, A_2, ..., A_n, \tag{5.1}$$

в которой $A_n = A$ и каждая из формул A_i, $i \in \overline{1,n}$, является или аксиомой, или формулой из T, или получается по некоторому правилу вывода из предыдущих формул последовательности (5.1).

При этом формула A называется выводимой из системы T, что записывается в виде $T \vdash A$ или

$$B_1, ..., B_k \vdash A,$$

если $T = \{B_1, ..., B_k\}$.

Особо важным является случай, когда T — пустое множество формул.

Определение 5.2. *Формула A логического исчисления $\mathcal{L}$, выводимая из пустого множества формул, называется доказуемой формулой, или теоремой исчисления $\mathcal{L}$, что обозначается в виде $\vdash A$.*

О выводах и доказательствах формул сделаем ряд замечаний.

Замечание 5.1. При выводах и доказательствах формул в исчислении $\mathcal{L}$, как и в исчислении высказываний, можно использовать не только аксиомы и посылки, но и любые доказанные ранее формулы.

Замечание 5.2. Так как все схемы аксиом и правил вывода исчисления высказываний сохраняются в исчислении $\mathcal{L}$, то все выводы и доказательства формул в исчислении высказываний могут быть без изменений повторены для любых формул исчисления $\mathcal{L}$.

Замечание 5.3. В исчислении $\mathcal{L}$ точно так же, как и в исчислении высказываний, доказывается следующая

Теорема 5.1 (теорема об ограниченной дедукции). *Для произвольных множества формул T и формулы A исчисления $\mathcal{L}$: если $T, A \vdash B$ и существует вывод формулы B из $T \cup \{A\}$ без использования правил $\forall$-введения и $\exists$-удаления, то*

$$T \vdash A \to B.$$

Из замечаний 5.1–5.3 следует, что в исчислении $\mathcal{L}$ имеют место все вспомогательные правила вывода, доказанные в исчислении высказываний.

Замечание 5.4. В исчислении $\mathcal{L}$ точно так же, как и в исчислении высказываний, определяется и обозначается равносильность формул. Формулы A и B называются равносильными, если доказуемы формулы $A \to B$ и $B \to A$.

Приведем примеры доказуемых формул, специфических для исчисления $\mathcal{L}$.

Теорема 5.2. *Для любой формулы A, содержащей свободные вхождения переменной x доказуемы следующие формулы:*

$$\textit{а)} \vdash \overline{\exists x A(x)} \to \forall x \overline{A(x)}, \ \textit{б)} \vdash \forall x \overline{A(x)} \to \overline{\exists x A(x)},$$

$$\textit{в)} \vdash \overline{\forall x A(x)} \to \exists x \overline{A(x)}, \ \textit{г)} \vdash \exists x \overline{A(x)} \to \overline{\forall x A(x)}.$$

□ а) $A(x) \to \exists x A(x)$ — V.2;
$\overline{\exists x A(x)} \to \overline{A(x)})$ — пр. контрапозиции;
$\overline{\exists x A(x)} \to \forall x \overline{A(x)})$ — пр. $\forall$-введения.
б)$\forall x \overline{A(x)} \to \overline{A(x)}$ — V.1;
$\overline{\overline{A(x)} \to \overline{\forall x \overline{A(x)}}}$ — пр. контрапозиции;
$A(x) \to \overline{\overline{A(x)}}$ — IV.2;
$A(x) \to \overline{\forall x \overline{A(x)}}$ — пр. силлогизма;
$\exists x A(x) \to \overline{\forall x \overline{A(x)}}$ — пр. $\exists$-удаления;
$\overline{\overline{\forall x \overline{A(x)}}} \to \overline{\exists x A(x)}$ — пр. контрапозиции;
$\forall x \overline{A(x)} \to \overline{\overline{\forall x \overline{A(x)}}}$ — IV.2;
$\forall x \overline{A(x)} \to \exists x A(x)$ — пр. силлогизма.

Доказательства формул *а)* и *d)* в некотором смысле двойственны соответственно доказательствам формул *а)*, *,)*. Доказательство *г)* следует начать с аксиомы V.1 для формулы $A(x)$, а доказательство *в)* — с аксиомы V.2 для формулы $\overline{A(x)}$. Распишите эти доказательства в качестве упражнения. □

Следствие 1. *В исчислении $\mathcal{L}$ имеют место следующие равносильности:*

$$\overline{\forall x A(x)} \sim \exists x \overline{A(x)},$$

$$\overline{\exists x A(x)} \sim \forall x \overline{A(x)}.$$

При доказательствах и выводах формул в исчислении предикатов может использоваться также следующая теорема дедукции.

Теорема 5.3. *Пусть T — произвольное множество формул, B — любая формула и A — любая замкнутая формула исчисления $\mathcal{L}$. Тогда в исчислении $\mathcal{L}$*

$$T, A \vdash B$$

в том и только том случае, когда

$$T \vdash A \to B.$$

□ Эта теорема доказывается по той же схеме, что и теорема дедукции в исчислении высказываний (см. 3.1) — индукцией по длине кратчайшего вывода формулы B из T, A. Здесь необходимо лишь при переходе от $m \leq n$ к $m = n + 1$ дополнительно рассмотреть два новых случая:

а) B получена из формулы B_i по правилу $\forall$-введения;

б) B получена из B_i по правилу $\exists$-удаления. Тогда $B_i = A \to B_i$, и по предположению

В случае *а)*:

$$B_i = C \to D(x)\ ,\ B = C \to \forall x D(x),$$

причем x не имеет свободных вхождений в C по определению правила $\forall$-введения. По предположению индукции имеем

$$T \vdash A \to (C \to D(x)).$$

Дополним вывод формулы $A \to (C \to D(x))$ из T до вывода формулы $A \to B$ следующими формулами:

$C \& D \to D(x)$ — правило умножения посылок;

$A \& C \to \forall x D(x)$ — пр. $\forall$-введения (применение этого правила возможно, поскольку формула $A \& C$ не содержит свободных вхождений x);

$A \to (C \to \forall x D(x))$ — пр. разделения посылок.

В случае *б)* рассуждения аналогичны, но вместо $\forall$-введения используется правило $\exists$-удаления. $\square$

В качестве примера на применение теоремы дедукции докажем одно утверждение, которое понадобится нам в следующем параграфе.

Теорема 5.4. *Если T — любое множество формул, B — любая формула, A — замкнутая формула исчисления $\mathcal{L}$ и*

$$T, A \vdash B\&\overline{B},$$

то

$$T \vdash \overline{A}.$$

$\square$ Приведем вывод формулы $\overline{A}$ из T.
$A \to \overline{B}\&B$ — теорема дедукции;
$\overline{\overline{B}\&B} \to \overline{A}$ — пр. контрапозиции;
$\overline{\overline{B}} \vee \overline{B} \to \overline{A}$ — пр. де Моргана;
$B\&\overline{\overline{B}} \to B$ — II.1;
$\overline{B} \to \overline{B}$ — пример 3.1;
$B \to \overline{\overline{B}}$ — аксиома IV.1;
$B \vee \overline{B} \to \overline{\overline{B}} \vee \overline{B}$ — монотонность операции $\vee$;
$B \vee \overline{B} \to \overline{A}$ — пр. силлогизма;
$B \vee \overline{B}$ — закон исключенного третьего;
$\overline{A}$ — правило заключения. $\square$

5.3. СЕМАНТИКА ИСЧИСЛЕНИЯ ПРЕДИКАТОВ

Под семантикой любой формализованной теории понимается исследование, направленное на выяснение содержательного смысла основных ее понятий — формул, аксиом, правил вывода, доказуемых формул, доказательств, выводов и др. Одним из основных в этом направлении является вопрос о соотношении между множествами доказуемых формул и истинных утверждений неформализованной теории. В частности, здесь выясняется, все ли утверждения теории могут быть выражены формулами, все ли истинные утверждения могут быть доказаны, не будет ли рассматриваемая формализация теории

противоречивой, нельзя ли ее расширить, не приходя к противоречию, и т. п.

В главе 3 все эти вопросы были положительно решены для исчисления высказываний, если его рассматривать как формализацию алгебры высказываний. В данном параграфе некоторые из указанных вопросов будут решены для исчисления предикатов.

Теорема 5.5. *Если каждая формула некоторого множества T формул алгебры предикатов сигнатуры Σ истинна в некоторой алгебраической системе G сигнатуры Σ и $T \vdash A$, то A также истинна на системе G.*

□ Метод полной математической индукции по длине вывода формулы A сводит доказательство теоремы к проверке истинности на G аксиом исчисления предикатов и сохранения тождественной истинности формул правилами вывода I–III. Проверка истинности аксиом первых четырех групп и сохранения истинности формул правилом заключения проводится точно так же, как и при доказательстве теоремы 3.2.

Докажем, что истинна аксиома V.1

$$\forall x A(x) \to A(t),$$

где t — терм, свободный для x в $A(x)$.

Пусть $x_1, ..., x_m$ — суть все предметные переменные, отличные от x и участвующие в образовании терма t, а $y_1, ..., y_n$ — все предметные переменные, имеющие свободные вхождения в формулу $A(x)$, отличные от $x, x_1, ..., x_m$. Заметим, что x может входить, а может и не входить в t. В целях общности запишем терм t в виде $t(x, x_1, ..., x_m)$. Так как терм t свободен для x в $A(x)$, то все вхождения переменных в терм t останутся свободными при подстановке t вместо x в формулу $A(x)$. В общем случае некоторые из переменных $x_1, ..., x_m$ (а возможно, и все) также могут иметь свободные вхождения в формулу $A(x)$. В связи с этим формулы $A(x)$ и $A(t)$ запишем подробнее в виде

$$A(x) = A(x, x_1, ..., x_m, y_1, ..., y_n),$$

$$A(t) = A(t(x, x_1, ..., x_m), x_1, ..., x_m, y_1, ..., y_n).$$

Допустим, что формула V.1 не является истинной на G. Это означает, что в G найдутся значения переменных

$$x = a,\ x_1 = a_1, ..., x_m = a_m,\ y_1 = b_1, ...,\ y_n = b_n,$$

такие, что на G:

$$\forall x A(x, a_1, ..., a_m, b_1, ..., b_n) \equiv 1,$$

$$A(b, a_1, ..., a_m, b_1, ..., b_n) \equiv 0,$$

где b — значение терма t при $x = a$, $x_1 = a_1$,..., $x_m = a_m$. Полученные соотношения противоречат определению квантора $\forall$. Значит, наше допущение неверно, и аксиома V.1 — истинна на G. Аналогично доказывается этот факт и для аксиомы V.2.

Докажем теперь, что заключение $D = B \to \forall x A(x)$ правила $\forall$-введения будет истинным, если истинна его посылка $C = B \to A$. Пусть $x_1, ..., x_n$ — все предметные переменные, кроме x, имеющие свободные вхождения в формулу C. Тогда правило $\forall$-введения подробнее запишется в виде

$$\frac{B(x_1, ..., x_n) \to A(x, x_1, ..., x_n)}{B(x_1, ..., x_n) \to \forall x A(x, x_1, ..., x_n)}.$$

Пусть на G формула C истинна, а формула D не является истинной, т. е. найдутся значения переменных $x = a$, $x_1 = a_1$,..., $x_m = a_n$, такие что на G:

$$B(a_1, ..., a_n) \to A(a, a_1, ..., a_n) \equiv 1,$$

$$B(a_1, ..., a_n) \to \forall x A(x, a_1, ..., a_n) \equiv 0.$$

при любом $a \in G$. Получили снова противоречие с определением квантора $\forall$. Аналогично придем к противоречию с определением квантора $\exists$, допустив, что отношение истинности формул не сохраняется правилом $\exists$-введения. □

Следствие 1. *Любая формула, доказуемая в исчислении предикатов, является тождественно истинной в алгебре предикатов.*

Следствие 2. *Исчисление предикатов непротиворечиво, т. е. в нем не могут быть доказуемыми какая-либо формула и ее отрицание.*

Замечание 5.5. Теорема 5.5 и ее следствия теряют силу, если не накладывать ограничений в аксиомах V.1, V.2 на терм t, а в правилах вывода — на формулу B. Приведем примеры.

1. Пусть M есть множество натуральных чисел, $A(x) = \exists y\,(x < y)$ и $t = x + y$ — терм, не свободный для x в $A(x)$, поскольку свободное вхождение переменного x в $A(x)$ находится в области действия квантора $\exists$ по переменному y, входящему в терм t. Легко видеть, что формула

$$\forall x\, A(x) \to A(t),$$

или, подробнее,

$$\forall x(\exists y\,(x < y)) \to \exists y\,(x + y < y)$$

ложна в арифметике натуральных чисел.

2. В той же алгебраической системе при $B = (x < y)$, $A = (x < y+1)$ формула $B \to A$ истинна, а формула $B \to \forall x A$ ложна.

Естественно возникает вопрос о достаточности средств исчисления предикатов для доказательства всех тождественно истинных формул из алгебры предикатов. На этот вопрос положительно отвечает следующая теорема Гёделя о полноте исчисления предикатов.

Теорема 5.6. *Любая тождественно истинная формула алгебры предикатов доказуема в исчислении предикатов.*

Для доказательства этой теоремы введем ряд понятий и докажем одно утверждение, составляющее главную компоненту в доказательстве теоремы Гёделя о полноте.

Определение 5.3. *Множество формул T исчисления $\mathcal{L}$ называется противоречивым, если существует такая формула A, что*

$$T \vdash A\&\overline{A}.$$

В противном случае множество T называется непротиворечивым.

Определение 5.4. *Множество формул T сигнатуры Σ называется выполнимым, если найдется алгебраическая система сигнатуры Σ, в которой все формулы из T принимают истинное значение хотя бы при одной, общей для всех формул из T замене свободных вхождений предметных переменных.*

Определение 5.5. *Множество формул T сигнатуры Σ называется полным, если для каждой замкнутой формулы A сигнатуры Σ имеет место*

$$T \vdash A \text{ или } T \vdash \overline{A}.$$

Теорема 5.7. *Любое непротиворечивое множество замкнутых формул исчисления $\mathcal{L}$ выполнимо.*

□ Пусть T — любое непротиворечивое множество замкнутых формул сигнатуры Σ. Добавим к Σ счетное множество новых констант (т. е. символов нуль-арных операций)

$$\mathcal{B} = \{b_0, b_1, b_2, ...\}.$$

Получим новую сигнатуру Σ_1. Заменяя в аксиомах и правилах вывода исчисления $\mathcal{L}$ формулы в сигнатуре Σ формулами сигнатуры Σ_1, мы получим новое логическое исчисление $\mathcal{L}_1$. Из его определения видно, что все доказанные факты об исчислении $\mathcal{L}$ (в частности, все доказанные формулы и вспомогательные правила вывода) имеют место и в $\mathcal{L}_1$. Так как $\mathcal{L}_\infty$ включает в себя $\mathcal{L}$, то T можно рассматривать как множество замкнутых формул исчисления $\mathcal{L}_1$. Легко показать, что T будет непротиворечивым и в $\mathcal{L}_1$. Действительно, если в $\mathcal{L}_1$ из T выводима формула $A\&\overline{A}$, то, заменив в формулах вывода все константы из $\mathcal{B}$ символами предметных переменных, не участвующими в этом выводе, мы получим вывод в $\mathcal{L}$ из T формулы $B\&\overline{B}$, где B получена из A указанной выше заменой констант. Таким образом, множество формул T непротиворечиво в $\mathcal{L}_1$.

Теперь расширим определенным образом множество T. Для этого занумеруем все замкнутые формулы исчисления $\mathcal{L}_1$:

$$A_0, A_1, A_2, ...$$

По формуле A_0 и системе T построим множество формул T_0, добавив к системе T:

а) формулу $\overline{A_0}$, если $T \vdash \overline{A_0}$;

б) формулу A_0, если из T не выводима $\overline{A_0}$ и A_0 не имеет вида $\exists x\, B(x)$;

в) формулы A_0 и $B(b_{i_0})$, если из T не выводима $\overline{A_0}$ и $A_0 = \exists x\, B(x)$, где b_{i_0} — символ с наименьшим номером из $\mathcal{B}$, не встречающийся в A_0.

Из построения множества T_0 имеем

$$T_0 \vdash A \text{ или } T_0 \vdash \overline{A}.$$

Покажем, что T_0 непротиворечиво. Допустим, что в $\mathcal{L}_1$:

$$T_0 \vdash A\&\overline{A}. \tag{5.2}$$

В соответствии с определением T_0 рассмотрим три случая.

1. В этом случае $T \vdash \overline{A}_0$ и $T_0 = T \cup \{\overline{A_0}\}$. Из (5.2) следует, что

$$T, \overline{A_0} \vdash A\&\overline{A}.$$

Отсюда по теореме 5.4 с использованием правила двойного отрицания получаем $T \vdash A_0$, что невозможно в силу непротиворечивости T.

2. В этом случае $T_0 = T \cup A_0$, и потому $T_0 \vdash A_0$. Далее, из (5.2) мы, как и в предыдущем случае, получим $T_0 \vdash \overline{A_0}\}$, что противоречит условию.

3. Здесь $A_0 = \exists x B(x)$, $T_0 = T \cup \{A_0, B(b_{i_0})\}$. Как и выше, из (5.2) и теоремы дедукции получим

$$T \vdash A_0 \to \overline{B(b_{i_0})}.$$

Так как b_{i_0} не входит в A_0 и в формулы из T, то, повторив весь вывод формулы $A_0 \to \overline{B(b_{i_0})}$ из T с заменой b_{i_0} на переменную x_j, не участвующую в выводе, получим:

$$T \vdash A_0 \to \overline{B(x_j)}.$$

Дополним вывод формулы $A_0 \to \overline{B(x_j)}$ из T формулами:

$A_0 \to \forall x_j \overline{B(x_j)}$ — пр. $\forall$-введения;

$\forall x_j \overline{B(x_j)} \to \overline{B(x)}$ — V.1;
$A_0 \to \overline{B(x)}$ — пр. силлогизма;
$A_0 \to \forall x \overline{B(x)}$ — пр. $\forall$-введения;
$\forall x \overline{B(x)} \to \exists x B(x)$ — теорема 5.1;
$A_0 \to \overline{\exists x B(x)}$ — пр. силлогизма.

В итоге мы получили вывод формулы $A_0 \to \overline{A_0}$ из T. Следовательно, $T, A_0 \vdash \overline{A_0}$. Отсюда по правилу умножения формул получаем $T, A_0 \vdash A_0 \& \overline{A_0}$. Теперь, применяя теорему 5.4, получим $T \vdash \overline{A_0}$ — противоречие с условием.

Таким образом, доказана непротиворечивость множества T_0.

Далее по T_0 и A_1 аналогично построим непротиворечивое множество формул T_1, такое что

$$T_0 \subset T_1 \text{ и } T_1 \vdash A_1 \text{ или } T_1 \vdash \overline{A_1}.$$

Продолжая этот процесс, мы получим последовательность множеств формул исчисления $\mathcal{L}_1$:

$$T_0 \subset T_1 \subset T_2 \subset ...,$$

таких что $T_i \vdash A_i$ или $T_i \vdash \overline{A_i}$. Следовательно, объединение S всех множеств T_i есть непротиворечивое и полное множество формул исчисления $\mathcal{L}_\infty$.

Теперь построим алгебраическую систему сигнатуры Σ_1, в которой истинны все формулы из T. В качестве основного множества M системы возьмем множество всех термов сигнатуры Σ_1, не содержащих символов предметных переменных. Такие термы обычно называют замкнутыми. Их определение можно получить из определения 4.1, оставив в его пункте 1 лишь систему констант F_0. В рассматриваемом случае множество F_0 содержит, в частности, множество $\mathcal{B}$.

Определим на M операции $\widetilde{f}$ и предикаты $\widetilde{p}$, соответствующие символам операций f и предикатов p из Σ_1.

Если a — символ 0-арной операции из Σ_1, то положим $\widetilde{a} = a$. Если f — символ n-арной операции, p — символ n-арного предиката и $t_1, ..., t_n$ — термы из M, то положим:

$$\widetilde{f}(t_1, ..., t_n) = f(t_1, ..., t_n),$$

$$\widetilde{p}(t_1, ..., t_n) \equiv 1, \text{ если } S \vdash p(t_1, ..., t_n),$$
$$\widetilde{p}(t_1, ..., t_n) \equiv 0, \text{ если } S \vdash \overline{p(t_1, ..., t_n)}.$$

В итоге нами определена алгебраическая система (M, Σ_1). В дальнейшем для упрощения записей вместо $\widetilde{f}$ и $\widetilde{p}$ будем писать соответствующие f и p.

Индукцией по рангу формул докажем, что для любой замкнутой формулы A исчисления $\mathcal{L}_\infty$:

$$S \vdash A \text{ в } \mathcal{L}_1 \Longleftrightarrow A \equiv 1 \text{ в } (M, \Sigma_1). \tag{5.3}$$

Если $A = p(t_1, ..., t_n)$ — элементарная формула, то утверждение 5.2 верно по определению предиката p. Допустим, что оно верно для всех замкнутых формул ранга $r < k$, и пусть $r(A) = k$. В зависимости от последней операции в A возможны 6 случаев.

1. $A = A_1 \& A_2$. Используя определение конъюнкции, предположение индукции и правило умножения формул, получим

$$A \equiv 1 \Leftrightarrow A_1 \equiv 0 \ , \ A_2 \equiv 1 \Leftrightarrow S \vdash A_1 \ , \ S \vdash A_2 \Leftrightarrow S \vdash A_1 \& A_2.$$

2. $A = A_1 \vee A_2$. Если $A = A_1 \equiv 1$, то в силу замкнутости формулы A имеем: $A_1 \equiv 1$ или $A_2 \equiv 1$. Тогда по предположению индукции $S \vdash A_1$ или $S \vdash A_2$, а потому и $S \vdash A = A_1 \vee A_2$.

Обратно, пусть $S \vdash A = A_1 \vee A_2$. Если $S \vdash A_1$ или $S \vdash A_2$, то по предположению индукции $A_1 \equiv 1$ или $A_2 \equiv 1$, а потому и $A \equiv 1$. В противном случае в силу полноты системы S имеем

$$S \vdash \overline{A_1} \ , \ S \vdash \overline{A_2}.$$

Тогда по правилам умножения формул и де Моргана получим

$$S \vdash \overline{A_1} \& \overline{A_2} \ , \ S \vdash \overline{A_1 \vee A_2}.$$

В итоге имеем

$$S \vdash A_1 \vee A_2 \ , \ S \vdash \overline{A_1 \vee A_2},$$

что невозможно в силу непротиворечивости S.

3. $A = \overline{A_1}$. Учитывая предположение индукции и полноту системы формул S, получим

$$A \equiv \text{И} \iff A_1 \equiv \text{Л} \iff S \vdash \overline{A_1} \iff S \vdash A.$$

4. $A = A_1 \to A_2$. В этом случае, используя предположение индукции, правила введения посылки, контрапозиции и двойного отрицания, а также равносильности $A_1 \to A_2 \equiv \overline{a_2} \to \overline{A_1}$ и $A_1 \to A_2 \equiv \overline{a_1} \to A_2$, получим

$$A \equiv 1 \iff A_1 \equiv 0 \text{ или } A_2 \equiv 1 \iff S \vdash \overline{A_1} \text{ или } S \vdash A_2 \iff$$

$$\iff S \vdash \overline{A_2} \to \overline{A_1} \text{ или } S \vdash A_1 \vee A_2 \iff S \vdash A.$$

5. $A = \forall x A_1(x)$. Если $A \equiv 1$, то $A_1(t) \equiv 1$ при любом $t \in M$. Тогда по предположению индукции $S \vdash A_1(t)$ для всех $t \in M$. Допустим, что $S \vdash \overline{\forall x A_1(x)}$. Тогда по теореме 5.2 $S \vdash \exists x \overline{A_1(x)}$ и по построению множества S в S существует формула $\overline{A_1(b)}$ при некоторой константе $b \in M$. В итоге мы получили противоречие: $S \vdash A_1(t)$ при всех $t \in M$ и $S \vdash \overline{A_1(b)}$ при некотором $b \in M$. Следовательно, наше допущение неверно, и в силу полноты системы S имеем

$$S \vdash \forall x A_1(x), \text{ т. е. } S \vdash A.$$

Обратно, пусть $S \vdash A$. Если $A \equiv 0$, то $A_1(t) \equiv 0$ при некотором $t \in M$. Тогда по предположению индукции $S \vdash \overline{A_1(t)}$. Отсюда, используя аксиому V.2 и теорему 5.2, получим

$$S \vdash \exists x \overline{A_1(x)} \text{ и } S \vdash \overline{\forall x A_1(x)}.$$

Последнее невозможно в силу непротиворечивости S.

6. $A = \exists x A_1(x)$. Если $A \equiv 1$, то $A_1(t) \equiv 1$ при некотором $t \in M$, и по предположению индукции $S \vdash A_1(t)$. Отсюда, используя аксиому V.2, получим $S \vdash \exists x A_1(x)$, т. е. $S \vdash A$.

Обратно, если $S \vdash A$, то в силу непротиворечивости S имеем $S \nvdash \overline{A}$. Тогда по построению системы S в S содержится формула $A_1(b)$ при некотором $b \in M$. По предположению индукции $A_1(b) \equiv \text{И}$, и потому $\exists x A_1(x) \equiv \text{И}$.

Таким образом, все замкнутые формулы исчисления $\mathcal{L}_\infty$, выводимые из S, истинны на алгебраической системе (M, Σ_1).

А так как $T \subset S$, то и все формулы из T истинны на (M, Σ_1), т. е. множество T выполнимо.

Теперь можно доказать и теорему Гёделя о полноте исчисления $\mathcal{L}$.

□ Пусть A — формула исчисления предикатов $\mathcal{L}$ и $A \equiv 1$ в алгебре предикатов, т. е. в любой алгебраической системе сигнатуры Σ. Если $x_1, ..., x_n$ — суть все переменные, имеющие свободные вхождения в A, то формула $B = \forall x_1, ..., x_n A$ также истинна, а потому формула $\overline{B}$ ложна в алгебре предикатов. Отсюда следует, что множество формул $\{\overline{B}\}$ невыполнимо. Тогда по теореме 5.7 множество $\{\overline{B}\}$ противоречиво, т. е. существует формула C исчисления $\mathcal{L}$, такая что

$$\overline{B} \vdash C \& \overline{C}.$$

Отсюда по теореме 5.4 $\vdash \overline{\overline{B}}$, а потому и $\vdash B$. Теперь, применяя n раз аксиому V.1 и правило заключения, получим $\vdash A$. □

В качестве других следствий теоремы 5.7 докажем так называемую теорему о совместной выполнимости и теорему компактности, являющуюся частным случаем известной локальной теоремы Мальцева.

Теорема 5.8 (о совместной выполнимости). *Если замкнутая формула A сигнатуры Σ истинна на любой алгебраической системе, на которой истинны все формулы некоторого множества T замкнутых формул сигнатуры Σ, то $T \vdash A$.*

□ Из условия видно, что множество формул $S = T \cup \{\overline{A}\}$ невыполнимо. Значит, по теореме 5.7 оно противоречиво, т. е.

$$T, \overline{A} \vdash B \& \overline{B}$$

для некоторой формулы B. Отсюда по теореме 5.4 $T \vdash A$. □

Теорема 5.9 (компактности). *Множество замкнутых формул T исчисления $\mathcal{L}_1$ выполнимо тогда и только тогда, когда выполнимо любое его конечное подмножество.*

□ Выполнимость конечных подмножеств выполнимого множества формул очевидна. Докажем обратное утверждение. Допустим, что выполнимо любое конечное подмножество

формул множества T, а само S невыполнимо. Тогда по теореме 5.7 оно противоречиво, т. е. найдется такая формула A, что

$$T \vdash A\overline{A}.$$

По определению выводимости при выводе формулы $A\overline{A}$ может использоваться лишь конечное множество формул. Поэтому формула $A\overline{A}$ выводима из конечного подмножества S множества T. Следовательно, найдется противоречивое конечное множество формул из T. Если бы S было выполнимым, то по теореме 5.5 нашлась бы алгебраическая система, в которой была бы истинной формула $A\overline{A}$, что невозможно. □

Замечание 5.6. Если сигнатура Σ содержит непустое множество символов нуль-арных предикатов, то, применяя к ним лишь пункты 1 и 3 определения формулы предикатов, мы получим множество формул, называемых формулами алгебры высказываний. По определению предикатов нуль-арный предикат является просто высказыванием, значит алгебру высказываний и исчисление высказываний можно рассматривать как составные части соответственно алгебры предикатов и исчисления предикатов подходящей сигнатуры.

5.4. ПОНЯТИЕ О ТЕОРИИ МОДЕЛЕЙ

Теория моделей возникла на стыке между алгеброй и математической логикой. Она занимается в основном изучением таких свойств классов алгебраических систем, которые могут быть записаны формулами какого-либо логического исчисления (чаще всего исчисления предикатов). При этом сами алгебраические системы изучаемого класса становятся интерпретациями или моделями соответствующего исчисления. Из анализа связей между формальным языком (формальными выводами, доказательствами и т. п.) и его интерпретациями черпает теория моделей и главные средства исследований. В самостоятельную область математики теория моделей выделилась в 1950-х гг. Ее основоположниками являются А. Тарский и А. И. Мальцев.

К настоящему времени теория моделей является достаточно развитой областью математики со своей проблематикой и

собственными методами исследований (см. [54, 22]). Многие из разработанных в ней методов успешно используются в других областях математики.

С некоторыми такими методами мы познакомимся в данном параграфе.

Предварительно сделаем одно замечание о предикате равенства и дадим несколько определений.

Пусть K — некоторый класс алгебраических систем сигнатуры Σ. Тогда можно построить логическое исчисление в сигнатуре Σ и записывать свойства систем из класса K формулами полученного исчисления. Однако формализация свойств систем зачастую существенно облегчается, если в сигнатуру исчисления ввести еще двухместный предикат равенства $=$. Полученную сигнатуру будем обозначать в виде $\widehat{\Sigma}$. Сразу отметим, что, введя предикат равенства, мы должны пополнить и систему аксиом исчисления, введя в нее аксиомы, характеризующие предикат равенства. В качестве таких аксиом выбираются формулы

$$\forall x(x = x),$$

$$\forall x_1, x_2(x_1 = x_2 \to x_2 = x_1),$$

$$\forall x_1, x_2, x_3(x_1 = x_2 \& x_2 = x_3 \to x_1 = x_3),$$

характеризующие отношение равенства как отношение эквивалентности, а также все формулы вида

$$\forall x_1, ..., x_k, y_1, ..., y_k(x_1 = y_1 \& ... \& x_k = y_k) \to \\ \to f(x_1, ..., x_k) = f(y_1, ..., y_k)),$$

$$\forall x_1, ..., x_n, y_1, ..., y_n(x_1 = y_1 \& ... \& x_n = y_n) \to \\ \to p(x_1, ..., x_n) = p(y_1, ..., y_n)),$$

где f — функциональный, а p — предикатный символы из Σ соответствующих арностей. Последние формулы характеризуют согласованность отношения равенства со всеми предикатами и операциями из Σ.

Определение 5.6. *Пусть K — класс алгебраических систем сигнатуры Σ. Множество замкнутых формул исчисления*

предикатов сигнатуры $\widehat{\Sigma}$, истинных на всех системах из K, называется элементарной теорией класса K и обозначается $th(K)$ или $th(A)$, если K состоит из одной алгебраической системы A.

Например, можно говорить об элементарной теории арифметики натуральных чисел, об элементарной теории групп, об элементарной теории конечных полей и т. д.

Заметим, что слово «элементарная» в термине «элементарная теория» объясняется тем, что любая формула исчисления предикатов выражает свойства систем через указание связей между их элементами, поскольку кванторы навешиваются лишь на предметные переменные.

Задача описания элементарной теории $th(K)$ для разных классов K является важнейшей задачей теории моделей. В частности, теория $th(K)$ называется разрешимой, если существует алгоритм, позволяющий для любой формулы в сигнатуре Σ выяснять, содержится она в $th(K)$ или нет. В противном случае элементарная теория называется неразрешимой. В решении вопросов о разрешимости или неразрешимости элементарных теорий теория моделей смыкается с теорией алгоритмов, которая будет изучаться позже.

Определение 5.7. *Пусть T — произвольное множество формул исчисления предикатов сигнатуры $\widehat{\Sigma}$. Алгебраическая система F сигнатуры Σ, на которой истинны все формулы из T, называется моделью для T, что записывается в виде*

$$F \models T.$$

Вопросы, связанные с построением и исследованием моделей для заданных систем формул, также занимают важное место в теории моделей. Алгебраические системы сигнатуры Σ называют элементарно эквивалентными, если на них истинны одни и те же формулы исчисления предикатов сигнатуры $\widehat{\Sigma}$. Очевидно, что из изоморфизма систем следует их элементарная эквивалентность. Обратное же не всегда верно.

В связи с этим укажем на одну из долго стоящих и широко известных проблем А. Тарского (проблему о полноте

для элементарной теории свободных групп): будут ли элементарно эквивалентными свободные группы рангов m и n при $m > n > 1$?

Определение 5.8. *Класс K алгебраических систем сигнатуры Σ называется аксиоматизируемым (конечно аксиоматизируемым), если существует такое множество (конечное множество) формул T исчисления предикатов сигнатуры $\widehat{\Sigma}$, что*

$$F \in K \Longleftrightarrow F \vdash T$$

для любой алгебраической системы F сигнатуры Σ. При этом множество T называется системой аксиом класса K.

Очевидно, что наличие хорошей системы аксиом, описывающей класс алгебраических систем, играет немаловажную роль для изучения этого класса. В связи с этим в теории моделей вопросу аксиоматизируемости уделяется серьезное внимание. Приведем несколько примеров аксиоматизируемых классов.

Пример 5.1. Класс всех групп в сигнатуре $\Sigma =< \cdot, 1 >$ конечно аксиоматизируем. Системой аксиом может служить следующее множество формул T_{gr}.

$$1)\ \forall x_1, x_2, x_3((x_1 \cdot x_2) \cdot x_3 = x_1 \cdot (x_2 \cdot x_3)),$$

$$2)\ \forall x_1(x_1 \cdot 1 = x_1),$$

$$3)\ \forall x_1 \exists x_2(x_1 \cdot x_2 = 1).$$

Добавление к ним формулы

$$4)\ \forall x_1, x_2(x_1 \cdot x_2 = x_2 \cdot x_1)$$

приводит к аксиоматике класса всех абелевых групп.

Пример 5.2. Класс всех полей в сигнатуре $\Sigma =< +, \cdot, 0, 1 >$ конечно аксиоматизируем. Его система аксиом T_f может быть получена добавлением к аксиомам абелевых групп в сигнатуре $< +, 0 >$ формул 1, 2, 4 из примера 5.1 и еще формулы

$$\forall x_1(x_1 \neq 0 \to \exists x_2(x_1 \cdot x_2 = 1)).$$

Добавив еще формулу Φ_p:

$$\forall x((1 + ... + 1) \cdot x = 0),$$

где $1 + ... + 1 = p$ — простое число, мы выделим из всех полей класс полей характеристики p.

Класс полей характеристики 0 задается системой аксиом, полученной добавлением к системе T_f отрицаний формул Φ_p по всем простым числам P. Эта система аксиом бесконечна, известно, что класс полей характеристики 0 не является конечно аксиоматизируемым.

Пример 5.3. Арифметику натуральных чисел можно рассматривать как алгебраическую систему сигнатуры $\Sigma_A = =< +, \cdot, s, 1 >$, где s — символ унарной операции («следовать за»). Объединяя систему аксиом Пеано с требованиями к операциям сложения и умножения по Грассману (см. 1.5), получим следующую систему аксиом:

$$\forall x_1 \overline{s(x_1 = 1)},$$

$$\forall x_1, x_2(x_1 = x_2 \to s(x_1) = s(x_2)),$$

$$\forall x_1, x_2(s(x_1) = s(x_2) \to x_1 = x_2),$$

$$\forall x_1(x + 1 = s(x_1)),$$

$$\forall x_1, x_2(x_1 + s(x_2) = s(x_1 + x_2)),$$

$$\forall x_1(x_1 \cdot 1 = x_1),$$

$$A(1)\&\forall x_1(A(x_1) \to A(s(x_1)) \to \forall x_2 A(x_2)),$$

где A — произвольная формула сигнатуры Σ_A.

Пусть K — произвольный аксиоматизируемый класс алгебраических систем сигнатуры Σ и T_K — его система аксиом. Тогда, добавив к аксиомам исчисления предикатов аксиомы равенства и все аксиомы из T_K, мы построим новое логическое исчисление $\mathcal{L}_K$. В нем, как и в исчислении предикатов $\mathcal{L}$, можно определить понятия формулы, выводимой из заданной системы формул, доказуемой формулы и т. д. Алгебраические системы класса K будут являться такими интерпретациями

исчисления $\mathcal{L}_K$, в которых истинны все формулы из T_K, т. е. будут моделями системы T_K.

Для исчисления $\mathcal{L}_K$, как и для исчисления $\mathcal{L}$, доказывается теорема о полноте: для любой замкнутой формулы A исчисления $\mathcal{L}_K$:

$$A \in th(K) \Longleftrightarrow \vdash A \text{в} \mathcal{L}_K.$$

Эту теорему, с одной стороны, можно использовать для проверки истинности формулы A на системах класса K, а с другой стороны, из нее следует ряд общих методов получения утверждений об алгебраических системах. Один из таких методов основан на теореме компактности.

Первые примеры на применение этой теоремы были указаны в теории групп самим автором теоремы — А. И. Мальцевым [39]. Им на основании теоремы компактности было показано, что для любой группы G и любого группового свойства S, записываемого конечной системой замкнутых формул в подходящей сигнатуре и сохраняющегося при переходе к подгруппам, группа G обладает свойством S в том и только том случае, когда им обладают все конечно порожденные подгруппы группы G.

Из этого утверждения, записывая те или иные групповые свойства формулами, можно получать различные конкретные результаты теории групп. На этом пути А. И. Мальцевым были получены, в частности, следующие результаты.

Если каждая конечно порожденная подгруппа группы G разрешима ступени k, то и G — разрешимая группа ступени k.

Если каждая конечно порожденная подгруппа группы G содержит нормальный делитель индекса $k \leq n$, то G также содержит нормальный делитель индекса $k \leq n$.

Проиллюстрируем на примере еще один метод теории моделей для получения утверждений об алгебраических системах, называемый методом переноса.

Теорема 5.10. *Если B — замкнутая формула в сигнатуре теории полей и B истинна на всех полях характеристики 0, то существует такое простое число p_0, что B — истинна на любом поле характеристики $p > p_0$.*

□ Действительно, если B истинна на всех полях характеристики 0, то B выводима из множества формул T, являющегося объединением множества аксиом исчисления предикатов в сигнатуре $\widehat{\Sigma}$, где $\Sigma =< +, \cdot, 0, 1 >$, и системы аксиом $T_f \cup \{\overline{\Phi}_p : p = 2, 3, 5, ...\}$, определяющих класс полей характеристики 0. Однако в выводе формулы B используется лишь конечное множество аксиом вида $\overline{\Phi}_p$. Если такие аксиомы использовались лишь при $p \leq p_0$, то B будет истинной на всех полях характеристики $p > p_0$. □

Выбирая в качестве B различные формулы, мы будем получать различные конкретные утверждения. В частности, взяв в качестве B формулу, которой записывается некоторая конечная система алгебраических уравнений над кольцом целых чисел, можно получить следующий интересный факт.

Если система алгебраических уравнений над кольцом $\mathbb{Z}$ не имеет решений в любом расширении поля рациональных чисел, то существует такое простое число p_0, что при каждом простом $p > p_0$ система уравнений, полученная заменой в исходной системе всех коэффициентов их вычетами по модулю p, не имеет решений в любом конечном поле характеристики p.

5.5. ПРОБЛЕМА РАЗРЕШИМОСТИ В ЛОГИКЕ ПРЕДИКАТОВ

При оценках рассуждения того или иного человека нередко приходится слышать (или читать): «Его рассуждения логичны» или, наоборот, «Его рассуждения нелогичны».

Одна из основных задач математической логики, да и любой логики вообще, заключается в указании метода построения правильных умозаключений. С формальной точки зрения это означает, что каждое из произносимых утверждений должно быть выводимым из предыдущих утверждений и некоторых общепринятых фактов (аксиом) по правилам логики. В связи с этим важной задачей математической логики является так называемая проблема выводимости, заключающаяся в построении эффективного метода (алгоритма), позволяющего для любого заданного конечного множества замкнутых

формул T и для любой замкнутой формулы A выяснять, выводима A из T или нет. Уточнение понятия алгоритма будет дано ниже.

Из теоремы дедукции (см. 5.3) следует, что проблема выводимости в исчислении предикатов эквивалентна проблеме разрешимости, т. е. проблеме распознавания доказуемости или недоказуемости формулы.

Из теорем 5.5 и 5.6 следует, что для решения этой проблемы достаточно уметь решать вопрос: является любая заданная замкнутая формула A тождественно истинной или нет? Для ответа на этот вопрос, необходимо уметь проверять истинность формулы A на всех алгебраических системах соответствующей сигнатуры. Однако множество таких систем, даже с точностью до изоморфизма, может быть бесконечным, да и сами системы могут быть бесконечными. Поэтому этот путь решения проблемы разрешимости не представляется возможным. Надежда на решение этой проблемы средствами самого исчисления предикатов также оказалась несостоятельной. В 1936 г. американским математиком и логиком А. Черчем было доказано, что в общем случае в узком исчислении предикатов алгоритма, устанавливающего доказуемость или недоказуемость любой фиксированной формулы, не существует. В связи с этим, естественно, возникает вопрос о решении указанной проблемы для формул некоторых частных типов. Так, например, проблема легко решается для формул, составленных только из нуль-местных и одноместных предикатов. Приведем схему доказательства этого факта.

Сначала доказывается, что если такая формула A содержит ровно n одноместных предикатов $p_1, ..., p_n$ и истинна на какой-либо алгебраической системе $F = (M, p_1, ..., p_n)$, то она истинна на некоторой системе, содержащей не более 2^n элементов. В качестве основного множества такой системы можно взять фактор-множество M/R, где R — отношение эквивалентности на M, такое что для любых $a, b \in M$:

$$aRb \Longleftrightarrow p_i(a) = p_i(b), i = 1, ..., n.$$

Значения предикатов $p_1, ..., p_n$ на классах определяются путем сведения к представителям классов.

Так как число попарно неизоморфных алгебраических систем ограниченного порядка конечно, то в итоге поиск моделей для формулы A сводится к просмотру лишь конечного числа конечных алгебраических систем: если A истинна на всех своих интерпретациях порядков, не превосходящих 2^n, то она тождественно истинна. Действительно, в противном случае будет выполнима формула $\overline{A}$, которая по доказанному будет иметь модель порядка $m \leq 2^n$, что противоречит условию.

Остается доказать существование эффективной процедуры распознавания истинности формулы A на заданной конечной системе $F = (M, p_1, ..., p_n)$. Для этой цели используется следующий очевидный факт. Если $M = \{a_1, ..., a_m\}$, то формулы вида $\forall x B(x)$ и $\exists x B(x)$ истинны на F тогда и только тогда, когда истинны соответственно формулы

$$B(a_1)\&...\&B(a_m) \text{ и } B(a_1) \vee ... \vee B(a_m).$$

Таким образом нашу замкнутую формулу A можно преобразовать в формулу, состоящую лишь из постоянных высказываний. После этого остается найти ее значение, пользуясь лишь определениями операций алгебры высказываний.

В 1930 г. французским математиком и логиком Дж. Эрбраном был указан подход к автоматическому доказательству тождественной истинности некоторых формул алгебры предикатов. Он разработал эффективный метод (алгоритм) распознавания доказуемости замкнутой формулы, который давал правильный ответ лишь в том случае, когда исходная формула действительно доказуема или, что то же самое, тождественно истинна. В противном случае алгоритм работал бесконечно и не давал никакого ответа.

В последующий период подход Эрбрана улучшался с точки зрения повышения его эффективности и упрощения программирования многими авторами. Наиболее существенные улучшения внесли М. Дэвис и Г. Патнем в 1960 г. (см. [64]) и Дж. Робинсон в 1965 г. (на русском языке опубликовано в 1970 г., см. [55]). Свой метод Робинсон назвал методом резолюций.

Во всех модификациях алгоритма Эрбрана решение вопроса о тождественной истинности замкнутой формулы A осно-

вывалось на опровержении допущения о выполнимости отрицания формулы A. Иначе говоря, допускалось (так и мы будем далее считать), что формула A не является тождественно истинной, т. е. что ее отрицание выполнимо, и после этого находилось противоречие, если в действительности формула A была тождественно истинной. Это противоречие и доказывало ее тождественную истинность. При этом в каждом случае проводились следующие предварительные преобразования формулы $\overline{A}$.

1. Приведение формулы $\overline{A}$ к равносильной предваренной нормальной форме, что можно эффективно сделать в силу теоремы 4.3. В итоге получится формула вида

$$A_1 = Q_1 x_1 ... Q_k x_k A_2 \ , \ k \geq 0,$$

где Q_i — квантор $\forall$ или $\exists$, $i = 1, ..., k$, $Q_1 x_1 ... Q_k x_k$ — кванторная приставка формулы A_1, которая считается пустой при $k = 0$, A_2 — бескванторная часть, или матрица, формулы A_1. При $k > 0$ в нее входят предметные переменные $x_1, ..., x_k$ и не входят никакие другие предметные переменные. При $k = 0$ в ней вообще нет предметных переменных, и, значит, входящие в нее термы строятся с использованием лишь функциональных символов, среди которых обязательно должны быть и нуль-местные. Напомним, что такие термы называются замкнутыми, а нуль-местные функциональные символы — константами.

2. Приведение матрицы A_2 формулы A_1 к конъюнктивной нормальной форме A_3, которая определяется и находится так же, как и в алгебре высказываний, только здесь роль переменных высказываний играют элементарные подформулы. В итоге получится формула $A_4 = Q_1 x_1 ... Q_k x_k A_3$, в которой A_3 есть конъюнкция формул вида

$$C_1 \vee ... \vee C_r \ , \ r \geq 1, \tag{5.4}$$

а каждая из C_i — элементарная (атомарная) формула или ее отрицание. Каждая из таких формул C_i называется литералом. Два литерала, из которых один является отрицанием другого, назовем противоположными.

Заметим, что если в формуле (5.4) встечаются противоположные литералы, то она тождественно истинна, и ее, не теряя общности, можно из A_3 удалить. Если в формуле (5.4) встечаются два одинаковых литерала, то один из них можно удалить в силу закона идепотентности для дизъюнкции.

Учитывая отмеченные факты, мы можем далее, не теряя общности, считать, что среди формул $C_1, ..., C_r$ нет одинаковых или противоположных литералов. При этих условиях формулы вида (5.4) будут называться дизъюнктами. Введем еще обозначение Λ для пустого дизъюнкта, не содержащего ни одного литерала. В дальнейшем при операциях с дизъюнктами могут появляться дизъюнкции литералов, содержащие одинаковые или противоположные литеры. В каждом таком случае автоматически из равных литералов будет оставляться лишь один, а дизъюнкции с противоположными литералами будут удаляться.

3. Преобразование формулы A_4 к так называемой стандартной сколемовской форме, т. е. к предваренной формуле, не содержащей кванторов существования.

Это преобразование, применимое к любой предваренной нормальной форме B, осуществляется следующим образом.

Если кванторная приставка формулы B не содержит кванторов существования, то B остается без изменений, и процесс ее преобразования закончен.

Если первый квантор имеет вид $\exists$ и навешивается по переменной x_1, то он аннулируется, а в матрице формулы B каждое вхождение x_1 заменяется одним и тем же символом нуль-местной функции, не использованным в формуле B. Если же самый первый квантор $\exists$ навешивается по переменной x_{m+1} и перед ним находятся кванторы $\forall x_1, ..., \forall x_m$, то квантор $\exists x_{m+1}$ аннулируется, а в матрице формулы B каждое вхождение x_{m+1} заменяется одним и тем же термом $f(x_1, ..., x_m)$, где f — любой символ m-арной функции, не встречающийся в формуле B.

После аннулирования в формуле B первого встретившегося в ней квантора существования мы снова получим замкнутую формулу. Далее ту же процедуру применяем к полученной формуле, пока не аннулируем все кванторы существования.

Полученная в итоге формула B_1 и называется стандартной сколемовской формой для формулы B по имени предложившего ее норвежского математика Т. А. Сколема (или Скулема). Заметим, что B_1 будет в общем случае формулой в сигнатуре, отличной от сигнатуры формулы B. Сигнатура формулы B_1 будет расширением сигнатуры формулы B путем добавления новых (сколемовских) функциональных символов.

Нетрудно видеть, что формула B_1 выполнима тогда и только тогда, когда выполнима формула B. Докажите это в качестве упражнения индукцией по числу кванторов $\exists$ в формуле B.

В нашем случае $B = A_4$ и ее матрица является конъюнктивной нормальной формой. Поэтому и полученная из нее сколемовская стандартная форма будет, очевидно, также иметь матрицу в конъюнктивной нормальной форме. Множество всех дизъюнктов этой КНФ обозначим через S. При этом мы должны помнить, что каждая предметная переменная системы S, как и формулы B_1, связана кванторами $\forall$. Условимся еще сигнатуру формулы B_1 и системы S обозначать буквой Σ.

В итоге имеем: формула $\overline{A}$ невыполнима тогда и только тогда, когда невыполнима система дизъюнктов S.

Одна из замечательных идей Эрбрана заключалась в том, что для проверки выполнимости системы дизъюнктов S можно ограничиться ее проверкой лишь на алгебраических системах с некоторым фиксированным основным множеством $U(S)$, называемым теперь универсумом Эрбрана. Это множество состоит из всех замкнутых термов сигнатуры $F \subset \Sigma$, если в ней множество F_0 констант, т. е. символов нуль-местных операций, не пусто. В противном случае оно определяется так же, но для сигнатуры, полученной из Σ добавлением одной константы. Заметим, что если сигнатура Σ содержит хотя бы один функциональный символ арности, большей нуля, то множество $U(S)$ счетно и является алгеброй сигнатуры F, в противном случае оно конечно и совпадает с F_0 при $F_0 \neq \varnothing$ и состоит из единственного добавленного символа при $F_0 = \varnothing$.

Пусть V — непустое подмножество из $U(S)$. Обозначим через $V(S)$ множество всех дизъюнктов, которые можно получить путем подстановок элементов из V вместо предметных

переменных в дизъюнкты из S, при условии, что в любом фиксированном дизъюнкте из S все вхождения одного и того же переменного заменяются одним и тем же термом из V.

Эрбраном была доказана теорема, которую, используя введенные обозначения, можно сформулировать следующим образом (см. [55]).

Теорема Эрбрана. Система дизъюнктов S невыполнима тогда и только тогда, когда найдется конечное подмножество V из $U(S)$, при котором невыполнимо множество $V(S)$.

Заметим, что $V(S)$ конечное множество формул без предметных переменных и его можно рассматривать как множество формул алгебры высказываний. Поэтому его проверка на невыполнимость может быть осуществлена за конечное время, например с использованием теоремы 2.9. И все же воспользоваться теоремой Эрбрана для проверки на невыполнимость системы S затруднительно, поскольку в общем случае число конечных подмножеств множества $U(S)$ бесконечно и нужное множество V можно искать сколь угодно долго, даже если оно существует.

В связи с этим представляет особый интерес предложенный Робинсоном метод резолюций для опровержения выполнимости системы дизъюнктов S. Для его описания введем необходимые понятия.

Определение 5.9. *Пусть в дизъюнктах C, D из S можно произвести такие замены предметных переменных термами, что в полученных дизъюнктах C', D' появятся противоположные литеры соответственно C_1, D_1. Тогда, удалив их и взяв дизъюнкцию оставшихся от C', D' частей, мы получим дизъюнкт, называемый резольвентой дизъюнктов C, D.*

Нетрудно видеть, что дизъюнкты C, D либо совсем не имеют резольвент, либо имеют лишь конечное число резольвент.

Определение 5.10. *Резолюцией системы дизъюнктов S называется множество $R(S)$, состоящее из всех дизъюнктов из S и всех резолюций всех пар дизъюнктов из S.*

Далее индуктивно определяются множества $R^n(S)$ для любого натурального числа n. При $n > 1 : R^n(S) = R^{n-1}(S)$.

Дж. Робинсоном доказана

Теорема 5.11 (основная теорема о резолюции). *Произвольное множество дизъюнктов S невыполнимо, если найдется такое n, при котором $R^n(S)$ содержит пустой дизъюнкт Λ.*

Ограничимся доказательством этой теорема для формул алгебры высказываний.

□ Пусть $M = \{C_1, ..., C_k\}$ — множество дизъюнктов алгебры высказываний. Оно выполнимо тогда и только тогда, когда выполнима КНФ $A = C_1 \vee ... \vee C_k$. Будем заменять в формуле A и в получаемых из нее КНФ последовательно конъюнкции пар дизъюнктов, содержащих противоположные литеры, их резольвентами до тех пор, пока не получим КНФ B, которая либо имеет пустой дизъюнкт, либо не имеет ни одной пары дизъюнктов, содержащих противоположные литералы. В итоге получим конечную последовательность формул, эквивалентных формуле A. Если формула B содержит пустой дизъюнкт, то предыдущая формуле должна иметь противоположные сомножители типа x_i и $\overline{x_i}$. Тогда эта формула тождественно ложна и, значит, формула A невыполнима. Если же формула B не содержит пустого дизъюнкта, то входящие в нее переменные высказывания разбиваются на два непересекающихся класса так, что переменные одного класса входят в нее только без отрицаний, а переменные второго класса — только с отрицанием. Не исключается случай, когда один из этих классов пуст. В любом случае при значениях переменных из первого класса, равных 1, а переменных из второго класса, равных 0, формула B примет значение 1, и, значит, формула A выполнима.□

Ниже использованное в доказательстве преобразование формулы A в формулу B нам еще встретится. Условимся говорить, что B получена из A по правилу резолюций.

Доказательство основной теоремы о резолюции для формул алгебры предикатов в общем случае усложняется тем, что в нем используемые в литералах предикаты могут браться

от различных термов, зависящих от различных предметных переменных, и тогда для применения использованного выше преобразования формул необходимо предварительно производить определенные замены предметных переменных во всей формуле на подходящие термы. В связи с этим доказательство получается весьма громоздким и мы его не приводим. Его можно найти, например, в оригинальной статье Дж. Робинсона [55] и в книге [64].

Для иллюстрации применения метода резолюций мы приведем один пример из этой книги [64], изменив лишь обозначения используемых предикатов (см. пример 5.22 на с. 93).

Пусть известно, что: 1) таможенные чиновники обыскивают каждого, кто въезжает в страну, кроме высокопоставленных лиц; 2) некоторые люди, способствующие провозу наркотиков, въезжали в страну и были обысканы исключительно людьми, также способствующими провозу наркотиков; 3) никто из высокопоставленных лиц не способствовал провозу наркотиков.

Доказать, что некоторые из таможенников способствовали провозу наркотиков.

Введем обозначения для следующих предикатов:

$E(x)$ — «x въезжал в страну»;

$V(x)$ — «x – высокопоставленное лицо»;

$Q(x,y)$ — «y обыскивал x»;

$C(x)$ — «x – таможенник»;

$P(x)$ — «x способствовал провозу наркотиков».

Теперь условия задачи можно записать формулами:

$A_1 = \forall x(E(x)\&\overline{V(x)} \to \exists y(Q(x,y)\&C(y)))$,

$A_2 = \exists x(P(x)\&E(x)\&\forall y(Q(x,y) \to P(y)))$,

$A_3 = \forall x(P(x) \to \overline{V(x)})$.

Заключение задачи запишется формулой

$A = \exists x(P(x)\&C(x))$.

Для решения задачи нужно показать, что из формул A_1, A_2, A_3 не выводима формула $\overline{A}$, т. е. невыполнима формула

$B = A_1\&A_2\&A_3\&\overline{A}$.

Так как B есть конъюнкция формул, а в итоге нас интересует КНФ ее стандартной сколемовской формы, то преоб-

разования 1–3 можно производить для каждого сомножителя отдельно.

1. Находим для формул A_1, A_2, A_3, B предваренные нормальные формы, пользуясь схемой доказательства теоремы 4.3:

$A_1' = \forall x \exists y(\overline{(E(x)} \vee V(x) \vee Q(x,y) \& C(y))$,
$A_2' = \exists x \forall y(P(x) \& E(x) \& (\overline{Q(x,y} \vee P(y)))$,
$A_3' = \forall x(\overline{P(x)} \vee \overline{V(x)})$,
$\overline{A} \equiv \forall x(\overline{P(x)} \vee \overline{C(x)})$.

2. Заменим матрицы полученных формул на их КНФ. Заметим, что при этом изменения требуются только для формулы A_1'. В этом случае получим

$A_1'' = \forall x \exists y(\overline{E(x)} \vee V(x) \vee Q(x,y)) \& (\overline{E(x)} \vee V(x) \vee C(y))$.

3. Приведем полученные формулы к стандартной сколемовской форме. Для этого в формуле A_1'' необходимо убрать квантор $\exists y$ и заменить y на функцию $f(x)$. В формуле A_2' нужно убрать квантор $\exists x$ и заменить x на константу a.

Мы не будем выписывать полученные при этом формулы, а выпишем сразу соответствующие им дизъюнкты.

(1) $\overline{(E(x)} \vee V(x) \vee Q(x, f(x))$,
(2) $\overline{E(x)} \vee V(x) \vee C(f(x))$,
(3) $P(a)$,
(4) $\underline{E(a)}$,
(5) $\overline{Q(a,y)} \vee P(y)$,
(6) $\overline{P(x)} \vee \overline{V(x)}$,
(7) $\overline{P(x)} \vee \overline{C(x)}$.

Теперь будем находить резольвенты формул, условившись резольвенты формул с номерами (i),(j) обозначать через $R(i,j)$.

(8) $R(3,6) = \overline{V(a)}$,
(9) $R(2,4) = V(a) \vee C(f(a))$,
(10) $R(8,9) = C(f(a))$,
(11) $R(1,4) = V(a) \vee Q(a, f(a))$,
(12) $R(8,11) = Q(a, f(a))$,
(13) $R(5,12) = \underline{P(f(a))}$,
(14) $R(7,13) = \overline{C(f(a))}$,
(15) $R(10,14) = \Lambda$.

Отсюда на основании теоремы о резолюции заключаем: формула B невыполнима, значит формула $\overline{A}$ не выводима из формул A_1, A_2, A_3.

В заключение данного параграфа отметим, что описанный метод резолюций является в чистом виде также малоэффективным, поскольку порядки множеств резолюций $R^n(S)$, как правило, растут очень быстро с ростом n. Вместе с тем в ходе его применения может порождаться большое число ненужных дизъюнктов. В связи с этим было придумано много приемов, позволяющих существенно уменьшать число используемых в алгоритме дизъюнктов. Подробно с ними можно познакомиться в книге [64]. Здесь же для примера мы приведем один из таких приемов, называемый правилом поглощения.

Непустой дизъюнкт C поглощается дизъюнктом D, если в C можно так заменить предметные переменные на термы, что полученный дизъюнкт будет частью дизъюнкта D. Правило поглощения заключается в том, что из любого используемого в методе резолюций множества дизъюнктов можно вычеркивать дизъюнкт, поглощаемый некоторым из остальных дизъюнктов.

Часть II

Дискретные функции

В учебной и научной литературе встречаются различные определения понятия дискретной функции. В данном пособии дискретными мы будем называть такие функции, у которых области определения и значений — конечные множества. В приложениях область определения дискретной функции часто представляется в виде декартового произведения множеств относительно небольшой мощности. Если $f : A \to B$ — дискретная функция и $A = A_1 \times \ldots \times A_n$, то f обозначают в виде

$$f(x_1, \ldots, x_n)$$

и называют дискретной функцией от n переменных $x_1, \ldots, x_n$. При этом переменная x_i принимает всевозможные значения из множества A_i, $i \in \overline{1,n}$. В том случае, когда $A_1 = \ldots = A_n = B$ и B является двухэлементным множеством, функция f называется двоичной или булевой функцией по имени ирландского математика Дж. Буля. В середине XIX века им были введены операции умножения, сложения и отрицания высказываний, принимающих два значения — «истина» и «ложь» (И и Л), а также рассмотрены в связи с этим двоичные функции на множестве {И, Л}. В настоящее время вместо букв {И, Л} обычно используют символы {1, 0}, поэтому под булевой функцией от n переменных понимают отображение $\Omega^n \to \Omega$, где $\Omega = \{0, 1\}$. Учитывая «логическое» происхождение булевых функций их называют также функциями алгебры логики, или функциями алгебры высказываний.

Класс булевых функций является важнейшим для приложений классом дискретных функций и исследовался многими отечественными и зарубежными специалистами. Так, например, К. Шеннон ([66, 67])изучал булевы функции в связи с их применением к вопросам защиты информации.

Математиками наряду с двоичной логикой изучались также k-значные логики при $k > 2$ и даже бесконечнозначные логики. Функции k-значной логики также нашли многочисленные приложения в теории конечных автоматов и других областях дискретной математики. В связи с использованием дискретных функций в компьютерной технике особую роль играют функции 2^m-значной логики.

В отечественной литературе первый и достаточно полный и подробный к тому времени обзор работ по функциям k-значной логики при $k \geq 2$ был опубликован С. В. Яблонским [69]. Вслед за К. Шенноном многие специалисты уделяли большое внимание вопросам сложности реализации дискретных функций и построенных на их основе различных управляющих систем. Больших успехов в этом направлении добились О. Б. Лупанов и многие другие участники отечественной научной школы по математической кибернетике, созданной С. В. Яблонским и О. Б. Лупановым.

Много интересных проблем в области теории вероятностей и комбинаторного анализа возникает при изучении дискретных функций в связи с их приложениями в теории корректирующих кодов и криптографии. Некоторые из таких вопросов рассмотрены в монографиях [35, 57].

Глава 1

ДИСКРЕТНЫЕ ФУНКЦИИ И СПОСОБЫ ИХ ЗАДАНИЯ

1.1. СПОСОБЫ ЗАДАНИЯ БУЛЕВЫХ ФУНКЦИЙ

Условимся использовать следующие обозначения:

Ω_2 (или просто Ω) $= \{0, 1\}$, где $0, 1$ — символы, под которыми можно понимать, в зависимости от контекста, «ложь» и «истину», числа, элементы поля и т. п.;

$X = \{x_1, x_2, \ldots\}$ — счетное множество символов переменных со значениями из Ω;

латинские буквы a, b, c, d, возможно с индексами, — элементы из Ω;

греческие буквы $\alpha, \beta, \gamma, \delta$, возможно с индексами, — элементы из Ω^n при известном n, называемые наборами или векторами-строками длины n в алфавите Ω;

α^T (или $\alpha^{\downarrow}$) — вектор, транспонированный к α, то есть набор α, записанный в виде столбца;

$\|\alpha\|$ — вес набора α, т. е. число символов 1, встречающихся в наборе α;

$\widetilde{\alpha}$ для набора $\alpha = (a_1, \ldots, a_n) \in \Omega^n$ — целое число $\sum a_i 2^{n-i}$, называемое величиной набора α.

Определение 1.1. *Булевой функцией (БФ) от n переменных, или n-местной булевой функцией, $n \in \mathbb{N}$ называется любое отображение $f : \Omega^n \to \Omega$.*

0-местными булевыми функциями называются константы из Ω.

Для n-местной БФ f обычно используют обозначение $f(x_1, \ldots, x_n)$.

Для БФ $f(x_1, \ldots, x_n)$ стандартным образом определяется понятие значения $f(a_1, \ldots, a_n)$ на наборе $(a_1, \ldots, a_n) \in \Omega^n$. Равенство БФ f и g понимается как равенство отображений и обозначается в виде $f = g$. При использовании подробных записей всегда будет предполагаться, что равные БФ являются функциями от одних и тех же переменных.

Множество всех булевых функций обозначим F_2, а множество всех булевых функций от n переменных — $F_2(n)$.

Рассмотрим вопрос о способах задания булевых функций.

1. Табличное задание

Из определения БФ $f(x_1, \ldots, x_n)$ видно, что для ее задания нужно указать значения $f(a_1, \ldots, a_n)$ на всех наборах $\alpha = (a_1, \ldots, a_n)$ из Ω^n. При этом сами наборы α можно не указывать, если считать, что они расположены в некотором фиксированном порядке, например по возрастанию величины $\widetilde{\alpha}$. Тогда БФ f будет определяться вектором-столбцом f^T, i-я координата которого есть значение $f(\alpha)$ при $\widetilde{\alpha} = i$. Таким образом, каждая БФ от n переменных будет задана набором длины 2^n элементов из Ω. Такое задание БФ называется *табличным*.

Проверьте самостоятельно справедливость следующего утверждения.

Утверждение 1.1. *Для любого* $n \in \mathbb{N}$:

$$|F_2(n)| = 2^{2^n}.$$

Пример. Множество $F_2(2)$ содержит 16 функций, которые мы представим столбцами их значений, занесенными в таблицу 2.1:

Таблица 2.1
Булевы функции от двух переменных

x_1x_2	f_0	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}
0 0	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
0 1	0	1	0	1	1	0	0	1	1	0	1	0	0	1	1	0
1 0	0	1	1	0	0	1	0	1	1	0	1	0	1	0	0	1
1 1	0	1	1	0	1	0	1	0	1	0	0	1	0	1	0	1

Определение 1.2. *Весом функции f называют величину*

$$\|f\| = |\{\alpha \in \Omega^n | f(\alpha) = 1\}|.$$

Множество наборов $\alpha \in \Omega^n$, для которых $f(\alpha) = 1$, будем обозначать N_f.

Очевидно, что для n-местной функции f выполняется неравенство $0 \le \|f\| \le 2^n$. Если для f выполнено $\|f\| = 2^{n-1}$, то такую функцию будем называть *устойчивой* (*равновероятной*).

Поскольку функции от двух переменных задают бинарные операции на множестве Ω, то их удобно записывать в виде так называемых «связок» (так же, как это имело место в алгебре высказываний). Приведем такую форму записи для этих функций с указанием их названий (некоторые, знакомые нам по первой части пособия, имеют по нескольку обозначений и названий):

$f_6(x_1, x_2) = x_1 \& x_2 = x_1 \wedge x_2 = x_1 \cdot x_2 = x_1 AND x_2$ —
конъюнкция, произведение по модулю 2, логическое «И»;

$f_7(x_1, x_2) = x_1 | x_2$ — штрих Шеффера;

$f_8(x_1, x_2) = x_1 \vee x_2 = x_1 OR x_2$ —
дизъюнкция, логическое «ИЛИ»;

$f_9(x_1, x_2) = x_1 \uparrow x_2$ — стрелка Пирса;

$f_{10}(x_1, x_2) = x_1 \oplus x_2 = x_1 XOR x_2$ —
сложение по модулю 2, исключающее «ИЛИ»;

$f_{11}(x_1, x_2) = x_1 \leftrightarrow x_2$ — эквиваленция;

$f_{13}(x_1, x_2) = x_1 \to x_2$ — импликация.

Следует отметить, что задание функции $f \in F_2(n)$ вектором f^T, являясь наиболее естественным, трудноосуществимо при больших значениях n. Одно из возможных упрощений задания связано с тем, что БФ $f \in F_2(n)$ однозначно задается любым из множеств:

$$N_f = \{\alpha \in \Omega^n : f(\alpha) = 1\} \text{ и } \Omega^n \setminus N_f.$$

Такое задание является удобным в тех случаях, когда одно из указанных множеств имеет малую мощность или достаточно простое описание, например:

— функция, называемая «счётчиком четности» принимает значение 1 на тех и только тех наборах α, для которых $\|\alpha\|$ — чётное число;

— функция называемая «счётчиком голосования» от n переменных принимает значение 1 на тех и только тех наборах α, для которых $\|\alpha\| > \frac{n}{2}$.

На практике более распространенным является способ задания БФ формулами алгебры логики и многочленами над полем $GF(2) = \{0, 1\}$ определенного вида, называемыми многочленами Жегалкина в честь русского математика, профессора Московского университета И. И. Жегалкина (1869–1947).

2. Задание БФ формулами алгебры логики

Если в алгебре высказываний вместо значений высказываний «И» и «Л» использовать соответственно символы 1 и 0, то каждая формула $A(x_1, \ldots, x_n)$ будет задавать (представлять) БФ $f(x_1, \ldots, x_n)$, определяемую условиями:

$$\forall \alpha \in \Omega_2^n : f(\alpha) = A(\alpha).$$

В связи с этим саму формулу $A(x_1, \ldots, x_n)$ естественно считать булевой функцией от переменных $x_1, \ldots, x_n$. Так, например, $x_1 \vee x_2$, $x_1 \& x_2$, $x_1 \to x_2$ являются одновременно и формулами, и функциями. Учитывая такое соглашение, ДНФ (КНФ) формулы называют также ДНФ (КНФ) соответствующей функции. В частности, сокращенную и минимальные ДНФ (КНФ) формулы $A(x_1, \ldots, x_n)$ называют сокращенной и минимальными ДНФ (КНФ) функции $f(x_1, \ldots, x_n)$, задаваемой формулой $A(x_1, \ldots, x_n)$.

Возникает естественный вопрос: любую ли БФ можно задать формулой алгебры высказываний? На этот вопрос положительно отвечает теорема 2.7 части I. Действительно, каждая БФ $f(x_1, \ldots, x_n)$ определяется множеством N_f, а из равенства (2.7) видно, что при $f \not\equiv 0$ $N_f = N_A$ для формулы

$$A(x_1, \ldots, x_n) = \bigvee_{(c_1, \ldots, c_n) \in N_f} x_1^{c_1} \ldots x_n^{c_n},$$

которая и задает функцию $f(x_1, \ldots, x_n)$. В случае $f \equiv 0$ функцию f можно задать, например, формулой $x_1 \& \overline{x}_2$.

Таким образом, любая БФ задается формулой алгебры логики. Более того, из теорем 2.7, 2.8 части I следует, что любая БФ $f(x_1, \ldots, x_n)$, не равная тождественно нулю (единице), задается СДНФ (СКНФ) от переменных $x_1, \ldots, x_n$, причем такое задание однозначно с точностью до перестановки конъюнктов (дизъюнктов).

Непосредственно из указанного способа сопоставления формулам функций следует также, что формулы $A(x_1, \ldots, x_n)$, $B(x_1, \ldots, x_n)$ представляют одну и ту же БФ от переменных $x_1, \ldots, x_n$ тогда и только тогда, когда они эквивалентны. В частности, любая БФ $f(x_1, \ldots, x_n) \not\equiv 0$ представляется однозначно с точностью до перестановки конъюнктов сокращенной ДНФ от переменных $x_1, \ldots, x_n$.

В итоге установлено взаимно однозначное соответствие между всеми БФ от переменных $x_1, \ldots, x_n$ и классами эквивалентности формул алгебры логики от тех же переменных.

Представление БФ совершенной ДНФ является частным видом более общего задания БФ через так называемые ее подфункции.

Рассмотрим ограничение $f|_M$ БФ $f(x_1, \ldots, x_n)$ на множество

$$M = \{(a_1, \ldots, a_n) \in \Omega^n : a_{i_1} = b_1, \ldots, a_{i_k} = b_k\}$$

при фиксированных $i_1, \ldots, i_k \in \overline{1, n}$, $i_1 < \ldots < i_k$ и $b_1, \ldots, b_k \in \in \Omega$, $0 < k < n$. Напомним, что $f|_M$ действует на M так же, как и f, и не определена на $\Omega^n \setminus M$. Так как M не является декартовой степенью Ω, то $f|_M$ не является булевой функцией. Однако по ней легко определить БФ от переменных $x_{j_1}, \ldots, x_{j_{n-k}}$, где $\{j_1, \ldots, j_{n-k}\} = \overline{1, n} \setminus \{i_1, \ldots, i_k\}$ и $j_1 < \ldots < j_{n-k}$, положив

$$\varphi(x_{j_1}, \ldots, x_{j_{n-k}}) = \\ = f(x_1, \ldots, x_{i_1-1}, b_1, x_{i_1+1}, \ldots, x_{i_k-1}, b_k, x_{i_k+1}, \ldots x_n). \quad (1.1)$$

Определение 1.3. *Функция* $\varphi(x_{j_1},\dots,x_{j_{n-k}})$, *определенная по БФ* $f(x_1,\dots,x_n)$ *равенством* (1.1), *называется функцией, полученной из* f *фиксацией переменных* $x_{i_1},\dots,x_{i_k}$ *соответственно константами* $b_1,\dots,b_k$, *и обозначается в виде*

$$\varphi(x_{j_1},\dots,x_{j_{n-k}}) = f^{b_1\dots b_k}_{i_1\dots i_k}(x_1,\dots,x_n).$$

Все булевы функции, полученные из f *фиксацией любых переменных константами, будем называть общим термином — подфункции функции* f.

Имеет место

Теорема 1.2. *Любая БФ* $f(x_1,\dots,x_n)$ *для произвольного* $0 < k \le n$ *представима в виде*

$$f(x_1,\dots,x_n) \equiv \bigvee_{(b_1,\dots,b_k)\in\Omega^k} x_{i_1}^{b_1}\cdots x_{i_k}^{b_k} f^{b_1\dots b_k}_{i_1\dots i_k}(x_1,\dots,x_n) \quad (1.2)$$

при любых фиксированных k, $i_1,\dots,i_k \in \overline{1,n}$, $i_1 < \dots < i_k$.

□ Доказательство этой теоремы проводится по той же схеме, что и доказательство теоремы 2.7 части I с учетом очевидных утверждений:

$$a_{i_1}^{b_1}\dots a_{i_k}^{b_k} = 1 \Leftrightarrow a_{i_1} = b_1,\dots,a_{i_k} = b_k, \text{ и}$$

$$f(a_1,\dots,a_n) = f^{a_{i_1}\dots a_{i_k}}_{i_1\dots i_k}(a_1,\dots,a_n). \quad \square$$

Из теоремы 1.2 при $k = n$ получаем представление функции $f(x_1,\dots,x_n) \not\equiv 0$ в виде СДНФ от переменных $x_1,\dots,x_n$.

При $k = 1$ получается еще одно практически важное представление БФ:

$$f(x_1,\dots,x_n) \equiv \overline{x}_i \cdot f(x_1,\dots,x_{i-1},0,x_{i+1},\dots,x_n) \vee \\ \vee\, x_i \cdot f(x_1,\dots,x_{i-1},1,x_{i+1},\dots,x_n). \quad (1.3)$$

Из очевидных равенств

$$a \vee b \equiv a \oplus b \oplus ab \text{ и } a\overline{a} \equiv 0$$

следует, что равенство (1.3) сохранится при замене в нем $\vee$ на $\oplus$.

Завершая разговор о задании БФ формулами алгебры высказываний, сделаем следующее замечание. При определении формул в главе 2 части I подчеркивалось, что одну и ту же формулу можно рассматривать от разных наборов переменных, к которым предъявляется только одно требование — каждый из них обязан содержать все переменные, встречающиеся в записи исходной формулы. Такая неоднозначность выбора переменных сохраняется и для булевых функций при их представлении формулами. В связи с этим переменные $x_1, \ldots, x_n$ БФ $f(x_1, \ldots, x_n)$ делятся на два класса — существенные и несущественные переменные.

Определение 1.4. *Переменная x_i булевой функции $f(x_1, \ldots, x_n)$ называется несущественной (фиктивной), если для любого набора элементов $a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n$ из Ω справедливо равенство:*

$$f(a_1, \ldots, a_{i-1}, 0, a_{i+1}, \ldots, a_n) = f(a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_n).$$

В противном случае говорят, что x_i — существенная переменная.

Из определения 1.4 видно, что значение функции не зависит от значений фиктивных переменных.

Пример. Дизъюнкцию $x_1 \vee x_2$ можно рассматривать как функцию от переменных $x_1, \ldots, x_n$ при любом $n \geq 2$. Тогда для нее переменные x_1 и x_2 — существенные, а остальные — фиктивные.

В этом примере существенные переменные совпали с переменными из записи формулы. Однако это совсем не обязательно.

Пример. Пусть $f(x_1, x_2, x_3) = x_1 \vee x_1 x_2 \overline{x}_3$. Очевидно, что значение функции f равно 1, если $x_1 = 1$, и равно 0, если $x_1 = 0$, независимо от того, какие значения принимают переменные x_2, x_3. Это означает, что для функции f переменная x_1 — существенная, а x_2, x_3 — несущественные (фиктивные).

Этот факт можно было заметить и не вычисляя значения функции, а воспользовавшись известным законом поглощения, по которому

$$x_1 \vee x_1 x_2 \overline{x}_3 \equiv x_1.$$

Пусть БФ $f(x_1, \ldots, x_n)$ представляется формулой $A(x_1, \ldots, x_n)$ и x_i — несущественная переменная функции f. Тогда из равенства (1.3) следует, что формула $A(x_1, \ldots, x_n)$ эквивалентна формуле $A(x_1, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_n)$, не содержащей переменной x_i. Отсюда легко следует

Утверждение 1.3. *Любая булева функция представляется формулой алгебры высказываний, содержащей лишь существенные переменные функции f.*

3. Представление БФ многочленами Жегалкина

Здесь под множеством $\{0, 1\}$ будем понимать поле с операциями сложения $\oplus$ и умножения $\cdot$ по модулю 2.

В алгебре многочленом Жегалкина от переменных $x_1, \ldots, x_n$ принято называть любой приведенный многочлен из кольца многочленов от переменных $x_1, \ldots, x_n$ над полем $(\Omega, \oplus, \cdot)$. Не обращаясь к алгебре, многочлены Жегалкина можно определить следующим образом.

Определение 1.5. *Многочленом Жегалкина от переменных $x_1, \ldots, x_n$ называют формулы 0, 1, а также все формулы вида*

$$\begin{array}{c} \sum\limits_{\substack{\{i_1, \ldots, i_k\} \in I \\ i_1 < \ldots < i_k}} \oplus \; x_{i_1} \ldots x_{i_k} \quad u \\ \sum\limits_{\substack{\{i_1, \ldots, i_k\} \in I \\ i_1 < \ldots < i_k}} \oplus \; x_{i_1} \ldots x_{i_k} \; \oplus 1, \; 0 < k \leq n \end{array}, \tag{1.4}$$

где $\sum \oplus$ — сумма по модулю 2, а суммирование ведется по любому непустому множеству I подмножеств из $\overline{1, n}$.

При этом дизъюнкты $x_{i_1} \ldots x_{i_k}$, а также слагаемое 1, если оно есть, формул (1.4) называются членами многочлена Жегалкина. Ясно, что каждый член, в свою очередь, является

многочленом, их иногда называют одночленами или мономами. Число k называется степенью члена $x_{i_1} \dots x_{i_k}$, а максимальная из степеней членов — степенью многочлена. Степени многочленов 1, 0 считаются равными соответственно 0 и $-\infty$.

Как и любая формула алгебры логики, многочлен Жегалкина $a(x_1, \dots, x_n)$ определяет булеву функцию от переменных $x_1, \dots, x_n$, которую обычно обозначают также через $a(x_1, \dots, x_n)$.

Теорема 1.4. *Любая БФ $f(x_1, \dots, x_n)$ представляется многочленом Жегалкина от $x_1, \dots, x_n$, и такое представление единственно с точностью до перестановки его членов.*

□ Так как каждая функция представляется формулой алгебры высказываний, то достаточно доказать, что любая формула $A(x_1, \dots, x_n)$ алгебры высказываний эквивалентна некоторому многочлену Жегалкина. При этом, не теряя общности, можно считать, что $A(x_1, \dots, x_n)$ — приведенная формула (см. параграф 2.4 части I).

Используя эквивалентности $a \vee b \equiv a \oplus b \oplus ab$ и $\overline{a} \equiv a \oplus 1$, мы перейдем от A к эквивалентной ей формуле B, в которой в качестве операций используются лишь $0, 1, \oplus, \cdot$. Далее, пользуясь свойствами операций поля $(\Omega, \oplus, \cdot)$, приведем формулу к эквивалентному ей многочлену Жегалкина от $x_1, \dots, x_n$. Возможность представления показана.

Покажем единственность. Заметим вначале, что в силу коммутативности операции $\oplus$ все многочлены Жегалкина разбиваются на классы попарно эквивалентных многочленов, отличающихся лишь перестановкой членов. Каждый класс, таким образом, определяется набором одночленов, составляющих входящие в данный класс многочлены. Поскольку различных одночленов (включая 1) от переменных $x_1, \dots, x_n$ существует 2^n, то число классов многочленов равно 2^{2^n}. Поскольку многочлены одного класса представляют только одну функцию, то существует взаимно однозначное соответствие между множеством булевых функций от n переменных и множеством классов многочленов. Таким образом, любая функция представляется многочленом Жегалкина однозначно с точностью до перестановки его одночленов. □

Представление БФ многочленами Жегалкина дает возможность классификации БФ по степеням нелинейности. Простейшими являются функции, имеющие степень нелинейности не больше 1. Они носят название *аффинных* функций. Каждая такая функция (отличная от 0 и 1) представляется многочленом вида

$$\text{а) } x_{i_1} \oplus \ldots \oplus x_{i_k} \text{ или б) } x_{i_1} \oplus \ldots \oplus x_{i_k} \oplus 1.$$

При этом функции вида *а)* и функция 0 называются *линейными*.

В дальнейшем нам также понадобится понятие коэффициентов многочлена Жегалкина.

Определение 1.6. *Коэффициентами многочлена Жегалкина* $a(x_1, \ldots, x_n)$ *называются элементы* a_0, $a_{i_1 \ldots i_k} \in \Omega$, *где* $k, i_1, \ldots, i_k \in \overline{1,n}$, $i_1 < \ldots < i_k$, *определяемые следующим образом:* $a_{i_1 \ldots i_k} = 1$, *если одночлен* $x_{i_1} \ldots x_{i_k}$ *входит в* $a(x_1, \ldots, x_n)$. *В противном случае* $a_{i_1 \ldots i_k} = 0$. *Коэффициент* a_0 *равен 1 тогда и только тогда, когда в* $a(x_1, \ldots, x_n)$ *входит член 1.*

4. Задание БФ многочленами над полем $\mathbb{R}$

Здесь под 0, 1 будем понимать действительные числа нуль и единицу.

Определение 1.7. *Будем говорить, что многочлен* $a(x_1, \ldots, x_n) \in \mathbb{R}[x_1, \ldots, x_n]$ *задает булеву функцию* $f(x_1, \ldots, x_n)$, *если для любого набора* $\alpha \in \Omega^n$ *значение* $a(\alpha)$ *содержится в* Ω *и* $a(\alpha) = f(\alpha)$. *Обозначение:*

$$f(x_1, \ldots, x_n) \simeq a(x_1, \ldots, x_n). \tag{1.5}$$

Иначе говоря, соотношение (1.5) означает, что БФ f есть ограничение многочлена a (как функции) на множество Ω^n:

$$f(x_1, \ldots, x_n) = a|_{\Omega^n}.$$

Пример.

$$x_1 \& x_2 \simeq x_1 x_2;\ x_1 \vee x_2 \simeq x_1 + x_2 - x_1 x_2;$$

$$x_1 \oplus x_2 \simeq x_1 + x_2 - 2x_1x_2;\ \overline{x} \simeq 1 - x.$$

Заметим, что представление БФ многочленом над $\mathbb{R}$ не однозначно:

$$\overline{x} \simeq 1 - x;\ \overline{x} \simeq 1 - 2x + x^2;\ \overline{x} \simeq 1 - x^2.$$

Однако при введении дополнительных условий на вид многочлена можно добиться однозначности представления функции действительным многочленом.

Теорема 1.5. *Любая булева функция $f(x_1, \ldots, x_n)$ однозначно представляется многочленом из кольца $\mathbb{R}[x_1, \ldots, x_n]$, имеющим следующий вид:*

$$\begin{gathered} a_0 + \sum_{\{i_1,\ldots,i_k\}} a_{i_1\ldots i_k} \cdot x_{i_1} \ldots x_{i_k}, \\ 1 \le i_1 < \ldots i_k \le n,\ a_0, a_{i_1\ldots i_k} \in \mathbb{R}, \end{gathered} \tag{1.6}$$

где суммирование ведется по всем непустым подмножествам $\{i_1, \ldots, i_k\}$ множества $\overline{1, n}$.

□ Доказательство проведем методом полной математической индукции по n.

В случае $n = 0$ доказательство очевидно.

Пусть утверждение теоремы верно при любом $n \le m$. Докажем, что оно справедливо при $n = m + 1$. Справедливо равенство

$$f(x_1, \ldots, x_n) \equiv \overline{x}_1 \cdot f(0, x_2, \ldots, x_n) + x_1 \cdot f(1, x_2, \ldots, x_n),$$

где $+$ — знак сложения в $\mathbb{R}$. Функции $f(0, x_2, \ldots, x_n)$ и $f(1, x_2, \ldots, x_n)$ зависят от $n - 1$ переменной, поэтому для них справедливо предположение индукции и существуют реализующие их многочлены $P_0(x_2, \ldots, x_n)$ и $P_1(x_2, \ldots, x_n)$ над $\mathbb{R}$. Тогда

$$f(x_1, \ldots, x_n) \simeq (1 - x_1) \cdot P_0(x_2, \ldots, x_n) + x_1 \cdot P_1(x_2, \ldots, x_n)$$

и выражение в правой части после раскрытия скобок и приведения подобных членов будет являться многочленом с действительными коэффициентами. Таким образом, доказана возможность представления.

Докажем однозначность. Пусть есть два различных многочлена $P(x_1,\ldots,x_n)$ и $Q(x_1,\ldots,x_n)$ вида (1.6), представляющие функцию $f(x_1,\ldots,x_n)$. Тогда, очевидно, многочлены $P(0,x_2,\ldots,x_n)$ и $Q(0,x_2,\ldots,x_n)$ представляют функцию $f(0,x_2,\ldots,x_n)$, а многочлены $P(1,x_2,\ldots,x_n)$ и $Q(1,x_2,\ldots,x_n)$ — функцию $f(1,x_2,\ldots,x_n)$. По предположению индукции $P(0,x_2,\ldots,x_n)=Q(0,x_2,\ldots,x_n)$ и $P(1,x_2,\ldots,x_n)=Q(1,x_2,\ldots,x_n)$.

Так как $P(x_1,\ldots,x_n)$ и $Q(x_1,\ldots,x_n)$ имеют вид (1.6), то они могут быть представлены в виде

$$P(x_1,\ldots,x_n)=x_1P'(x_2,\ldots,x_n)+P''(x_2,\ldots,x_n),$$

$$Q(x_1,\ldots,x_n)=x_1Q'(x_2,\ldots,x_n)+Q''(x_2,\ldots,x_n).$$

При этом, очевидно,

$$P''(x_2,\ldots,x_n)=P(0,x_2,\ldots,x_n),$$
$$Q''(x_2,\ldots,x_n)=Q(0,x_2,\ldots,x_n),$$

следовательно,

$$P''(x_2,\ldots,x_n)=Q''(x_2,\ldots,x_n).$$

Кроме того,

$$P(1,x_1,\ldots,x_n)=P'(x_2,\ldots,x_n)+P''(x_2,\ldots,x_n)$$

и

$$Q(1,x_1,\ldots,x_n)=Q'(x_2,\ldots,x_n)+Q''(x_2,\ldots,x_n),$$

следовательно,

$$P'(x_2,\ldots,x_n)=Q'(x_2,\ldots,x_n).$$

Таким образом, многочлены $P(x_1,\ldots,x_n)$ и $Q(x_1,\ldots,x_n)$ совпадают. □

В дальнейшем многочлен вида (1.6), представляющий булеву функцию f, будем называть *действительным многочленом БФ* f.

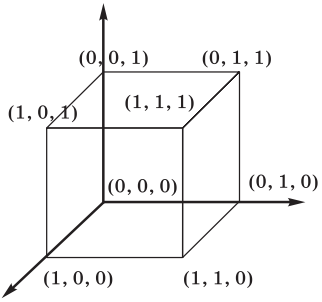

Рис. 1.1
3-мерный куб

5. Геометрическое представление булевых функций

Определение 1.8. *n-мерным кубом называют множество точек пространства $\mathbb{R}^n$ с координатами $(a_1, \ldots, a_n)$, где $a_i \in \{0, 1\}$, $i \in \overline{1, n}$.*

При малых n n-мерный куб представляют графически, изображая его двумерную проекцию и соединяя вершины, соответствующие соседним наборам.

Для задания булевой функции $f(x_1, \ldots, x_n)$ на n-мерном кубе отмечают вершины, соответствующие множеству N_f, выделяя при этом и ребра, соединяющие соседние вершины из этого множества.

В дальнейшем мы будем использовать также понятие грани n-мерного куба:

Определение 1.9. *Гранью ранга k (или, что то же самое, размерности $n - k$) n-мерного куба называется множество его вершин, соответствующее множеству N_φ, где φ — произвольная элементарная конъюнкция ранга k, т. е. $\varphi = x_{i_1}^{a_1} \ldots x_{i_k}^{a_k}$.*

Заметим, что гранью ранга k является множество вершин, соответствующих наборам, у которых k элементов принимают фиксированные значения, а остальные — произвольные. Таким образом, грань ранга k (размерности $n - k$) содержит 2^{n-k} вершин.

Приведем утверждение, описывающее свойства множества N_f, легко получаемое с использованием геометрического представления:

Утверждение 1.6. *Пусть f и g — булевы функции от одних и тех же переменных. Справедливы следующие утверждения:*

а) $f = \varphi \Leftrightarrow N_f = N_\varphi$;

б) $N_{f \cdot \varphi} = N_f \cap N_\varphi$;

в) $N_{f \vee \varphi} = N_f \cup N_\varphi$;

г) $f \vee \varphi \equiv f \Leftrightarrow N_\varphi \subseteq N_f$;

д) $f \equiv \bigvee\limits_{i=1}^{m} \psi_i \Leftrightarrow \bigcup\limits_{i=1}^{m} N_{\psi_i} = N_f$.□

В заключение проиллюстрируем применение геометрического представления к задаче минимизации ДНФ, задающей некоторую булеву функцию.

Из утверждения 1.6 следует, что если функция f представима ДНФ $\psi_1 \vee \ldots \vee \psi_r$, где $\psi_1, \ldots, \psi_r$ — элементарные конъюнкции, то N_f есть объединение граней $N_{\psi_1} \cup \ldots \cup N_{\psi_r}$ соответствующих размерностей. Таким образом, нахождение минимальной ДНФ функции f сводится к нахождению представления множества N_f в виде объединения таких граней, сумма рангов которых минимальна (т. е. сумма размерностей максимальна).

Заметим, что простым импликантам в геометрическом представлении соответствуют такие грани, которые не входят собственной частью ни в какую другую грань большей размерности, лежащую в N_f (будем называть их *максимальными* гранями). В противном случае импликанта будет содержать собственную часть, также являющуюся импликантой функции f, и, следовательно, не будет простой. Таким образом, для того чтобы по геометрическому представлению найти сокращенную ДНФ, надо найти все максимальные грани, входящие в N_f. Для нахождения же минимальной ДНФ в наборе максимальных граней надо выделить подмножество максимальных граней, дающих в объединении множество N_f

и имеющих минимальную сумму рангов (максимальную сумму размерностей).

Такой метод является весьма наглядным, однако он применим лишь для функций от небольшого числа переменных.

1.2. ПОЛНЫЕ СИСТЕМЫ И ЗАМКНУТЫЕ КЛАССЫ БУЛЕВЫХ ФУНКЦИЙ

В этом параграфе мы рассмотрим широко использующийся на практике метод построения булевых функций, суть которого можно кратко описать следующим образом. Пусть имеется множество исходных функций. Подставляя значения одной из них в качестве аргумента другой, мы будем получать новые функции, вообще говоря, от другого количества переменных, нежели исходные функции. Такой способ построения функций от большого числа переменных часто использовался в то время, когда элементная база вычислительной техники находилась на начальном этапе развития, однако и в настоящее время он продолжает оставаться актуальным.

Для того чтобы корректно определить множество функций, получаемых таким методом из функций исходного множества, нам потребуется понятие *формулы над классом функций*, являющееся обобщением уже введенного ранее понятия формулы алгебры высказываний (см. определение 2.2 части I).

Пусть задано непустое множество функций $\mathcal{K} = \{f_i, i \in \mathfrak{I}\}$ (в дальнейшем будем называть его классом или системой БФ) и множество символов переменных $X = \{x_1, x_2, \ldots\}$.

Определение 1.10. *1. Любой символ переменной из X есть формула над классом $\mathcal{K}$.*

2. Если $\mathcal{K}$ содержит 0-местные функции, то любой символ 0-местной функции есть формула над $\mathcal{K}$.

3. Если f_j — символ m-местной функции из $\mathcal{K}$, $m \in \mathbb{N}$, а $A_1, \ldots, A_m$ — формулы над $\mathcal{K}$, то выражение

$$f_j(A_1, \ldots, A_m) \tag{1.7}$$

есть формула над $\mathcal{K}$.

4. Других формул нет.

Множество всех формул над классом $\mathcal{K}$ обозначим $\Phi(\mathcal{K})$.

Понятия ранга и длины формулы, а также подформулы определяются по аналогии с соответствующими понятиями для формулы алгебры высказываний (см. 2.2, 2.3 части I).

Аналогично тому, как определяется соответствие между булевыми функциями и формулами алгебры высказываний, определяется соответствие между булевыми функциями и формулами над произвольным классом БФ. Таким образом, любой класс $\mathcal{K}$ определяет новый класс, состоящий из тех БФ, которые представляются формулами из $\Phi(\mathcal{K})$. Введем определение:

Определение 1.11. *Замыканием класса $\mathcal{K}$ булевых функций называют множество всех булевых функций, представимых формулами над $\mathcal{K}$. Обозначение:* $[\mathcal{K}]$.

Читателю предоставляется самостоятельно проверить следующие свойства замыкания класса:

Утверждение 1.7. *Справедливы утверждения:*

1) $\mathcal{K} \subseteq [\mathcal{K}]$;

2) $\mathcal{K}_1 \subseteq \mathcal{K}_2 \Rightarrow [\mathcal{K}_1] \subseteq [\mathcal{K}_2]$;

3) $[[\mathcal{K}]] = [\mathcal{K}]$.

Важнейшее место среди всех классов БФ занимают так называемые *полные системы БФ*:

Определение 1.12. *Система $\mathcal{K}$ БФ называется полной, если* $[\mathcal{K}] = F_2$.

Ниже мы приведем примеры полных системы БФ, а также докажем критерий полноты системы БФ.

Теорема 1.8. *Следующие системы булевых функций полны:*

а) $\mathcal{K}_0 = \{x_1x_2 = x_1\&x_2,\ x_1 \vee x_2,\ \overline{x}\}$;

б) $\mathcal{K}_1 = \{x_1x_2,\ \overline{x}_1\}$;

в) $\mathcal{K}_2 = \{x_1 \vee x_2,\ \overline{x}_1\}$;

г) $\mathcal{K}_3 = \{x_1x_2, x_1 \oplus x_2, 1\}$;

д) $\mathcal{K}_4 = \{x_1|x_2\}$;

е) $\mathcal{K}_5 = \{x_1 \uparrow x_2\}$.

□ Полнота систем $\mathcal{K}_0$, $\mathcal{K}_1$, $\mathcal{K}_2$ следует из того, что любая БФ представляется формулой алгебры высказываний, системы $\mathcal{K}_3$ — из теоремы 1.4. Докажем пункты *д)* и *е)* (соответствующие функции определены на с. 163).

Принимая во внимание утверждение 1.7 и полноту систем $\mathcal{K}_1$ и $\mathcal{K}_2$, достаточно доказать включения $\mathcal{K}_1 \subset [\mathcal{K}_4]$ и $\mathcal{K}_2 \subset [\mathcal{K}_5]$. Но это очевидно, поскольку

$$\overline{x}_1 \equiv x_1|x_1,\ x_1 \cdot x_2 \equiv (x_1|x_2)|(x_1|x_2)$$

и

$$\overline{x}_1 \equiv x_1 \uparrow x_1,\ x_1 \vee x_2 \equiv (x_1 \uparrow x_2) \uparrow (x_1 \uparrow x_2).\square$$

Определение 1.13. *Класс $\mathcal{K}$ булевых функций называется замкнутым, если $\mathcal{K} = [\mathcal{K}]$.*

Полное описание всех замкнутых классов булевых функций было дано Э. Постом (подробно это описание изложено в [70]). Он показал, что их число счетно и в каждом замкнутом классе $\mathcal{K}$ содержится конечный подкласс $\mathcal{K}'$, такой что $\mathcal{K} = [\mathcal{K}']$.

Опишем важнейшие замкнутые классы булевых функций.

1. Класс T_0 всех функций, *сохраняющих константу 0*:

$$T_0 = \{f(x_1, \ldots, x_n)|f(0, \ldots, 0) = 0\}.$$

2. Класс T_1 всех функций, *сохраняющих константу 1*:

$$T_1 = \{f(x_1, \ldots, x_n)|f(1, \ldots, 1) = 1\}.$$

3. Класс L_0 всех *линейных* функций:

$$L_0 = \{f(x_1, \ldots, x_n)|f(x_1, \ldots, x_n) = a_1x_1 \oplus \ldots a_nx_n,\ a_i \in \Omega\}.$$

4. Класс L всех *аффинных* функций:

$$L = \{f(x_1, \ldots, x_n)|f(x_1, \ldots, x_n) = a_1x_1 \oplus \ldots a_nx_n \oplus a_0,\ a_i \in \Omega\}.$$

5. Класс S всех *самодвойственных* функций:

$$S = \{f(x_1, \ldots, x_n)|f(x_1, \ldots, x_n) \equiv \overline{f}(\overline{x}_1, \ldots, \overline{x}_n)\}.$$

6. Класс M всех *монотонных* функций.

Для того чтобы определить класс M, введем на множестве Ω^n отношение частичного порядка «$\preccurlyeq$» следующим образом: для наборов $\alpha = (a_1, \ldots, a_n)$ и $\beta = (b_1, \ldots, b_n)$

$$\alpha \preccurlyeq \beta \Leftrightarrow \forall i \in \overline{1,n} \;\; a_i \leq b_i.$$

Здесь мы подразумеваем, что в Ω выполняется неравенство $0 \leq 1$.

Определение 1.14. *Булева функция $f(x_1, \ldots, x_n)$ называется монотонной, если из того, что $\alpha \preccurlyeq \beta$, следует, что $f(\alpha) \leq f(\beta)$.*

Определение 1.15. *Наборы $\alpha = (a_1, \ldots, a_n)$ и $\beta = (b_1, \ldots, b_n)$ из Ω^n называют соседними, если они отличаются ровно в одной позиции, т. е.*

$$\exists i \in \overline{1,n} : \; a_i \neq b_i \text{ и } \forall j \neq i \; : a_j = b_j.$$

В дальнейшем нам понадобится следующая лемма.

Лемма 1.9. *Булева функция $f(x_1, \ldots, x_n)$ не является монотонной тогда и только тогда, когда существуют соседние наборы α и β, такие что $\alpha \preccurlyeq \beta$ и $f(\alpha) > f(\beta)$.*

□ Если такие наборы α и β существуют, то функция f не монотонна по определению. Докажем обратное утверждение. Пусть f — не монотонна. Следовательно, найдутся наборы $\alpha = (a_1, \ldots, a_n)$ и $\beta = (b_1, \ldots, b_n)$, такие что $\alpha \preccurlyeq \beta$ и $f(\alpha) > f(\beta)$. Если они — соседние, то все доказано. Предположим, что они не являются соседними. Пусть $i_1, \ldots, i_s$ — все такие числа из множества $\overline{1,n}$, $s > 1$, что $a_{i_k} \neq b_{i_k}$, $k \in \overline{1,s}$. Поскольку $\alpha \preccurlyeq \beta$, то $a_{i_k} = 0$, $b_{i_k} = 1$. Рассмотрим последовательность наборов

$$\alpha_0 = \alpha, \alpha_1, \alpha_2, \ldots, \alpha_s = \beta,$$

где α_{k+1} получается из α_k заменой i_{k+1}-й координаты с 0 на 1. Очевидно, что α_k и α_{k+1} — соседние при $k \in \overline{0, s-1}$ и

$$\alpha = \alpha_0 \preccurlyeq \alpha_1 \preccurlyeq \alpha_2 \preccurlyeq \ldots \preccurlyeq \alpha_s = \beta.$$

Рассмотрим теперь последовательность $f(\alpha_0), \ldots, f(\alpha_s)$. Так как

$$f(\alpha_0) = f(\alpha) > f(\beta) = f(\alpha_s),$$

то найдется такое $m \in \overline{0, s-1}$, что $f(\alpha_k) > f(\alpha_{k+1})$. Тогда α_m и α_{m+1} — искомые соседние наборы. □

Теорема 1.10. *Классы T_0, T_1, L_0, L, S, M — замкнуты.*

□ Поскольку схемы доказательств одинаковы для всех классов, ограничимся доказательством только одного случая, для класса S. Покажем, что любая функция, представимая формулой A над классом S, сама принадлежит S.

Доказательство проводится индукцией по рангу формулы A.

1. Если $r(A) = 0$, то $A = x_i$, и так как $x_i \equiv \overline{\overline{x}_i}$, то x_i — самодвойственная функция.
2. Пусть любая функция, представленная формулой ранга не выше m над классом S, является самодвойственной.
3. Пусть $f(x_1, \ldots, x_n) \equiv A$, $A \in \Phi(S)$, $r(A) = m + 1$. В этом случае формула A имеет вид

$$A = g(A_1, \ldots, A_t),$$

где $g \in S$, а $A_1, \ldots, A_t$ — формулы ранга, не большего m над S. По предположению индукции, каждая из формул A_k представляет некоторую самодвойственную функцию φ_k. Таким образом, для любого набора $(a_1, \ldots, a_n) \in \Omega^n$ имеем

$$\begin{aligned}\overline{f}(\overline{a}_1, \ldots, \overline{a}_n) &= \overline{g}(\varphi_1(\overline{a}_1, \ldots, \overline{a}_n), \ldots, \varphi_t(\overline{a}_1, \ldots, \overline{a}_n)) = \\ &= \overline{g}(\overline{\overline{\varphi}}_1(\overline{a}_1, \ldots, \overline{a}_n), \ldots, \overline{\overline{\varphi}}_t(\overline{a}_1, \ldots, \overline{a}_n)) = \\ &= \overline{g}(\overline{\varphi}_1(a_1, \ldots, a_n), \ldots, \overline{\varphi}_t(a_1, \ldots, a_n)) = \\ &= g(\varphi_1(a_1, \ldots, a_n), \ldots, \varphi_t(a_1, \ldots, a_n)) = f(a_1, \ldots, a_n).\end{aligned}$$

Следовательно, f — самодвойственная функция. □

Докажем теперь критерий полноты системы булевых функций.

Теорема 1.11 (Э. Пост). *Система булевых функций K полна тогда и только тогда, когда она содержит хотя бы по*

одной функции каждого из классов $F_2 \setminus T_0$, $F_2 \setminus T_1$, $F_2 \setminus L$, $F_2 \setminus S$, $F_2 \setminus M$.

□ Пусть $\mathcal{G}$ — произвольный замкнутый класс, не совпадающий с F_2 и $\mathcal{K} \cap (F_2 \setminus \mathcal{G}) = \emptyset$. Тогда $\mathcal{K} \subseteq \mathcal{G}$, и поэтому

$$[\mathcal{K}] \subseteq [\mathcal{G}] = \mathcal{G} \neq F_2.$$

Следовательно, система $\mathcal{K}$ не полна. Этим теорема доказана в одну сторону, поскольку T_0, T_1, L, S, M — замкнутые классы, отличные от F_2.

Обратно, пусть $\mathcal{K}$ не лежит полностью ни в одном из классов T_0, T_1, L, S, M. Покажем, что тогда замыкание системы $\mathcal{K}$ содержит функции $x_1 \cdot x_2$ и $\overline{x}$. Тогда полнота системы $\mathcal{K}$ будет следовать из теоремы 1.8(б). Пусть f_1, f_2, f_3, f_4, f_5 — функции (не обязательно различные) из системы $\mathcal{K}$, причем $f_1 \notin T_0$, $f_2 \notin T_1$, $f_3 \notin L$, $f_4 \notin S$, $f_5 \notin M$.

Рассмотрим два случая в зависимости от значения $f_1(1, \ldots, 1)$. Покажем, что в обоих случаях система $[\mathcal{K}]$ содержит функции 0, 1, $\overline{x}$.

1. $f_1(1, \ldots, 1) = 1$. Так как $f_1(0, \ldots, 0) = 1$, то в этом случае формула $f(x, \ldots, x)$ представляет константу 1, т. е. $1 \in [\mathcal{K}]$. Тогда и константа $0 \equiv f_2(1, \ldots, 1)$ также лежит в $[\mathcal{K}]$.

Поскольку функция f_5 — не монотонна, то по лемме 1.9 найдутся такие соседние наборы α и β, что $\alpha \preccurlyeq \beta$ и $f_5(\alpha) = 1$, $f_5(\beta) = 0$. Пусть f_5 — функция от n переменных и

$$\alpha = (a_1, \ldots, a_{k-1}, 0, a_{k+1}, \ldots, a_n),$$

$$\beta = (a_1, \ldots, a_{k-1}, 1, a_{k+1}, \ldots, a_n).$$

Тогда, очевидно, формула $f_5(a_1, \ldots, a_{k-1}, x, a_{k+1}, \ldots, a_n)$ представляет функцию $\overline{x}$. А так как $x, 0, 1$ и f_5 содержатся в $[\mathcal{K}]$, то и $\overline{x} \in [\mathcal{K}]$.

2. $f_1(1, \ldots, 1) = 0$. Так как $f_1(0, \ldots, 0) = 1$, то в этом случае формула $f_1(x, \ldots, x)$ представляет функцию $\overline{x}$. Таким образом, $\overline{x} \in [\mathcal{K}]$.

Так как функция f_4 не является самодвойственной, то найдется набор $\alpha = (a_1, \ldots, a_n)$:

$$f_4(a_1, \ldots, a_n) = f_4(\overline{a}_1, \ldots, \overline{a}_n) = C,$$

где $C \in \Omega$. Рассмотрим функцию $f_4(x^{a_1}, \ldots, x^{a_n})$. При $x = 0$ она принимает значение $f_4(0^{a_1}, \ldots, 0^{a_n}) = f_4(\overline{a}_1, \ldots, \overline{a}_n) = C$, при $x = 1$ она принимает значение $f_4(1^{a_1}, \ldots, 1^{a_n}) = f_4(a_1, \ldots, a_n) = C$, следовательно, эта функция — константа C. Значит, $C \in [\mathcal{K}]$. Константа $\overline{C}$ также принадлежит $[\mathcal{K}]$, так как уже показано, что в $[\mathcal{K}]$ содержится $\overline{x}$.

Таким образом, показано, что в любом случае $\{0,\ 1,\ \overline{x}\} \subset [\mathcal{K}]$. Остается показать, что $x_1 \cdot x_2 \in [\mathcal{K}]$.

Рассмотрим функцию f_3. Так как она не лежит в L, то она представляется многочленом Жегалкина степени выше 1, т. е. в этот многочлен входит член, содержащий произведение по крайней мере двух переменных. Пусть это переменные x_1 и x_2. Тогда f_3 можно представить следующим образом:

$$f_3(x_1, \ldots, x_n) = x_1 x_2 \varphi_1(x_3, \ldots, x_n) \oplus \\ \oplus\, x_1 \varphi_2(x_3, \ldots, x_n) \oplus x_2 \varphi_3(x_3, \ldots, x_n) \oplus \\ \oplus\, \varphi_4(x_3, \ldots, x_n),$$

причем функция φ_1 отлична от константы 0, т. е. существует набор $(a_3, \ldots, a_n)$, такой что $\varphi_1(a_3, \ldots, a_n) = 1$. Функция $g(x_1, x_2)$, представимая формулой

$$g(x_1, x_2) \equiv f_3(x_1, x_2, a_3, \ldots, a_n) \equiv \\ \equiv x_1 x_2 \oplus C_1 x_1 \oplus C_2 x_2 \oplus C_3, \quad C_i \in \Omega, \quad i \in \overline{1,3},$$

принадлежит $[\mathcal{K}]$, так как ранее доказано, что в этом классе лежат константы 0 и 1. Тогда

$$g(x_1^{\overline{C}_2}, x_2^{\overline{C}_1}) \equiv (x_1 \oplus C_2)(x_2 \oplus C_1) \oplus \\ \oplus\, C_1(x_1 \oplus C_2) \oplus C_2(x_2 \oplus C_1) \oplus C_3 \equiv \\ \equiv x_1 x_2 \oplus C_3 \oplus C_1 C_2.$$

Таким образом, классу $[\mathcal{K}]$ принадлежит либо функция $x_1 x_2$, либо $\overline{x_1 x_2}$, а поскольку уже известно, что в этом классе содержится «отрицание», то в любом случае $x_1 x_2 \in [\mathcal{K}]$. $\square$

1.3. ФУНКЦИИ k-ЗНАЧНОЙ ЛОГИКИ И СПОСОБЫ ИХ ЗАДАНИЯ. ПОЛНЫЕ СИСТЕМЫ

Пусть Ω_k — множество, состоящее из k элементов, которые мы будем обозначать в виде $0, 1, \ldots, k-1$.

Определение 1.16. *Функцией k-значной логики от n переменных (n-местной функцией) называется любое отображение $f : \Omega_k^n \to \Omega_k$, $n \in \mathbb{N}$. При $n = 0$ функциями k-значной логики называют константы* $0, 1, \ldots, k-1$.

Множество всех функций k-значной логики обозначим F_k, а множество n-местных функций k-значной логики — $F_k(n)$. Аналогично булевому случаю определяются равенство функций, существенные и несущественные переменные, представление функций формулами, замыкание системы функций, полные системы и замкнутые классы. Простейшим способом задания функции k-значной логики является табличный. Мощность множества $F_k(n)$ равняется k^{k^n}.

Множество Ω_k можно рассматривать как кольцо вычетов $\mathbb{Z}_k$, считая, что на нем заданы операции сложения и умножения по модулю k. Тогда каждый многочлен из $\mathbb{Z}_k[x_1, \ldots, x_n]$ стандартным образом определяет некоторую функцию k-значной логики от n переменных. Если k — простое число, то $\mathbb{Z}_k$ является полем и, как хорошо известно (см., например, [14], т. 1, стр. 180), любая функция над $\mathbb{Z}_k$ может быть представлена многочленом (ниже мы вернемся к обсуждению этого случая). Однако в случае составного k это не так:

Утверждение 1.12. *Если k — составное число, то для любого n существуют функции $f \in F_k(n)$, не представимые многочленом над $\mathbb{Z}_k$.*

□ Пусть $P(x_1, \ldots, x_n)$ — многочлен из кольца $\mathbb{Z}_k[x_1, \ldots, x_n]$ и d — собственный делитель числа k. Тогда если для наборов $(a_1, \ldots, a_n)$ и $(b_1, \ldots, b_n)$ из $\mathbb{Z}_k^n$ выполняются сравнения $a_i \equiv b_i \pmod d$, $i \in \overline{1,n}$, то из простейших свойств сравнений следует соотношение

$$P(a_1, \ldots, a_n) \equiv P(b_1, \ldots, b_n) \pmod d.$$

Если некоторая функция k-значной логики не удовлетворяет такому соотношению, то она не представима многочленом. Примером может служить функция f, определенная следующим образом:

$$f(x_1,\dots,x_n)=\begin{cases}0, & \text{если } x_i<d \text{ для всех } i\in\overline{1,n}\\ 1, & \text{в противном случае.}\end{cases}$$

На наборах $(0,\dots,0)$ и $(d,\dots,d)$ эта функция принимает значения 0 и 1 соответственно, и эти значения не сравнимы по модулю d. Следовательно f не представима многочленом над $\mathbb{Z}_k$. $\square$

При больших значениях k и n табличное задание функций из $F_k(n)$ реально не осуществимо. Поэтому основным способом их задания является способ, при котором функция представляется в виде суперпозиции некоторых легко описываемых функций от небольшого числа переменных.

Перечислим некоторые часто используемые функции k-значной логики.

1. $\delta_a(x)=\begin{cases}1, & \text{если } x=a\\ 0, & \text{в противном случае.}\end{cases}$
2. $J_a(x)=\begin{cases}k-1, & \text{если } x=a\\ 0, & \text{в противном случае.}\end{cases}$
3. Сложение по модулю k: x_1+x_2.
4. Умножение по модулю k: $x_1\cdot x_2$.
5. Аналог дизъюнкции: $x_1\vee x_2=\max\{x_1,x_2\}$.
6. Аналог конъюнкции: $x_1\wedge x_2=\min\{x_1,x_2\}$.
7. Отрицание Поста: $\bar{x}=x+1 \pmod k$.
8. Отрицание Лукашевича: $\sim x=k-1-x \pmod k$.
9. Функция Вэбба: $V_k=\overline{x_1\vee x_2}=\max\{x_1,x_2\}+1 \pmod k$.

С помощью этих функций строятся простейшие полные системы функций k-значной логики.

Утверждение 1.13. *Следующие системы функций k-значной логики полны:*

а) $\{0,1,\dots,k-1,\delta_0(x),\dots\delta_{k-1}(x),x_1\cdot x_2,x_1+x_2\}$;

б) $\{0,1,\dots,k-1,\delta_0(x),\dots\delta_{k-1}(x),x_1\cdot x_2,x_1\vee x_2\}$;

в) $\{0,1,\dots,k-1,J_0(x),\dots J_{k-1}(x),x_1\vee x_2,x_1\wedge x_2\}$.

□ а) Покажем, что любая функция $f(x_1, \ldots, x_n)$ из $F_k(n)$ представима формулой

$$f(x_1, \ldots, x_n) \equiv \sum_{(a_1,\ldots,a_n)\in\Omega_k^n} \delta_{a_1}(x_1) \cdot \ldots \cdot \delta_{a_n}(x_n) f(a_1, \ldots, a_n).$$

Действительно, пусть $x_i = b_i$, $i \in \overline{1,n}$. Тогда среди слагаемых в правой части равны нулю все, кроме слагаемого, для которого $a_i = b_i$, $i \in \overline{1,n}$. Это слагаемое, очевидно, принимает значение $f(b_1, \ldots, b_n)$, значит, правая и левая части принимают на одинаковых наборах одинаковые значения.

Для доказательства полноты систем из пунктов *б)* и *в)* достаточно провести аналогичные рассуждения для формул:

$$f(x_1, \ldots, x_n) \equiv \bigvee_{(a_1,\ldots,a_n)\in\Omega_k^n} \delta_{a_1}(x_1) \cdot \ldots \cdot \delta_{a_n}(x_n) f(a_1, \ldots, a_n) \text{ и}$$

$$f(x_1, \ldots, x_n) \equiv \bigvee_{(a_1,\ldots,a_n)\in\Omega_k^n} J_{a_1}(x_1) \wedge \ldots \wedge J_{a_n}(x_n) \wedge f(a_1, \ldots, a_n).$$

Читателю предлагается проделать это самостоятельно. □

Теорема 1.14. *Система функций $\mathcal{K} = \{\bar{x}_1, x_1 \vee x_2\}$ («система Поста») — полна.*

□ 1. Замыкание системы $\mathcal{K}$ содержит любую функцию вида $x_1 + a$, $a \in \overline{1, k-1}$, поскольку каждая такая функция может быть получена a-кратным применением отрицания Поста ($\bar{x}_1$).

2. Константа $k-1$ содержится в замыкании $\mathcal{K}$ в силу справедливости тождества $k - 1 \equiv \bigvee_{a=0}^{k-1}(x_1 + a)$.

3. Из пунктов 1 и 2 следует, что все константы $0, \ldots, k-1$ лежат в замыкании $\mathcal{K}$.

4. Замыкание $\mathcal{K}$ содержит функцию $J_a(x)$, $a \in \Omega_k$. Действительно, нетрудно проверить, что для любого $b \in \Omega_k$ справедливо равенство

$$J_a(b) = \max\{b + i |\ i \in \Omega_k \setminus \{k - 1 - a\}\} + 1,$$

поскольку при $x = a$ максимальный элемент из множества в правой части равен $k - 2$ и правая часть принимает значение

$k-1$, а при $x \neq a$ максимальный элемент из множества в правой части равен $k-1$ и правая часть принимает значение 0. Таким образом, справедливо тождество

$$J_a(x) \equiv \bigvee_{i \neq k-1-a} (x+i) + 1.$$

5. Замыкание $\mathcal{K}$ содержит функцию

$$f_{s,a}(x) = \begin{cases} s, & \text{если } x = a \\ 0, & \text{в противном случае.} \end{cases} \qquad s, a \in \Omega_k.$$

Это следует из тождества

$$f_{s,a}(x) \equiv \max\{J_a(x), k-1-s\} + s + 1.$$

Действительно, если $x = a$, то $\max\{J_a(x), k-1-s\} = k-1$ и правая часть принимает значение s, если же $x \neq a$, то $\max\{J_a(x), k-1-s\} = k-1-s$ и правая часть принимает значение 0.

6. Замыкание $\mathcal{K}$ содержит отрицание Лукашевича ($\sim x$). Это следует из тождества

$$\sim x \equiv \bigvee_{s=0}^{k-1} f_{k-1-s,s}(x).$$

Действительно, $f_{k-1-s,s}(a)$ не равно нулю тогда и только тогда, когда $s = a$. Таким образом, для любого $a \in \Omega_k$ правая часть при $x = a$ принимает значение $f_{k-1-a,a}(a) = k-1-a = \sim a$.

7. Замыкание $\mathcal{K}$ содержит функцию $x_1 \wedge x_2$. Это следует из тождества

$$x_1 \wedge x_2 \equiv \sim ((\sim x_1) \vee (\sim x_2)).$$

Таким образом, из пунктов 3, 4, 7 следует, что замыкание $\mathcal{K}$ содержит все функции, составляющие систему из пункта *в)* утверждения 1.13, и $\mathcal{K}$ — полная система. □

Утверждение 1.15. *Функция Вэбба $V_k(x_1, x_2)$ образует полную систему.*

□ Для доказательства утверждения заметим, что $\bar{x}_1 \equiv V_k(x_1, x_1)$, следовательно, $x + k - 1 \in [\{V_k\}]$. Кроме того, $x_1 \vee x_2 \equiv V_k(x_1, x_2) + k - 1$, и утверждение следует теперь из теоремы 1.14. □

1.4. КРИТЕРИИ ПОЛНОТЫ СИСТЕМ ФУНКЦИЙ k-ЗНАЧНОЙ ЛОГИКИ

Дадим определение, играющее важную роль как в процессе дальнейшего изложения, так и в вопросах полноты систем дискретных функций в самой общей постановке. Вначале напомним, что m-арным отношением на множестве A называется любое подмножество R множества A^m. В частности, бинарным отношением называется любой набор упорядоченных пар множества A, а унарным — любое подмножество из A.

Определение 1.17. *Говорят, что система $\mathcal{K}$ функций k-значной логики сохраняет m-арное отношение R, если для любой функции $f(x_1, \ldots, x_n)$ из $\mathcal{K}$ и любых $a_{ij} \in \Omega_k$, $i \in \overline{1,n}$, $j \in \overline{1,m}$ из того, что*

$$(a_{11}, \ldots, a_{1m}) \in R, \ldots, (a_{n1}, \ldots, a_{nm}) \in R$$

следует

$$(f(a_{11}, \ldots, a_{n1}), \ldots, f(a_{1m}, \ldots, a_{nm})) \in R.$$

В частности, некоторое бинарное отношение R сохраняется, если из того, что $a_1 R b_1, \ldots, a_n R b_n$ следует

$$f(a_1, \ldots, a_n) R f(b_1, \ldots, b_n),$$

а некоторое унарное отношение R сохраняется, если из того, что $a_1 \in R, \ldots, a_n \in R$ следует $f(a_1, \ldots, a_n) \in R$.

Как было отмечено выше, в случае составного k система, состоящая из констант, сложения и умножения по модулю k, не является полной. Следующая теорема дает критерий полноты системы, замыкание которой содержит данные функции. Приводимое доказательство предложено А. А. Нечаевым:

Теорема 1.16 (А. А. Нечаев). *Система функций k-значной логики $\mathcal{K}$ полна тогда и только тогда, когда одновременно выполняются условия:*

1) система $\mathcal{K}$ не сохраняет бинарного отношения сравнимости по каждому из собственных делителей числа k;

2) замыкание системы $\mathcal{K}$ содержит функции $x_1 + x_2$, $x_1 \cdot x_2$ и 1.

□ Необходимость условия 2 для полноты системы $\mathcal{K}$ очевидна по определению полной системы. Предположим, условие 1 не выполнено, т. е. каждая функция из $\mathcal{K}$ сохраняет отношение сравнимости по некоторому собственному делителю числа k. Индукцией по рангу формулы нетрудно доказать, что таким свойством будет обладать любая функция из замыкания $\mathcal{K}$. Однако, как было показано в утверждении 1.12, в F_k всегда найдется функция, не сохраняющая этого отношения. Таким образом, условие 1 также выполняется.

Для доказательства обратного утверждения нам понадобится следующая лемма.

Лемма 1.17. *Если система $\mathcal{K}$ удовлетворяет условиям 1 и 2, то для любых различных a и b из Ω_k (рассматриваемого как кольцо $\mathbb{Z}_k$) множество $M_{a,b}$, определяемое равенством*

$$M_{a,\,b} = \{\, f(a) - f(b) \mid f(x_1) \in [\mathcal{K}]\},$$

совпадает со всем Ω_k.

◁ Из условия 2 следует, что $M_{a,\,b}$ является идеалом кольца $\mathbb{Z}_k$. Действительно, если f и g — функции одного переменного из $[\mathcal{K}]$, то $(f(a) - f(b)) + (g(a) - g(b)) = (f(a) + g(a)) - (f(b) + {} + g(b)) \in M_{a,\,b}$ и для любого r из $\mathbb{Z}_k$:

$$r \cdot (f(a) - f(b)) = r \cdot f(a) - r \cdot f(b) \in M_{a,\,b}.$$

Как хорошо известно (см., например, [14]), идеал кольца $\mathbb{Z}_k$ имеет вид $d\mathbb{Z}_k$ для некоторого делителя d числа k. Покажем, что в данном случае $d = 1$.

Очевидно, что $d \neq k$, поскольку множество $M_{a,\,b}$ всегда содержит ненулевой элемент $a - b$. Предположим, $M_{a,\,b} = d\mathbb{Z}_k$ и

$1 < d < k$. Тогда в системе $\mathcal{K}$ найдется функция $F(x_1, \ldots, x_n)$, не сохраняющая отношения сравнимости по d. А именно, найдутся такие наборы $\alpha = (a_1, \ldots, a_n)$ и $\beta = (b_1, \ldots, b_n)$ из Ω_K^n, что $a_i \equiv b_i \pmod d$, $i \in \overline{1,n}$, но $F(\alpha) \not\equiv F(\beta)(\mathrm{mod} d)$. Поскольку элементы $a_i - b_i$ лежат в $M_{a,b}$, то в $[\mathcal{K}]$ найдутся функции одного переменного $f_1, \ldots, f_n$, такие что

$$a_i - b_i = f_i(a) - f_i(b),\ i \in \overline{1,n}.$$

Рассмотрим функцию одного переменного

$$G(x) = F(b_1 + f_1(x) - f_1(b), \ldots, b_n + f_n(x) - f_n(b)).$$

Несложно заметить, что $G(x) \in [\mathcal{K}]$ и $G(a) = F(\alpha)$, $G(b) = F(\beta)$. Тогда

$$G(a) - G(b) = F(\alpha) - F(\beta) \not\equiv 0(\mathrm{mod} d),$$

и имеем противоречие с тем, что $G(a) - G(b)$ содержится в $M_{a,b} = d\mathbb{Z}_k$. Из полученного противоречия следует, что $d = 1$ и лемма доказана. $\triangleright$

Перейдем к доказательству теоремы. Из леммы следует, что для любых различных a и b в $[\mathcal{K}]$ найдется функция одного переменного φ, такая что $\varphi(a) - \varphi(b) = 1$. Рассмотрим функцию $h_{a,b}(x) = \varphi(x) - \varphi(b)$. Очевидно, $h_{a,b}(a) = 1$, $h_{a,b}(b) = 0$ и $h_{a,b}(x) \in [\mathcal{K}]$. Но тогда в $[\mathcal{K}]$ лежит функция $\delta_a(x)$, поскольку $\delta_a(x) \equiv \prod\limits_{b \in \Omega_k \setminus \{a\}} h_{a,b}(x)$. Таким образом, $[\mathcal{K}]$ содержит в себе систему из утверждения 1.12(а) и, следовательно, $[\mathcal{K}]$ — полная система. $\square$

Следствие 1. *Система $\{x_1 \cdot x_2, x_1 + x_2, 1\}$ из F_k полна тогда и только тогда, когда k — простое число.* $\square$

Пример Пусть $k = 2^n$, $\Omega_k = \{0, 1, \ldots, 2^n - 1\}$. Любой элемент из Ω_k, рассматриваемый как целое число, может быть представлен следующим образом:

$$a = \gamma_{n-1}(a) \cdot 2^{n-1} + \ldots + \gamma_1(a) \cdot 2 + \gamma_0(a),\ \gamma_i(a) \in \{0, 1\}.$$

Рассмотрим функцию $t(x) = 2 \cdot x + \gamma_{n-1}(x) \pmod{2^n}$ («циклический сдвиг»). Поскольку $t(2^{n-1}) = 1$ и $t(0) = 0$, то $t(x)$ не

сохраняет отношения сравнимости ни по одному собственному делителю числа 2^n (все они имеют вид 2^s, $s \in \overline{1, n-1}$).

Приведем без доказательства еще один важный и часто применяемый критерий полноты:

Теорема 1.18 (Критерий Слупецкого). *Пусть $k \geq 3$ и замыкание системы функций k-значной логики K содержит все функции одного переменного. Система K полна тогда и только тогда, когда она содержит функцию, существенно зависящую от не менее чем двух переменных и принимающую все значения из Ω_k.* $\square$

Заметим, что оба приведенных критерия имеют частный характер, поскольку содержат дополнительные предположения: в первом случае — наличие в замыкании системы арифметических операций, во втором — всех функций одного переменного. Такого недостатка лишен критерий Поста полноты системы булевых функций, поскольку для его проверки надо рассматривать свойства функций из самой системы, а не ее замыкания.

После введения понятия сохранения отношения, критерий Поста может быть сформулирован в несколько другой форме. Для этого рассмотрим на множестве Ω_2 набор отношений:

1) унарное отношение R_0, $R_0 = \{0\} \subset \Omega_2$;
2) унарное отношение R_1, $R_1 = \{1\} \subset \Omega_2$;
3) бинарное отношение R_M, $aR_Mb \leftrightarrow a \leq b$;
4) бинарное отношение R_S, $aR_Sb \leftrightarrow a = \bar{b}$.

Несложно заметить теперь, что булева функция f принадлежит классу T_0, T_1, M, S тогда и только тогда, когда она сохраняет соответственно отношения R_0, R_1, R_M, R_S.

Для класса аффинных функций проследить такую закономерность не столь просто, однако и здесь существует соответствующее отношение:

5) 4-арное отношение R_L, $(a, b, c, d) \in R_L \leftrightarrow a \oplus b = c \oplus d$.

Читателю предлагается самостоятельно проверить, что булева функция f принадлежит классу L аффинных функций тогда и только тогда, когда она сохраняет отношение R_L. Таким образом, критерий Поста можно теперь сформулировать так:

Система K булевых функций полна тогда и только тогда, когда она не сохраняет отношений R_0, R_1, R_M, R_S, R_L.

Для случая функций k-значной логики существует аналогичный критерий, известный как теорема Розенберга. Доказан он был в 1960-х гг. Его формулировка заключается в том, что полнота системы равносильна несохранению этой системой набора отношений специального вида, разбитых на шесть групп. Доказательство этой теоремы достаточно объемно и выходит за рамки данного пособия.

1.5. ПОЛИНОМИАЛЬНОЕ ПРЕДСТАВЛЕНИЕ ФУНКЦИЙ k-ЗНАЧНОЙ ЛОГИКИ

Как было показано в утверждении 1.12, в случае составного k всегда найдется функция из F_k, не представимая многочленом над кольцом $\mathbb{Z}_k$. В то же время, если $k = p^t$, где p — простое, а t — натуральное, то множество Ω_k может рассматриваться как множество элементов конечного поля $GF(p^t)$ и тогда, как известно из курса алгебры, любая функция может быть задана полиномиально. Здесь, однако, необходимо добавить, что сложение и умножение в этом случае будет осуществляться не по модулю k, а в соответствии с операциями данного поля. В общем случае представление функции многочленом над полем неоднозначно. Однозначности представления можно добиться, если наложить дополнительные условия на вид многочлена.

Теорема 1.19. *Пусть $f(x_1, \ldots, x_n) \in F_q(n)$, где $q = p^t$, p — простое, $n \in \mathbb{N}$. Тогда f однозначно с точностью до перестановки слагаемых представима полиномом над $GF(q)$ вида*

$$P(x_1 \ldots, x_n) = \sum_{\substack{(s_1, \ldots, s_n) \\ s_i \in \overline{0, q-1}}} a_{s_1 \ldots s_n} \cdot x_1^{s_1} \cdots x_n^{s_n}\ ,\ a_{s_1 \ldots s_n} \in GF(q). \quad (1.8)$$

□ Представимость. Найдем значение формулы вида

$$\sum_{(a_1\ldots,a_n)\in GF(q)} f(a_1,\ldots,a_n)\cdot\prod_{i=1}^{n}(1-(x_i-a_i))^{q-1} \qquad (1.9)$$

на наборе $(b_1,\ldots,b_n)$. Если для некоторого слагаемого внешней суммы, соответствующего набору $(a_1\ldots,a_n)$, найдется такое $j\in\overline{1,n}$, что $b_j\neq a_j$, то $(b_j-a_j)^{q-1}=1$ и данное слагаемое принимает значение 0. Если же для всех j выполняется $a_j=b_j$, то слагаемое, очевидно, примет значение $f(a_1,\ldots,a_n)=f(b_1,\ldots,b_n)$. Таким образом, данная формула представляет функцию $f(x_1\ldots,x_n)$. Несложно увидеть, что, если в формуле (1.9) раскрыть скобки и привести подобные слагаемые, мы получим многочлен вида (1.8).

Однозначность. Заметим, что число различных многочленов вида (1.8) равно k^{k^n}, т. е. совпадает с мощностью множества $F_k(n)$. Поскольку разным функциям соответствуют, очевидно, различные многочлены, то соответствие между множествами функций и многочленов является взаимно однозначным. □

Следующее утверждение дает удобный способ нахождения коэффициентов многочлена над $GF(q)$ вида (1.8), представляющего функцию q-значной логики одного переменного.

Утверждение 1.20. *Пусть* $f(x_1)=\sum\limits_{i=0}^{q-1}u_ix_1^i$ *— многочлен над* $GF(q)$, $q\geq 2$. *Тогда справедливо матричное равенство:*

$$\begin{pmatrix} u_0\\ u_1\\ \vdots\\ u_{q-1}\end{pmatrix}=\begin{pmatrix} 1-a_0^{q-1} & 1-a_1^{q-1} & \ldots & 1-a_{q-1}^{q-1}\\ -a_0^{q-2} & -a_1^{q-2} & \ldots & -a_{q-1}^{q-2}\\ \vdots & \vdots & \ddots & \vdots\\ -1 & -1 & \ldots & -1\end{pmatrix}\times$$

$$\times\begin{pmatrix} f(a_0)\\ f(a_1)\\ \vdots\\ f(a_{q-1})\end{pmatrix}, \qquad (1.10)$$

где $a_0, \ldots, a_{q-1}$ — занумерованные в некотором порядке элементы поля $GF(q)$.

□ Несложно видеть, что

$$A \cdot \begin{pmatrix} u_0 \\ u_1 \\ \vdots \\ u_{q-1} \end{pmatrix} = \begin{pmatrix} f(a_0) \\ f(a_1) \\ \vdots \\ f(a_{q-1}) \end{pmatrix},$$

где

$$A = \begin{pmatrix} 1 & a_0 & \ldots & a_0^{q-1} \\ 1 & a_1 & \ldots & a_1^{q-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_{q-1} & \ldots & a_{q-1}^{q-1} \end{pmatrix}.$$

Для завершения доказательства остается установить, что матрица A обратна к матрице, стоящей в правой части равенства (1.10). Этот факт легко проверяется непосредственно, проверка предоставляется читателю в качестве упражнения. □

Поскольку не любая функция k-значной логики представима многочленом над кольцом $\mathbb{Z}_k$, то представляет интерес задача нахождения необходимых и достаточных условий такой представимости. Как хорошо известно из курса алгебры, любое кольцо вычетов раскладывается в прямую сумму колец вычетов по примарным модулям, поэтому задача представимости функции многочленом над произвольным кольцом сводится к аналогичной задаче над кольцом вида $\mathbb{Z}_{p^n}$, где p — простое число.

Пусть $k = p_1^{m_1} \ldots p_t^{m_t}$. В силу изоморфизма

$$\mathbb{Z}_k \cong \mathbb{Z}_{p_1^{m_1}} \oplus \ldots \oplus \mathbb{Z}_{p_t^{m_t}}$$

каждой функции $F : \mathbb{Z}_k^n \to \mathbb{Z}_k$ соответствует единственный набор $(F_1, \ldots F_t)$ функций $F_i : \mathbb{Z}_k^n \to \mathbb{Z}_{p_i^{m_i}}$, $i \in \overline{1,t}$.

Если функция F сохраняет бинарные отношения сравнимости по модулю $p_i^{m_i}$, т. е. при всех наборах аргументов $(a_1, \ldots, a_n)$, $(b_1, \ldots, b_n) \in \mathbb{Z}_k^n$, таких что $a_j \equiv b_j \pmod{p_i^{m_i}}$, $j \in \overline{1,n}$, выполнены сравнения $F(a_1, \ldots, a_n) \equiv F(b_1, \ldots, b_n)$

$(\mathrm{mod}\ p_i^{m_i})$, $i \in \overline{1,t}$, то вместо функций F_i можно взять функции $F_i' : \mathbb{Z}^n_{p_i^{m_i}} \to \mathbb{Z}_{p_i^{m_i}}$, которые определяются однозначно по значениям функций F_i на остатках от деления аргументов на $p_i^{m_i}$, $i \in \overline{1,t}$.

Теорема 1.21. *Пусть $k = p_1^{m_1} \dots p_t^{m_t}$. Для полиномиальности функции $F(x_1, \dots, x_n)$ над $\mathbb{Z}_k$ необходимо и достаточно, чтобы выполнялись условия:*

1) F сохраняет отношения сравнимости по модулю $p_i^{m_i}$, $i \in \overline{1,t}$ и

2) в случае выполнимости условия 1 функции F_i' полиномиальны над $\mathbb{Z}^n_{p_i^{m_i}}$, $i \in \overline{1,t}$.

□ Необходимость условий 1 и 2 очевидна. Докажем достаточность.

В силу условия 1 функция F однозначно определяется набором функций $F_1', \dots, F_t'$. По условию 2 каждая из функций F_i' задается некоторым многочленом f_i над $\mathbb{Z}^n_{p_i^{m_i}}$, $i \in \overline{1,t}$. В силу изоморфизма

$$\mathbb{Z}_k[x_1, \dots, x_n] \cong \mathbb{Z}_{p_1^{m_1}}[x_1, \dots, x_n] \oplus \dots \oplus \mathbb{Z}_{p_t^{m_t}}[x_1, \dots, x_n]$$

набору многочленов $(f_1, \dots, f_t)$ соответствует однозначно некоторый многочлен $f(x_1, \dots, x_n)$ над $\mathbb{Z}_k$.

Покажем, что многочлен $f(x_1, \dots, x_n)$ задает функцию $F(x_1, \dots, x_n)$. Пусть ψ — изоморфизм

$$\psi : \mathbb{Z}_k \to \mathbb{Z}_{p_1^{m_1}} \oplus \dots \oplus \mathbb{Z}_{p_t^{m_t}}.$$

Тогда для любого набора аргументов $(a_1, \dots, a_n) \in \mathbb{Z}_k^n$ имеем

$$\psi(a_j) = (a_j^{(1)}, \dots, a_j^{(t)}), \quad j \in \overline{1,n},$$

$$\psi(F(a_1, \dots, a_n)) = (F_1'(a_1^{(1)}, \dots, a_n^{(1)}), \dots, F_t'(a_1^{(t)}, \dots, a_n^{(t)})) =$$

$$= (f_1(a_1^{(1)}, \dots, a_n^{(1)}), \dots, f_t(a_1^{(t)}, \dots, a_n^{(t)})) = \psi(f(a_1, \dots, a_n)).$$

Поскольку ψ — изоморфизм, получаем $F(a_1, \dots, a_n) = f(a_1, \dots, a_n)$. □

Решение вопроса о полиномиальности функции над кольцом $\mathbb{Z}_{p^n}$ в случае простого p мы рассмотрим для функции одного переменного.

Докажем вначале вспомогательную лемму.

Лемма 1.22. *Для любых неотрицательных целых чисел* r, c, l *справедливо утверждение*

$$r! \text{ делит число } \sum_{s=0}^{r}(-1)^{r-s}\left(\begin{array}{c} r \\ s \end{array}\right)(c+s)^l.$$

□ 1. Пусть $c = 0$. Из комбинаторного анализа хорошо известна формула для числа $D(l, r)$ сюръективных отображений множества X мощности l в множество Y мощности r (см., например, [14] или [57]):

$$D(l,r) = \sum_{s=0}^{r}(-1)^{r-s}\left(\begin{array}{c} r \\ s \end{array}\right)s^l.$$

Заметим, что при $l < r$ соответствующая сумма обращается в 0.

Пусть $\sigma : Y \to Y$ — биекция. Очевидно, что $f : X \to Y$ — сюръективное отображение тогда и только тогда, когда $f \cdot \sigma$ сюръективно. Зададим на множестве сюръективных отображений отношение $\sim$:

$f_1 \sim f_2$ тогда и только тогда, когда найдется биекция σ, такая что $f_1 = f_2 \cdot \sigma$.

Очевидно, что данное отношение является отношением эквивалентности. Оно разбивает все множество сюръективных отображений на классы, причем мощность каждого класса равна мощности множества всех биективных преобразований множества Y, т. е. $r!$. Следовательно, число $r!$ делит сумму $\sum\limits_{s=0}^{r}(-1)^{r-s}\left(\begin{array}{c} r \\ s \end{array}\right)s^l$.

2. Если $c \neq 0$, то справедливы равенства:

$$\sum_{s=0}^{r}(-1)^{r-s}\left(\begin{array}{c} r \\ s \end{array}\right)(c+s)^l = \\ = \sum_{s=0}^{r}(-1)^{r-s}\left(\begin{array}{c} r \\ s \end{array}\right)\sum_{j=0}^{l}\left(\begin{array}{c} l \\ j \end{array}\right)s^j c^{l-j} = \\ = \sum_{j=0}^{l}\left(\begin{array}{c} l \\ j \end{array}\right)c^{l-j}\sum_{s=0}^{r}(-1)^{r-s}\left(\begin{array}{c} r \\ s \end{array}\right)s^j,$$

и поскольку каждая внутренняя сумма в последнем выражении кратна $r!$, то $r!$ делит исходное число. □

Обозначим для данных p, n, r:

$$\mu(r) = \max\{t : p^t | r!\},\ \nu(r) = \min(n, \mu(r)).$$

Легко проверяется

Следствие 1. *Пусть $P(x)$ — многочлен над кольцом $\mathbb{Z}_{p^n}$, p — простое число. Тогда для любых неотрицательных целых чисел r, c*

$$\sum_{s=0}^{r}(-1)^{r-s}\left(\begin{array}{c} r \\ s \end{array}\right)P(c+s) \equiv 0 \pmod{p^{\nu(r)}}.\ \square$$

Сформулируем и докажем критерий полиномиальности функции одного переменного:

Теорема 1.23. *Функция $\varphi(x_1)$, заданная на кольце $\mathbb{Z}_{p^n}$, представима полиномом над этим кольцом тогда и только тогда, когда для любого $c \in \mathbb{Z}_{p^n}$ и для любого целого неотрицательного r выполняется сравнение*

$$\sum_{s=0}^{r}(-1)^{r-s}\left(\begin{array}{c} r \\ s \end{array}\right)\varphi(c+s) \equiv 0 \pmod{p^{\nu(r)}},$$

где сложение в аргументе функции ведется по модулю p^n.

□ Справедливость прямого утверждения вытекает из следствия к лемме 1.22. Докажем обратное утверждение. Введем обозначение[1]

$$\Delta^r\varphi(c)=\sum_{s=0}^{r}(-1)^{r-s}\left(\begin{array}{c} r \\ s \end{array}\right)\varphi(c+s).$$

Рассмотрим многочлен $f(x)$ над полем рациональных чисел следующего вида:

$$f(x)=\sum_{j=0}^{N}\frac{1}{j!}\Delta^j\varphi(0)x(x-1)\dots(x-j+1),$$

где N — некоторое натуральное число. Тогда для любого $a\in$ $\in\overline{0,p^n-1}$ имеем

$$\begin{aligned}
f(a)&=\sum_{j=0}^{N}\frac{1}{j!}\Delta^j\varphi(0)a(a-1)\dots(a-j+1)=\\
&=\sum_{j=0}^{N}\left(\begin{array}{c} a \\ j \end{array}\right)\Delta^j\varphi(0)=\sum_{j=0}^{N}\left(\begin{array}{c} a \\ j \end{array}\right)\sum_{s=0}^{j}(-1)^{j-s}\left(\begin{array}{c} j \\ s \end{array}\right)\varphi(s)=\\
&=\sum_{s=0}^{N}\varphi(s)\sum_{j=s}^{N}(-1)^{j-s}\left(\begin{array}{c} j \\ s \end{array}\right)\left(\begin{array}{c} a \\ j \end{array}\right)=\\
&=\sum_{s=0}^{N}\varphi(s)\sum_{j=s}^{N}(-1)^{j-s}\left(\begin{array}{c} a \\ s \end{array}\right)\left(\begin{array}{c} a-s \\ j-s \end{array}\right)=\\
&=\sum_{s=0}^{N}\left(\begin{array}{c} a \\ s \end{array}\right)\varphi(s)\sum_{j=s}^{N}(-1)^{j-s}\left(\begin{array}{c} a-s \\ j-s \end{array}\right).
\end{aligned}$$

Пусть $N>p^n-1$. Заметим, что при $j>a$ слагаемые суммы

$$\sum_{j=0}^{N}\frac{1}{j!}\Delta^j\varphi(0)a(a-1)\dots(a-j+1)$$

[1] На самом деле, символом Δ^r обозначается оператор конечной разности порядка r, а приводимое равенство представляет одно из свойств конечных разностей (см., например, [9]).

обращаются в нуль, поэтому суммировать по j можно только до значения a. Рассмотрим тогда внутреннюю сумму в выражении для $f(a)$. Имеем

$$\sum_{j=s}^{N}(-1)^{j-s}\left(\begin{array}{c} a-s \\ j-s \end{array}\right) = \\ = \sum_{j=s}^{a}(-1)^{j-s}\left(\begin{array}{c} a-s \\ j-s \end{array}\right) = \sum_{j'=0}^{a-s}(-1)^{j'}\left(\begin{array}{c} a-s \\ j' \end{array}\right).$$

Несложно видеть, что, в силу свойств биномиальных коэффициентов, последняя сумма равна 0, если $a \neq s$, и равна 1, если $a = s$. Таким образом, $f(a) = \varphi(a)$ для всех $a \in \overline{0, p^n - 1}$.

Теперь докажем, что найдется многочлен $F(x)$ над $\mathbb{Z}$, такой что $F(a) = f(a)$ для всех $a \in \overline{0, p^n - 1}$. Для этого рассмотрим сравнение

$$j!x \equiv \Delta^j\varphi(0)(\mathrm{mod}\, p^n). \tag{1.11}$$

Оно разрешимо для любого $j \in \overline{0, p^n - 1}$. Действительно, если $\nu(j) = n$, то $\Delta^j\varphi(0) \equiv 0(\mathrm{mod}\, p^n)$ и решением является $x = 0$. Если же $\nu(j) = \mu(j)$, то по определению $\mu(j)$ имеем $j! = p^{\mu(j)}A_j$, для некоторого $A_j \in \mathbb{Z}$, где $(A_j, p) = 1$ и по условию теоремы $\Delta^j\varphi(0) = p^{\mu(j)}B_j$ для некоторого B_j. Тогда сравнение (1.11) является следствием сравнения $A_j x \equiv B_j(\mathrm{mod}\, p^{n-\mu(j)})$, которое разрешимо в силу взаимной простоты p и A_j.

Для каждого j зафиксируем какое-либо решение c_j сравнения (1.11) и рассмотрим многочлен с целыми коэффициентами:

$$F(x) = \sum_{j=0}^{N} c_j x(x-1)\cdots(x-j+1).$$

Для $a \in \overline{0, p^n - 1}$ имеем

$$F(a) - f(a) = \sum_{j=0}^{N} \left(c_j - \frac{\Delta^j \varphi(0)}{j!} \right) a(a-1)\cdots(a-j+1) =$$

$$= \sum_{j=0}^{N} (c_j \cdot j! - \Delta^j \varphi(0)) \frac{a(a-1)\cdots(a-j+1)}{j!} =$$

$$= \sum_{j=0}^{N} p^n k_j \binom{a}{j}$$

при некоторых целых k_j. Таким образом, $F(a) \equiv f(a) \equiv \varphi(a) \pmod{p^n}$. Следовательно, функцию φ можно представить многочленом над $\mathbb{Z}_{p^n}$. Такой многочлен получается из $F(x)$ приведением его коэффициентов по модулю p^n. $\square$

На самом деле для выяснения вопроса о полиномиальной представимости функции нет необходимости проверять столь сильное достаточное условие. Очевидно, справедливо

Следствие 1. *Функция $\varphi(x_1)$, заданная на кольце $\mathbb{Z}_{p^n}$, представима полиномом над этим кольцом тогда и только тогда, когда*

$$\Delta^r \varphi(0) \equiv 0 \pmod{p^{\nu(r)}}$$

для всех r, $0 \leq r < p^n$.

Глава 2

ПРЕДСТАВЛЕНИЕ ДИСКРЕТНЫХ ФУНКЦИЙ В БАЗИСАХ ФУНКЦИОНАЛЬНЫХ ПРОСТРАНСТВ

В этой главе будет использован подход к описанию дискретной функции как элемента линейного пространства функций над некоторым полем, которое включает в себя область значений данной функции. Такой подход оказывается достаточно эффективным при получении различных, в том числе рассмотренных выше, представлений дискретных функций и описании их свойств.

2.1. БАЗИСЫ ЛИНЕЙНЫХ ФУНКЦИОНАЛЬНЫХ ПРОСТРАНСТВ. БАЗИС ХАРАКТЕРОВ

Определение 2.1. *Пусть K — произвольное поле, 0 и 1 — нуль и единица этого поля соответственно. Псевдобулевой функцией от n переменных (n-местной псевдобулевой функцией) называется произвольное отображение $f : (\Omega = \{0,1\})^n \to K$.*

Очевидно, что частным случаем псевдобулевых функций являются булевы функции. Мы будем также рассматривать и обобщения псевдобулевых функций, а именно отображения вида $f : (GF(p))^n \to K$, где $GF(p)$ — конечное поле из p элементов $\{0, \ldots, p-1\}$, p — простое число. Множество таких функций будем обозначать $K_p(n)$. На $K_p(n)$ зададим операции сложения и умножения на элементы поля K следующим образом:

$$(f_1 + f_2)(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x}),\ (r \cdot f_1)(\mathbf{x}) = r \cdot (f_1(\mathbf{x})).$$

Утверждение 2.1. *Относительно заданных операций множество $K_p(n)$ является векторным пространством над K размерности p^n.*

□ Проверка аксиом векторного пространства проводится непосредственно и предоставляется читателю. В качестве базиса данного пространства можно указать следующий:

$$\{f_\alpha(\mathbf{x})|\alpha \in \{0,\dots,p-1\}^n\},\quad f_\alpha(\mathbf{x}) = \begin{cases} 1, & \text{если } \mathbf{x} = \alpha \\ 0, & \text{если } \mathbf{x} \neq \alpha \end{cases}.$$

Действительно, линейная независимость такой системы очевидна. Кроме того, для любой функции $F(\mathbf{x})$:

$$F(\mathbf{x}) = \sum_{\alpha \in \overline{\{0,p-1\}}^n} F(\alpha) f_\alpha(\mathbf{x}). \quad \square$$

Введем в рассмотрение один из базисов пространства $K_p(n)$, разложение функций по которому часто используется на практике.

Рассмотрим аддитивную группу поля $GF(p^n)$. Как известно, она может быть представлена как группа векторов пространства $GF(p)^n$ по операции сложения.

Рассмотрим множество гомоморфизмов φ этой группы в мультипликативную группу поля комплексных чисел.

Теорема 2.2. *Множество всех гомоморфизмов φ, $\varphi : (GF(p)^n, +) \to (\mathbb{C}, \cdot)$, состоит из p^n различных гомоморфизмов χ_α, $\alpha = (a_1,\dots,a_n) \in GF(p)^n$, каждый из которых однозначно определяется своим действием на векторах $\mathbf{e_j}$, $j \in \overline{1,n}$ стандартного базиса следующим образом:*

$$\chi_\alpha(\mathbf{e_j}) = e^{\frac{2\pi i}{p} a_j}.$$

□ Так как φ — гомоморфизм, то для любого вектора $\gamma = (c_1,\dots,c_n) \in GF(p)^n$ имеем

$$\varphi(\gamma) = \varphi(\sum_{j=1}^{n} c_j \mathbf{e_j}) = \prod_{j=1}^{n} \varphi(\mathbf{e_j})^{c_j},$$

и, следовательно, любой гомоморфизм определяется действием на векторы $\mathbf{e_j}$. Далее, для любого j справедливы соотношения:

$$1 = \varphi(0) = \varphi(\underbrace{\mathbf{e_j} + \mathbf{e_j} + \ldots + \mathbf{e_j}}_{p \text{ слагаемых}}) = \varphi(\mathbf{e_j})^p.$$

Таким образом, $\varphi(\mathbf{e_j})$ является корнем степени p из единицы и $\varphi(\mathbf{e_j}) = e^{\frac{2\pi i}{p}k}$ для некоторого $k \in \overline{0, p-1}$. Следовательно, φ является одним из p^n гомоморфизмов вида χ_α, определенных в условии теоремы. $\square$

Заметим, что если $\beta = (b_1, \ldots, b_n)$, то

$$\chi_\alpha(\beta) = \prod_{j=1}^{n} (e^{\frac{2\pi i}{p}a_j})^{b_j} = e^{\frac{2\pi i}{p}(a_1b_1 + \ldots + a_nb_n)},$$

и, обозначая $< \alpha, \beta >= a_1b_1 + \ldots + a_nb_n$, где сумма берется по модулю p, имеем окончательно $\chi_\alpha(\beta) = e^{\frac{2\pi i}{p}<\alpha,\beta>}$. Гомоморфизмы χ_α называются *аддитивными характерами поля* $GF(p^n)$.

Одно из основных свойств характеров будет доказано в следующем утверждении.

Утверждение 2.3 (Соотношение ортогональности). *Для любых* α, β *из* $GF(p)^n$ *выполняется равенство*

$$\frac{1}{p^n} \sum_{\gamma \in GF(p)^n} \chi_\alpha(\gamma)\overline{\chi_\beta(\gamma)} = \delta_{\alpha,\beta} = \begin{cases} 1, & \text{если } \alpha = \beta \\ 0, & \text{если } \alpha \neq \beta \end{cases}.$$

$\square$ Рассмотрим выражение из левой части:

$$\frac{1}{p^n} \sum_{\gamma \in GF(p)^n} \chi_\alpha(\gamma)\overline{\chi_\beta(\gamma)} =$$

$$= \frac{1}{p^n} \sum_{\gamma \in GF(p)^n} e^{\frac{2\pi i}{p}<\alpha,\gamma>} e^{-\frac{2\pi i}{p}<\beta,\gamma>} =$$

$$= \frac{1}{p^n} \sum_{\gamma \in GF(p)^n} e^{\frac{2\pi i}{p}<\alpha-\beta,\gamma>}.$$

Если $\alpha = \beta$, то $< \alpha - \beta, \gamma >= 0$ и выражение принимает значение 1.

Пусть $\alpha - \beta = (d_1, \ldots, d_n) \neq (0, \ldots, 0)$. Тогда в поле $GF(p)$ уравнение $d_1 x_1 + \ldots + d_n x_n = C$ для любого C имеет ровно p^{n-1} решений. Поэтому, если γ пробегает все значения из $GF(p)^n$, то $< \alpha - \beta, \gamma >$ будет принимать все значения из $\overline{0, p-1}$ ровно по p^{n-1} раз. Следовательно,

$$\sum_{\gamma \in GF(p)^n} e^{\frac{2\pi i}{p} < \alpha - \beta, \gamma >} = p^{n-1} \sum_{k=0}^{p-1} e^{\frac{2\pi i}{p} k} = p^{n-1} \frac{e^{\frac{2\pi i}{p} p} - 1}{e^{\frac{2\pi i}{p}} - 1} = 0. \square$$

Покажем, что система характеров образует базис $\mathbb{C}_p(n)$.

Теорема 2.4. $\{\chi_\alpha | \, \alpha \in GF(p)^n\}$ — *базис пространства* $\mathbb{C}_p(n)$.

$\square$ Поскольку количество характеров равно размерности $\mathbb{C}_p(n)$, то достаточно показать линейную независимость системы характеров. Пусть для некоторого набора $\{C_\alpha\}_{\alpha \in GF(p)^n} \in \in \mathbb{C}$

$$\sum_{\alpha \in GF(p)^n} C_\alpha \chi_\alpha(\mathbf{x}) = 0.$$

Для произвольного $\beta \in GF(p)^n$ умножим обе части последнего равенства на $\overline{\chi_\beta}$ и просуммируем по всем $\mathbf{x} \in GF(p)^n$:

$$\sum_{\mathbf{x} \in GF(p)^n} \sum_{\alpha \in GF(p)^n} C_\alpha \chi_\alpha(\mathbf{x}) \overline{\chi_\beta(\mathbf{x})} = 0.$$

Поменяв порядок суммирования, получим

$$\sum_{\alpha \in GF(p)^n} C_\alpha \sum_{\mathbf{x} \in GF(p)^n} \chi_\alpha(\mathbf{x}) \overline{\chi_\beta(\mathbf{x})} = \\ = p^n \sum_{\alpha \in GF(p)^n} C_\alpha \delta_{\alpha,\beta} = p^n C_\beta = 0.$$

Таким образом, все коэффициенты C_α равны 0, и система линейно независима. $\square$

2.2. ПРЕОБРАЗОВАНИЕ ФУРЬЕ. КОЭФФИЦИЕНТЫ ФУРЬЕ И УОЛША–АДАМАРА

Определение 2.2. *Разложение произвольной функции f из $\mathbb{C}_p(n)$ по базису характеров $\{\chi_\alpha | \, \alpha \in GF(p)^n\}$*

$$f(\mathbf{x}) = \sum_{\alpha \in GF(p)^n} C_\alpha^f \chi_\alpha(\mathbf{x})$$

называется разложением f в ряд Фурье, а комплексное число C_α^f — коэффициентом Фурье, соответствующим набору α. Отображение $\mathbb{C}_p(n)$ в $\mathbb{C}^{2^n}$ сопоставляющее каждой функции набор (вектор) ее коэффициентов Фурье («спектр Фурье») называется преобразованием Фурье.

В некоторых источниках это преобразование называют преобразованием Уолша–Адамара первого рода (соответственно C_α^f — коэффициентами Уолша–Адамара первого рода, а их упорядоченный набор — спектром Уолша–Адамара первого рода).

В следующем утверждении приводится формула для вычисления коэффициентов Фурье:

Утверждение 2.5. *Пусть $\gamma \in GF(p)^n$. Тогда*

$$C_\gamma^f = \frac{1}{p^n} \sum_{\beta \in GF(p)^n} f(\beta)\overline{\chi_\gamma(\beta)}.$$

□ Доказательство следует из цепочки равенств:

$$\begin{aligned}
&\frac{1}{p^n} \sum_{\beta \in GF(p)^n} f(\beta)\overline{\chi_\gamma(\beta)} = \\
&\qquad = \frac{1}{p^n} \sum_{\beta \in GF(p)^n} \sum_{\alpha \in GF(p)^n} C_\alpha^f \chi_\alpha(\beta)\overline{\chi_\gamma(\beta)} = \\
&\qquad = \sum_{\alpha \in GF(p)^n} C_\alpha^f \frac{1}{p^n} \sum_{\beta \in GF(p)^n} \chi_\alpha(\beta)\overline{\chi_\gamma(\beta)} = \\
&\qquad\qquad = \sum_{\alpha \in GF(p)^n} C_\alpha^f \delta_{\alpha,\gamma} = C_\gamma^f. \; \square
\end{aligned}$$

В случае, когда $p = 2$, а f принимает только значения $\{0, 1\}$ (т. е. f — булева функция), удобно вычислять коэффициенты Фурье по следующим формулам.

Утверждение 2.6. *Пусть $f(\mathbf{x})$ — булева функция. Тогда для нулевого набора θ*

$$C_\theta^f = \frac{1}{2^n}\|f(\mathbf{x})\|,$$

а для $\alpha \neq \theta$

$$C_\alpha^f = \frac{1}{2^n}\|f(\mathbf{x}) \oplus < \alpha, \mathbf{x} > \| - \frac{1}{2}.$$

□ Выписывая выражение для коэффициента Фурье, в условиях теоремы имеем

$$C_\alpha^f = \frac{1}{2^n}\sum_{\beta\in\Omega_2^n} f(\beta)\overline{\chi_\alpha(\beta)} = \frac{1}{2^n}\sum_{\beta\in\Omega_2^n} f(\beta)(-1)^{<\alpha,\beta>} =$$

$$= \frac{1}{2^n}\left(\sum_{\beta\in\Omega'} 1 - \sum_{\beta\in\Omega''} 1\right),$$

где

$$\Omega' = \{\beta | f(\beta) = 1 \text{ и } < \alpha, \beta >= 0\},$$
$$\Omega'' = \{\beta | f(\beta) = 1 \text{ и } < \alpha, \beta >= 1\}.$$

Найдем мощности множеств Ω' и Ω'':

$$|\Omega'| = \|f(\mathbf{x}) \cdot (< \alpha, \mathbf{x} > \oplus 1)\| = \|f(\mathbf{x}) \cdot < \alpha, \mathbf{x} > \oplus f(\mathbf{x})\| =$$
$$= \|f(\mathbf{x}) \cdot < \alpha, \mathbf{x} > \| + \|f(\mathbf{x})\| - 2\|f(\mathbf{x}) \cdot < \alpha, \mathbf{x} > \cdot f(\mathbf{x})\| =$$
$$= \|f(\mathbf{x})\| - \|f(\mathbf{x}) \cdot < \alpha, \mathbf{x} > \|;$$
$$|\Omega''| = \|f(\mathbf{x}) \cdot < \alpha, \mathbf{x} > \|.$$

Таким образом,

$$C_\alpha^f = \frac{1}{2^n}\left(\|f(\mathbf{x})\| - 2 \cdot \|f(\mathbf{x}) \cdot < \alpha, \mathbf{x} > \|\right) =$$
$$= \frac{1}{2^n}\left(\|f(\mathbf{x})\| + \| < \alpha, \mathbf{x} > \| - \right.$$
$$\left. -2 \cdot \|f(\mathbf{x}) \cdot < \alpha, \mathbf{x} > \| - \| < \alpha, \mathbf{x} > \|\right).$$

В итоге, при $\alpha = \theta$ имеем $\| < \alpha, \mathbf{x} > \| = 0$ и

$$C_\theta^f = \frac{1}{2^n} \| f(\mathbf{x}) \|,$$

а при $\alpha \neq \theta$

$$C_\alpha^f = \frac{1}{2^n} \| f(\mathbf{x}) \oplus < \alpha, \mathbf{x} > \| - \frac{1}{2},$$

поскольку любая линейная ненулевая функция, очевидно, устойчива (равновероятна). □

В некоторых случаях вместо свойств булевой функции f удобнее рассматривать свойства функции $F = (-1)^{f(\mathbf{x})}$. Коэффициенты Фурье этой функции называют также коэффициентами Уолша–Адамара второго рода функции $f(\mathbf{x})$ и обозначают W_α^f (таким образом, $W_\alpha^f = C_\alpha^F$).

Справедливы равенства, доказательство которых предоставляется читателю:

Утверждение 2.7.

$$W_\alpha^f = 1 - \frac{1}{2^{n-1}} \| f(\mathbf{x}) \oplus < \alpha, \mathbf{x} > \|. \; \square$$

Следствие 1.

$$W_\alpha^f = \begin{cases} -2C_\alpha^f, & \alpha \neq \theta \\ 1 - 2C_\alpha^f, & \alpha = \theta \end{cases} . \quad \square$$

В дальнейшем нами также будут использоваться следующие свойства коэффициентов Уолша–Адамара второго рода:

Утверждение 2.8.

$$\sum_{\alpha \in \Omega_2^n} W_\alpha^f = (-1)^{f(\theta)} \; , \quad \sum_{\alpha \in \Omega_2^n} (W_\alpha^f)^2 = 1.$$

□ Требуемые соотношения получаются следующими преобразованиями:

$$\sum_{\alpha\in\Omega_2^n} W_\alpha^f = \sum_{\alpha\in\Omega_2^n} \frac{1}{2^n} \sum_{\beta\in\Omega_2^n} (-1)^{f(\beta)\oplus<\alpha,\beta>} =$$

$$= \frac{1}{2^n} \sum_{\beta\in\Omega_2^n} (-1)^{f(\beta)} \sum_{\alpha\in\Omega_2^n} (-1)^{<\alpha,\beta>} =$$

$$= \sum_{\beta\in\Omega_2^n} (-1)^{f(\beta)} \delta_{\beta,\theta} = (-1)^{f(\theta)}.$$

$$\sum_{\alpha\in\Omega_2^n} (W_\alpha^f)^2 = \sum_{\alpha\in\Omega_2^n} \left(\frac{1}{2^n} \sum_{\beta\in\Omega_2^n} (-1)^{f(\beta)\oplus<\alpha,\beta>} \times \right.$$

$$\left. \times \frac{1}{2^n} \sum_{\gamma\in\Omega_2^n} (-1)^{f(\gamma)\oplus<\alpha,\gamma>} \right) =$$

$$= \frac{1}{2^{2n}} \sum_{\beta\in\Omega_2^n} \sum_{\gamma\in\Omega_2^n} (-1)^{f(\beta)\oplus f(\gamma)} \sum_{\alpha\in\Omega_2^n} (-1)^{<\alpha,\beta\oplus\gamma>} =$$

$$= \frac{1}{2^n} \sum_{\beta\in\Omega_2^n} \sum_{\gamma\in\Omega_2^n} (-1)^{f(\beta)\oplus f(\gamma)} \delta_{\beta,\gamma} = 1.$$

Заметим, что равенство

$$\sum_{\alpha\in\Omega_2^n} (-1)^{<\alpha,\beta\oplus\gamma>} = \delta_{\beta,\gamma}$$

есть следствие соотношений ортогональности (утверждение 2.3). □

Следствие 1.

$$\frac{1}{2^{n/2}} \le \max_\alpha |W_\alpha^f| \le 1. \quad \square$$

2.3. МАТРИЧНЫЙ ПОДХОД К ПРЕДСТАВЛЕНИЮ БУЛЕВЫХ ФУНКЦИЙ

Зафиксируем некоторую обратимую $2^n \times 2^n$ матрицу A над полем K. Пусть $f^\downarrow$ — вектор-столбец значений функции f из $K_2(n)$ (при лексикографическом упорядочении векторов аргументов). Обозначим

$$\tilde{f}^{\downarrow} = A^{-1} \cdot f^{\downarrow} \text{ (т. е. } f^{\downarrow} = A \cdot \tilde{f}^{\downarrow}). \tag{2.1}$$

Для набора α из Ω_2^n обозначим $\tilde{f}(\alpha)$ координату вектора $\tilde{f}^{\downarrow}$, соответствующую этому набору. Очевидно, что равенством (2.1) задается биективное отображение множества $K_2(n)$ во множество векторов K^{2^n}. Вектор $\tilde{f}^{\downarrow}$ будем называть *представлением* функции f.

Если столбцы матрицы A занумеровать наборами α из Ω_2^n и обозначить их $g_\alpha^{\downarrow}$, то будет справедливо равенство

$$f^{\downarrow} = \sum_{\alpha \in \Omega_2^n} g_\alpha^{\downarrow} \tilde{f}(\alpha). \tag{2.2}$$

Каждый столбец $g_\alpha^{\downarrow}$ задает некоторую функцию g_α из $K_2(n)$, и так как матрица A невырождена, то набор $\{g_\alpha\}_{\alpha \in \Omega_2^n}$ образует базис функционального пространства $K_2(n)$, а вектор $\tilde{f}^{\downarrow}$ есть вектор координат функции f в разложении ее по этому базису и, значит, задает функцию однозначно. С каждой невырожденной матрицей, таким образом, может быть связано некоторое представление. Оказывается, что ряд уже известных нам представлений булевых функций может быть получен с помощью матриц специального вида. Ниже мы изложим свойства операции над матрицами, позволяющие строить матрицы, задающие основные представления булевых функций.

Тензорное (кронекерово) произведение матриц

Определение 2.3. *Пусть $A = (a_{ij})$ и $B = (b_{st})$ — матрицы размеров $m \times m$ и $n \times n$ над полем K соответственно. Тензорным (кронекеровым) произведением матриц A и B называется матрица C размера $mn \times mn$ следующего вида:*

$$C = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1m}B \\ a_{21}B & a_{22}B & \dots & a_{2m}B \\ \dots & \dots & \dots & \dots \\ a_{m1}B & a_{m2}B & \dots & a_{mm}B \end{pmatrix}.$$

Обозначение: $C = A \otimes B$. Условимся также обозначать k-ю тензорную степень матрицы A через $A^{[k]}$.

Теорема 2.9 (Свойства тензорного произведения). *Пусть A, B, C — квадратные матрицы подходящих размеров над полем K. Справедливы соотношения:*

1) $A \otimes (B \otimes C) = (A \otimes B) \otimes C$;

2) $(A+B)\otimes C = A\otimes C + B\otimes C$; $A\otimes(B+C) = A\otimes B + A\otimes C$;

3) если $A, C \in K_{m,m}$, *а* $B, D \in K_{n,n}$, *то* $(A\otimes B)(C\otimes D) = AC \otimes BD$;

4) матрица $A\otimes B$ *обратима тогда и только тогда, когда обратимы матрицы* A *и* B. *При этом* $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$;

5) $(AB)^{[n]} = A^{[n]}B^{[n]}$.

□ Свойства 1 и 2 легко проверяются непосредственно, проверка предоставляется читателю. Докажем свойства 3–5. Вначале опишем более формально элементы матрицы, являющейся тензорным произведением. Пусть $A = (a_{r,i}) \in K_{m,m}$, $B = (b_{sj}) \in K_{n,n}$ и $Q = A \otimes B = (q_{kl})_{mn\times mn}$. Тогда из определения тензорного произведения следует, что

$$q_{(r-1)n+s,(i-1)n+j} = a_{ri}b_{sj}.$$

Для доказательства свойства 3 введем обозначения: $G = (g_{kl}) = A \otimes B$, $F = (f_{lt}) = C \otimes D$. Тогда

$$g_{kl} = a_{ru}b_{sv},\ f_{lt} = c_{ui}d_{vj},$$

где

$$k = (r-1)n + s,\ l = (u-1)n + v,\ t = (i-1)n + j$$

и

$$1 \le r, u, i \le m;\ 1 \le s, v, j \le n.$$

Пусть $G \cdot F = H = (h_{kt})$. Тогда

$$h_{kt} = \sum_{l=1}^{mn} g_{kl}f_{lt} = \sum_{l=1}^{mn} a_{ru}b_{sv}c_{ui}d_{vj} = \sum_{u=1}^{m}\sum_{v=1}^{n} a_{ru}b_{sv}c_{ui}d_{vj} =$$

$$= \sum_{u=1}^{m} a_{ru}c_{ui} \sum_{v=1}^{n} b_{sv}d_{vj} = x_{ri}y_{sj}.$$

Из последнего равенства следует требуемое равенство: $H = X \otimes Y$, где $X = AC$ и $Y = BD$.

Докажем свойство 4. Пусть матрицы A и B порядков m и n соответственно обратимы. Тогда по пункту 3 имеем

$$(A \otimes B)(A^{-1} \otimes B^{-1}) = AA^{-1} \otimes BB^{-1} = E_{mn},$$

где E_{mn} — единичная матрица порядка mn. Отсюда вытекает требуемое утверждение.

Пусть теперь матрица $A \otimes B$ невырождена. Поскольку по пункту 3 выполняется равенство $(A \otimes E_n)(E_m \otimes B) = A \otimes B$, то ранги матриц $A \otimes E_n$ и $E_m \otimes B$ равны mn и, следовательно, эти матрицы невырождены. Рассмотрим матрицу

$$E_m \otimes B = \begin{pmatrix} B & \mathbb{O} & \dots & \mathbb{O} \\ \mathbb{O} & B & \dots & \mathbb{O} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbb{O} & \mathbb{O} & \dots & B \end{pmatrix}.$$

Ее определитель, очевидно, равен $|B|^m$, откуда следует невырожденность матрицы B.

Рассмотрим матрицу

$$A \otimes E_n = \begin{pmatrix} a_{11} & & \mathbb{O} & a_{12} & & \mathbb{O} & & a_{1m} & & \mathbb{O} \\ & \ddots & & & \ddots & & \dots & & \ddots & \\ \mathbb{O} & & a_{11} & \mathbb{O} & & a_{12} & & \mathbb{O} & & a_{1m} \\ & \dots & & & \dots & & \dots & & \dots & \end{pmatrix}.$$

Нетрудно проверить, что данную матрицу подходящей перестановкой строк и столбцов можно привести к виду

$$\begin{pmatrix} A & \mathbb{O} & \dots & \mathbb{O} \\ \mathbb{O} & A & \dots & \mathbb{O} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbb{O} & \mathbb{O} & \dots & A \end{pmatrix}.$$

Следовательно, $|A \otimes E_n| = \pm|A|^n$, и так как $|A \otimes E_n| \neq 0$, то матрица A невырождена.

Применим метод полной математической индукции. Для $n = 1$ утверждение очевидно. Дальнейшее доказательство сводится к проверке равенств:

$$(AB)^{[n+1]} = (AB) \otimes (AB)^{[n]} = (AB) \otimes A^{[n]}B^{[n]} = A^{[n+1]}B^{[n+1]}.$$

Второе равенство следует из предположения индукции, а последнее — из пункта 3. □

Следствие 1. $A^{[n]} = (A \otimes E \otimes \ldots E)(E \otimes A \otimes \ldots E)\ldots(E \otimes \oplus E \otimes \ldots A)$.

□ Снова применим метод полной математической индукции. Для $n = 2$ утверждение следует из пункта 3 доказанной теоремы. Рассмотрим $A^{[n+1]}$:

$$A^{[n+1]} = A \otimes A^{[n]} = (A \cdot E) \otimes E^{[n]}A^{[n]} = \\ = (A \otimes E^{[n]})(E \otimes A^{[n]}).$$

Второй сомножитель распишем, применив предположение индукции:

$$E \otimes A^{[n]} = E \otimes ((A \otimes E \otimes ...E)(E \otimes A \otimes ...E)... \\ ...(E \otimes E \otimes ...A)) = (E \otimes A \otimes E \otimes ...E)(E \otimes A'),$$

где $A' = (E \otimes A \otimes \ldots E)\ldots(E \otimes E \otimes \ldots A)$, и утверждение доказано. □

Заметим, что справедливо и такое равенство:

$$A^{[n]} = (E \otimes \ldots E \otimes A)(E \otimes \ldots \otimes A \otimes E)\ldots(A \otimes E \otimes \ldots E).$$

Базисы, пространства псевдобулевых функций, задаваемые тензорным произведением матриц

Как отмечалось выше, с каждой невырожденной матрицей размера $2^n \times 2^n$ над полем K связан базис функционального пространства $K_2(n)$, составленный из функций, столбцы значений которых совпадают со столбцами данной матрицы. Оказывается, что в случае, когда матрица получена как тензорное произведение n матриц размера 2×2, можно легко выписать вид этих базисных функций.

Лемма 2.10. *Пусть матрица A размера $2^n \times 2^n$ над полем K имеет вид $A = B \otimes A'$, где $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $a, b, c, d \in K$, а A' — матрица размера $2^{n-1} \times 2^{n-1}$ над полем K, причем B и A' — невырожденные матрицы. Пусть столбцы матриц A и A' задают базисы пространств $K_2(n)$ и $K_2(n-1)$, функции из которых мы будем обозначать $\{g_\alpha\}_{\alpha \in \Omega_2^n}$ и $\{g_{\alpha'}\}_{\alpha' \in \Omega_2^{n-1}}$. Тогда для любого $\alpha' \in \Omega_2^{n-1}$ справедливы равенства:*

$$\begin{aligned} g_{(0,\alpha')}(x_1, \ldots, x_n) &= (a\overline{x}_1 + cx_1) g_{\alpha'}(x_2, \ldots, x_n), \\ g_{(1,\alpha')}(x_1, \ldots, x_n) &= (b\overline{x}_1 + dx_1) g_{\alpha'}(x_2, \ldots, x_n). \end{aligned} \tag{2.3}$$

□ Матрица A имеет вид

$$A = \begin{pmatrix} aA' & bA' \\ cA' & dA' \end{pmatrix}.$$

Таким образом, функцию $g_{(0,\alpha')}$ задает столбец матрицы A, верхняя половина которого соответствует столбцу α' матрицы A', умноженному на a, а нижняя — тому же столбцу, умноженному на c. Аналогично, функцию $g_{(1,\alpha')}$ задает столбец матрицы A, верхняя половина которого соответствует столбцу α' матрицы A', умноженному на b, а нижняя — тому же столбцу, умноженному на d. Отсюда и из 1.3 следуют формулы (2.3). □

Заметим, что равенства (2.3) объединяются в одно следующим образом:

$$\begin{aligned} g_{(u,\alpha')}(x_1, \ldots, x_n) = \\ = (\overline{u}(a\overline{x}_1 + cx_1) + u(b\overline{x}_1 + dx_1)) g_{\alpha'}(x_2, \ldots, x_n). \end{aligned} \tag{2.4}$$

Теорема 2.11. *Пусть матрица A является тензорным произведением матриц $B_i = \begin{pmatrix} a_i & b_i \\ c_i & d_i \end{pmatrix}$, $i = 1, \ldots, n$, т. е. $A = B_1 \otimes B_2 \otimes \ldots \otimes B_n$. Тогда базисная функция g_ω, соответствующая столбцу матрицы A, занумерованному набором $\omega = (u_1, \ldots, u_n)$, имеет вид*

$$g_\omega(x_1, \ldots, x_n) = \prod_{i=1}^{n} [\overline{u}_i(a_i\overline{x}_i + c_i x_i) + u_i(b_i\overline{x}_i + d_i x_i)].$$

□ Доказательство проводится методом математической индукции с использованием леммы 2.10. □

Рассмотрим теперь различные варианты выбора матриц B_i. В рассмотренных примерах эта матрица будет одинакова при различных i ($B_i = B$):

1. Поле K — любое. $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

По теореме 2.11 базисные функции имеют вид

$$g_\omega(x_1, \ldots, x_n) = \prod_{i=1}^{n} (\overline{u}_i \overline{x}_i + u_i x_i) = \prod_{i=1}^{n} x_i^{u_i} = x_1^{u_1} \ldots x_n^{u_n},$$

тогда по равенству (2.2):

$$f(x_1, \ldots, x_n) = \sum_{(u_1, \ldots, u_n) \in \Omega_2^n} x_1^{u_1} \ldots x_n^{u_n} \tilde{f}(u_1, \ldots, u_n),$$

и вектор $\tilde{f}^{\downarrow}$ совпадает с вектором $f^{\downarrow}$, что, впрочем, было очевидно изначально. Этот пример приведен исключительно в целях иллюстрации действия теоремы 2.11. Следующие примеры являются более содержательными.

2. K — поле действительных чисел. $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$.

Базисные функции имеют вид

$$g_\omega(x_1, ..., x_n) = \prod_{i=1}^{n} (\overline{u}_i (\overline{x}_i + x_i) + u_i x_i) =$$

$$= \prod_{i=1}^{n} (\overline{u}_i + u_i x_i) = \prod_{i:\, u_i = 1} x_i$$

и

$$f(x_1, \ldots, x_n) = \sum_{(u_1, \ldots, u_n) \in \Omega_2^n} \prod_{i:\, u_i = 1} x_i \tilde{f}(u_1, \ldots, u_n),$$

таким образом, вектор $\tilde{f}^{\downarrow}$ состоит из коэффициентов действительного многочлена функции f (см. теорему 1.5).

3. $K = GF(2)$. $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$.

Рассуждения аналогичны предыдущему случаю, только здесь вектор $\tilde{f}$ состоит из коэффициентов многочлена Жегалкина функции f.

4. K — поле комплексных чисел чисел. $B = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

Базисные функции имеют вид

$$g_\omega(x_1, ..., x_n) =$$
$$= \prod_{i=1}^{n} (\overline{u}_i(\overline{x}_i + x_i) + u_i(\overline{x}_i - x_i)) = \prod_{i=1}^{n} (\overline{u}_i + u_i(-1)^{x_i}) =$$
$$= \prod_{i=1}^{n} (-1)^{x_i u_i} = (-1)^{\sum_{i=1}^{n} x_i u_i} = (-1)^{<\omega, \mathbf{x}>}$$

и

$$f(x_1, \ldots, x_n) = \sum_{\omega \in \Omega_2^n} (-1)^{<\omega, \mathbf{x}>} \tilde{f}(u_1, \ldots, u_n),$$

и вектор $\tilde{f}^{\downarrow}$ состоит из коэффициентов Фурье функции f.

Быстрое преобразование Фурье

В случае, когда матрица A представляет собой тензорную степень матрицы 2×2, представление $\tilde{f}^{\downarrow}$ из равенства (2.1) можно получить не обычным умножением матрицы A^{-1} на вектор $f^{\downarrow}$, которое требует, как известно, $O(2^{2n})$ операций поля K, а существенно более эффективным способом.

Теорема 2.12. *Пусть B — невырожденная матрица размера 2×2 над полем K, а $A = B^{[n]}$. Тогда существует алгоритм вычисления вектора $\tilde{f}^{\downarrow}$ по данному вектору $f^{\downarrow}$ (см. равенство (2.1)), имеющий сложность $O(n2^n)$ операций поля K.*

□ Пусть $B^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Из свойств тензорного произведения матриц следует, что

$$A^{-1} = (B^{-1})^{[n]} = D_n \cdot D_{n-1} \cdots D_1,$$

где

$$D_i = \left(E_2^{[n-i]} \otimes \begin{pmatrix} a & b \\ c & d \end{pmatrix} \otimes E_2^{[i-1]} \right), \; E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Обозначим $f_0^{\downarrow} = f^{\downarrow}$ и для каждого $i \in \overline{1,n}$: $f_i^{\downarrow} = D_i \cdot f_{i-1}^{\downarrow}$. Тогда, очевидно, $f_n^{\downarrow} = \tilde{f}^{\downarrow}$.

Покажем, что каждое из умножений D_i на $f_{i-1}^{\downarrow}$, $i \in \overline{1,n}$ может быть выполнено за $O(2^n)$ операций поля K. Тогда общее количество операций, необходимое для вычисления $\tilde{f}^{\downarrow}$ по данному вектору $f^{\downarrow}$, будет составлять $O(n2^n)$ операций. Несложно видеть, что

$$D_i = \left(E_{2^{n-i}} \otimes \begin{pmatrix} aE_{2^{i-1}} & bE_{2^{i-1}} \\ cE_{2^{i-1}} & dE_{2^{i-1}} \end{pmatrix} \right) = \begin{pmatrix} \widehat{D}_i & & \mathbb{O} \\ & \ddots & \\ \mathbb{O} & & \widehat{D}_i \end{pmatrix},$$

где через $\widehat{D}_i$ мы обозначили матрицу $\begin{pmatrix} aE_{2^{i-1}} & bE_{2^{i-1}} \\ cE_{2^{i-1}} & dE_{2^{i-1}} \end{pmatrix}$.

Пусть теперь $X^{\downarrow}$ — произвольный вектор длины 2^n над полем K. Опишем алгоритм умножения D_i на $X^{\downarrow}$.

1. Разбиваем вектор $X^{\downarrow}$ на 2^{n-i} частей длины 2^i:

$$X^{\downarrow} = \begin{pmatrix} X_1^{\downarrow} \\ \vdots \\ X_{2^{n-i}}^{\downarrow} \end{pmatrix}.$$

Тогда

$$D_i X^{\downarrow} = \begin{pmatrix} \widehat{D}_i X_1^{\downarrow} \\ \vdots \\ \widehat{D}_i X_{2^{n-i}}^{\downarrow} \end{pmatrix}.$$

2. Каждый из подвекторов $X_j^{\downarrow}$ разбиваем на два подвектора равной длины: $X_j^{\downarrow} = \begin{pmatrix} X_{j0}^{\downarrow} \\ X_{j1}^{\downarrow} \end{pmatrix}$. Тогда

$$\widehat{D}_i X_j^{\downarrow} = \begin{pmatrix} aE_{2^{i-1}} & bE_{2^{i-1}} \\ cE_{2^{i-1}} & dE_{2^{i-1}} \end{pmatrix} \begin{pmatrix} X_{j0}^{\downarrow} \\ X_{j1}^{\downarrow} \end{pmatrix} = \\ = \begin{pmatrix} aE_{2^{i-1}} X_{j0}^{\downarrow} + bE_{2^{i-1}} X_{j1}^{\downarrow} \\ cE_{2^{i-1}} X_{j0}^{\downarrow} + dE_{2^{i-1}} X_{j1}^{\downarrow} \end{pmatrix} = \\ = \begin{pmatrix} aX_{j0}^{\downarrow} + bX_{j1}^{\downarrow} \\ cX_{j0}^{\downarrow} + dX_{j1}^{\downarrow} \end{pmatrix}.$$

Несложно видеть, что для вычисления последнего вектора требуется $4 \cdot 2^{i-1}$ умножений и $2 \cdot 2^{i-1}$ сложений в поле K. Таким образом, вычисление выражения

$$\tilde{f}^{\downarrow} = D_n \cdot D_{n-1} \cdots D_1 \cdot f^{\downarrow}$$

может быть произведено за $O(n2^n)$ операций поля K. □

Описанный в теореме алгоритм изначально был разработан для осуществления преобразования Фурье, поэтому носит название *быстрое преобразование Фурье (БПФ)*. Между тем понятно, что этот алгоритм может применяться для нахождения и других представлений булевых функций.

Пример. Требуется вычислить коэффициенты многочлена Жегалкина функции $f = (x_1 \to x_2) \to x_3$.

Как отмечалось выше, представлению многочленом Жегалкина соответствует матрица A, являющаяся n-й тензорной степенью матрицы $B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. Несложно видеть, что матрица B^{-1} совпадает с матрицей B. Выполним последовательность действий, описанную в теореме 2.12 и соответствующую матрице B^{-1} (табл. 2.1):

Таблица 2.1

$x_1x_2x_3$	$f = f_0$	f_1	f_2	$f_3 = \tilde{f}$
0 0 0	0	0	0	0
0 0 1	1	1	1	1
0 1 0	0	0	0	0
0 1 1	1	1	0	0
1 0 0	1	1	1	1
1 0 1	1	0	0	1
1 1 0	0	0	1	1
1 1 1	1	1	1	1

Таким образом, в многочлене Жегалкина функции f равными единице будут коэффициенты $a_3, a_1, a_{13}, a_{12}, a_{123}$, и сам многочлен будет иметь вид

$$x_1 \oplus x_3 \oplus x_1x_2 \oplus x_1x_3 \oplus x_1x_2x_3.$$

Глава 3

КЛАССИФИКАЦИЯ ДИСКРЕТНЫХ ФУНКЦИЙ С ПОМОЩЬЮ ГРУПП ПРЕОБРАЗОВАНИЙ

Мощность множества n-местных функций k-значной логики значительно возрастает с ростом n. Таким образом, задача описания свойств всех функций существенно усложняется. Вместе с тем, появляется большое количество функций, свойства которых во многом идентичны. Поэтому логически обоснованным является подход, связанный с объединением функций в классы по некоторым признакам и дальнейшим изучением свойств функций по классам. Принципы такой группировки могут быть различными. Ниже мы рассмотрим некоторые способы группировки и методы изучения свойств полученных классов функций.

3.1. ЭКВИВАЛЕНТНОСТЬ ФУНКЦИЙ. ГРУППЫ ИНЕРЦИИ

Условимся, в случае известного фиксированного n, обозначать функцию от n переменных следующим образом: $f(x_1, \ldots, x_n) = f(\mathbf{x})$.

Определение 3.1. *Пусть $f(\mathbf{x})$ и $h(\mathbf{x})$ функции из $F_k(n)$ и G — произвольная группа подстановок (биективных преобразований) множества Ω_k^n (т. е. G — подгруппа симметрической группы $S(\Omega_k^n)$). Говорят, что функция f эквивалентна функции h относительно группы G, если существует такая подстановка g из группы G, что для любого набора α из Ω_k^n справедливо равенство $f(\alpha) = h(g(\alpha))$. Обозначают: $f \overset{G}{\sim} h$.*

Утверждение 3.1. *Отношение* $\overset{G}{\sim}$ *является отношением эквивалентности на* $F_k(n)$.

□ Рефлексивность. $f \overset{G}{\sim} f$, поскольку группа G содержит тождественную подстановку.

Симметричность. Пусть $f \overset{G}{\sim} h$. Значит, существует g из G, такая что для любого набора α из Ω_k^n справедливо равенство $f(\alpha) = h(g(\alpha))$. Тогда, очевидно, $f(g^{-1}(\alpha)) = h(\alpha)$, и следовательно, $h \overset{G}{\sim} f$.

Транзитивность. Пусть $f \overset{G}{\sim} h$ и $h \overset{G}{\sim} p$. Тогда существуют g_1 и g_2 из G, такие что для любого α справедливы равенства $f(\alpha) = h(g_1(\alpha))$, $h(\alpha) = p(g_2(\alpha))$. Тогда $f(\alpha) = p((g_1 g_2)(\alpha))$, и следовательно, $f \overset{G}{\sim} p$. □

Таким образом, множество $F_k(n)$ разбивается на классы эквивалентности. Класс, содержащий функцию f, будем обозначать $[f]_G$ и называть *G-типом функции f*. Очевидно неравенство $1 \leq |[f]_G| \leq |G|$.

Определение 3.2. *Функция* $f(\mathbf{x}) \in F_k(n)$ *называется инвариантной относительно подстановки* g *из группы* $G < S(\Omega_k^n)$, *если* $f(g(\mathbf{x})) = f(\mathbf{x})$.

Функция $f(\mathbf{x}) \in F_k(n)$ *называется инвариантной относительно группы* $G < S(\Omega_k^n)$, *если она инвариантна относительно любой подстановки из* G.

Читателю предоставляется самостоятельно доказать

Утверждение 3.2. *Множество подстановок из группы* G, *относительно которых функция* f *инвариантна, является подгруппой в* G. □

Эта подгруппа называется *группой инерции функции* f *в группе* G и обозначается $I_G(f)$.

Теорема 3.3. *Если* $f \in F_k(n)$ *и* G *— подгруппа в группе* $S(\Omega_k^n)$, *то* $|[f]_G| = \frac{|G|}{|I_G(f)|}$.

□ Пусть g и g' — подстановки из G. Равенство $f(g(\mathbf{x})) = f(g'(\mathbf{x}))$ выполняется тогда и только тогда, когда $f(\mathbf{x}) =$

$= f((g^{-1}g')(\mathbf{x}))$, что равносильно условию $g^{-1}g' \in I_G(f)$. Следовательно, функции $f(g(\mathbf{x}))$ и $f(g'(\mathbf{x}))$ различны тогда и только тогда, когда g и g' принадлежат различным смежным классам группы G по подгруппе $I_G(f)$. Следовательно,

$$[f]_G = |G : I_G(f)| = \frac{|G|}{|I_G(f)|}. \quad \square$$

Таким образом, класс $[f]_G$ максимален по мощности, если f имеет тривиальную (состоящую только из единичной подстановки) группу инерции в G. Наоборот, класс $[f]_G$ состоит из единственной функции f, если $I_G(f) = G$. Обе эти граничные ситуации мы рассмотрим в следующих параграфах.

3.2. ФУНКЦИИ, ИНВАРИАНТНЫЕ ОТНОСИТЕЛЬНО ГРУППЫ

Напомним, что орбитой группы подстановок $G < S(\Omega_k^n)$, содержащей элемент α, называется множество

$$\Delta_\alpha = \{\beta \in \Omega_k^n | \exists g \in G : \beta = g(\alpha)\}.$$

Утверждение 3.4. *Функция f инвариантна относительно группы G тогда и только тогда, когда на элементах каждой орбиты она принимает постоянные значения (т. е. для любого $\beta \in \Delta_\alpha : f(\beta) = f(\alpha)$).*

$\square$ Пусть f инвариантна относительно G. Если β принадлежит орбите Δ_α, то найдется такая подстановка g из G, что $\beta = g(\alpha)$, и тогда $f(\alpha) = f(g(\alpha)) = f(\beta)$.

Предположим, что f не инвариантна относительно группы G. Тогда существует такая подстановка g, что найдется α из Ω_k^n, для которого $f(\alpha) \neq f(g(\alpha))$, и имеем противоречие с тем, что f принимает постоянные значения на Δ_α. $\square$

Следствие 1. *Если группа G транзитивна, то функция f инвариантна относительно G тогда и только тогда, когда f — константа.* $\square$

Следствие 2. *Число функций k-значной логики, инвариантных относительно данной группы G, равно $k^{\nu(G)}$, где $\nu(G)$ — число орбит группы G.* $\square$

В качестве группы G мы будем рассматривать следующие подгруппы симметрической группы $S(\Omega_k^n)$:
1. Группа S_n перестановок координат векторов $\alpha \in \Omega_k^n$. Эта группа состоит из подстановок вида

$$g_s(a_1, \ldots, a_n) = (a_{i_1}, \ldots, a_{i_n}),$$

где $s = (i_1, \ldots, i_n)$ — некоторая перестановка чисел $1, \ldots, n$.
2. Группа сдвигов Σ_n. Пусть $\alpha = (a_1, \ldots, a_n) \in \Omega_k^n$. Тогда

$$\Sigma_n = \{g_\alpha | g_\alpha(c_1, \ldots, c_n) = (c_1 + a_1, \ldots, c_n + a_n),\ \alpha \in \Omega_k^n\},$$

где суммирование ведется по модулю k.
3. Группа Джевонса $Q_n =< S_n, \Sigma_n >$ (группа, порожденная всеми подстановками из групп $S_{n_1}\Sigma_n$).
4. Полная линейная группа $GL(n,k)$ над кольцом $\mathbb{Z}_k$.

$$GL(n,k) = \{g_A | g_A(a_1, \ldots, a_n) = (a_1, \ldots, a_n)A,\ A \in (\mathbb{Z}_k)_{n,n}^*\},$$

где $(\mathbb{Z}_k)_{n,n}^*$ — группа обратимых $n \times n$-матриц размером $n \times n$ над кольцом $\mathbb{Z}_k$.

Отметим, что для решения некоторых прикладных задач рассматривается также полная линейная группа $GL(n,q)$ над конечным полем из q элементов (см., например, [25]).
5. Полная аффинная группа $AL(n,k) =< GL(n,k), \Sigma_n >$ (группа, порожденная всеми подстановками из групп $GL(n,k)$, Σ_n).

Несложно видеть, что группы Σ_n, Q_n и $AL(n,k)$ транзитивны и, следовательно, инвариантными относительно них являются только константы.

Рассмотрим группу S_n. Ее орбитами являются множества векторов, имеющих одинаковый состав координат (т. е. таких векторов, в которых одинаковое количество координат принимает значение 0, одинаковое количество — значение 1 и т. д.). Таким образом, количество орбит группы S_n равно числу сочетаний с повторениями из k элементов по n. Напомним, что сочетанием с повторением из k элементов по n называется неупорядоченный набор $a_1, \ldots, a_n$ из n необязательно различных элементов множества Ω_k. Число таких сочетаний с повторениями равно $\binom{k+n-1}{k-1} = \binom{k+n-1}{n}$ (см., например, [57]).

Инвариантные относительно S_n функции называются симметрическими. Число их, согласно следствию 2 из утверждения 3.4, равно $k^{\binom{k+n-1}{k-1}}$. В случае $k=2$ это число равно 2^{n+1}. Это те функции, которые принимают одинаковые значения на векторах равного веса.

Рассмотрим группу $GL(n,k)$.

Теорема 3.5. *Орбитами группы $GL(n,k)$ являются все множества*

$$M_d = \{(a_1 \ldots, a_n) \in \Omega_k^n \mid НОД(a_1, \ldots, a_n, k) = d\},$$

где d — делитель числа k, $d < k$.

□ Пусть $\alpha = (a_1, \ldots, a_n)$ и $\beta = (b_1, \ldots, b_n)$, причем $\alpha A = \beta$ для некоторой невырожденной матрицы A над кольцом $\mathbb{Z}_k$. Тогда все координаты вектора β суть линейные комбинации координат вектора α, следовательно любой общий делитель $a_1, \ldots, a_n$ является общим делителем $b_1, \ldots, b_n$. Поскольку матрица A обратима, то справедливо и равенство $\beta A^{-1} = \alpha$, значит, общие делители наборов α и β совпадают. Следовательно

$$\text{НОД}(a_1, \ldots, a_n, k) = \text{НОД}(b_1, \ldots, b_n, k).$$

Таким образом, если α и β лежат на одной орбите, то они принадлежат одному и тому же множеству M_d.

Пусть теперь $\text{НОД}(a_1, \ldots, a_n, k) = \text{НОД}(b_1, \ldots, b_n, k) = d$. Покажем, что найдется невырожденная матрица A со свойством $\alpha A = \beta$. В наших рассуждениях понадобится следующее утверждение, доказательство которого предоставляется читателю в качестве самостоятельного упражнения:

Утверждение 3.6. *Если НОД $(a_1, \ldots, a_n, k) = d$, то существуют такие $c_2, \ldots, c_n$, что НОД $(a_1 + c_2 a_2 + \ldots + c_n a_n, k) = d$.* □

Пользуясь этим утверждением, найдем такие $c_2, \ldots, c_n$, что $(a_1 + c_2 a_2 + \ldots + c_n a_n, k) = d$, и такие $c'_2, \ldots, c'_n$, что $(b_1 + c'_2 b_2 + \ldots + c'_n b_n, k) = d$. Обозначим

$$\alpha_1 = (a_1 + c_2 a_2 + \ldots + c_n a_n, a_2, \ldots, a_n) \quad \text{и}$$

$$\beta_1 = (b_1 + c'_2 b_2 + \ldots + c'_n b_n, b_2, \ldots, b_n),\ \alpha_1, \beta_1 \in \Omega_k^n.$$

Тогда $\alpha_1 = \alpha \cdot A_1$ и $\beta_1 = \beta \cdot B_1$, где

$$A_1 = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ c_2 & & & \\ \vdots & & E_{n-1} & \\ c_n & & & \end{pmatrix} \text{ и } B_1 = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ c'_2 & & & \\ \vdots & & E_{n-1} & \\ c'_n & & & \end{pmatrix},$$

а E_{n-1} — единичная матрица порядка $n-1$. Пусть $a = a_1 + c_2 a_2 + \ldots + c_n a_n$. Рассмотрим сравнение $ax \equiv d(\mod k)$. Поскольку $(a,k) = d$, то оно разрешимо. Покажем, что найдется такое его решение x_1, что $(x_1, k) = 1$. Действительно, множество различных по модулю k решений этого сравнения состоит из чисел $x_0, x_0 + k_1, \ldots, x_0 + k_1(d-1)$, где $k_1 = \frac{k}{d}$, а x_0 — решение сравнения $a'x \equiv 1(\mathrm{mod} k_1)$, $a' = \frac{a}{d}$. Поскольку, очевидно, $(x_0, k_1) = 1$, то $(x_0, k_1, k) = 1$. Следовательно, по утверждению 3.6, найдется такое t, что $(x_0 + k_1 t, k) = 1$. Остается положить $x_1 = x_0 + k_1 t$. Рассуждая аналогично, получим, что для $b = b_1 + c'_2 b_2 + \ldots + c'_n b_n$ найдется такое y_1, что $by_1 \equiv d(\mathrm{mod} k)$ и $(y_1, k) = 1$. Рассмотрим матрицы

$$A_2 = \begin{pmatrix} x_1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & E_{n-1} & \\ 0 & & & \end{pmatrix} \text{ и } B_2 = \begin{pmatrix} y_1 & 0 & \ldots & 0 \\ 0 & & & \\ \vdots & & E_{n-1} & \\ 0 & & & \end{pmatrix}.$$

Несложно видеть, что $\alpha_2 = \alpha_1 \cdot A_2 = (d, a_2, \ldots, a_n)$ и $\beta_2 = \beta_1 \cdot B_2 = (d, b_2, \ldots, b_n)$. Положим $f_i = \frac{b_i - a_i}{d}$, $i \in \overline{2,n}$ и рассмотрим

$$F = \begin{pmatrix} 1 & f_2 & \ldots & f_n \\ 0 & & & \\ \vdots & & E_{n-1} & \\ 0 & & & \end{pmatrix}.$$

Тогда

$$\alpha_2 \cdot F = (d, f_2 d + a_2, \ldots, f_n d + a_n) = \beta_2,$$

и следовательно $\alpha \cdot A_1A_2F = \beta \cdot B_1B_2$. Поскольку матрицы A_1, A_2, B_1, B_2, F обратимы над $\mathbb{Z}_k$ (их определители взаимно просты с k), то матрица $A = A_1A_2FB_2^{-1}B_1^{-1}$ обратима, и $\alpha A = \beta$. □

Следствие 1. *Число функций k-значной логики от n переменных, инвариантных относительно группы $GL(n,k)$ равно $k^{\nu(k)}$, где $\nu(k)$ — число делителей числа k.*

3.3. ФУНКЦИИ С ТРИВИАЛЬНОЙ ГРУППОЙ ИНЕРЦИИ В АФФИННОЙ ГРУППЕ

В этом параграфе будет показано, что при достаточно большом n почти все функции из $F_k(n)$ имеют тривиальную группу инерции в аффинной группе. Точнее, имеет место

Теорема 3.7. *Пусть $T_k(n)$ — множество всех функций k-значной логики с тривиальной группой инерции в группе $AL(n,k)$. Тогда*

$$\lim_{n\to\infty} \frac{|T_k(n)|}{|F_k(n)|} = 1.$$

□ 1. Оценим сверху число неподвижных точек нетождественного аффинного преобразования $g \in AL(n,k)$. Рассмотрим уравнение $g(\mathbf{x}) = \mathbf{x}$, где $g(\mathbf{x}) = \mathbf{x}A + \alpha$, $\alpha = (a_1, \ldots, a_n) \in \mathbb{Z}_k^n$, A — невырожденная $n \times n$ матрица над $\mathbb{Z}_k$. В матричной форме уравнение запишется в виде

$$\mathbf{x}(E - A) = \alpha. \tag{3.1}$$

Если $A = E$, то уравнение не имеет решений, так как $\alpha \neq 0$ (g – нетождественное преобразование). Пусть $C = (c_{ij})_{n\times n} = E - A \neq \mathbb{O}$. Без ограничения общности можем считать, что $c_{11} \neq 0$. Тогда уравнение

$$c_{11}x_1 + c_{21}x_2 + \ldots + c_{n1}x_n = a_1 \tag{3.2}$$

в $\mathbb{Z}_k$ является следствием уравнения (3.1). Если зафиксировать $x_2, \ldots, x_n$ произвольными значениями из $\mathbb{Z}_k$, то полученное уравнение будет иметь не более (c_{11}, k) решений, а так

как $c_{11} \neq 0$, то число решений при произвольной фиксации $x_2, \ldots, x_n$ не превосходит $k/2$. Таким образом, количество решений уравнения (3.2), а следовательно и (3.1), не превосходит $k^n/2$.

2. Поскольку количество неподвижных точек аффинной подстановки не превосходит $k^n/2$, то количество независимых циклов в ее разложении не может превосходить $3k^n/4$ ($k^n/2$ циклов длины 1 и $k^n/4$ циклов длины 2). Следовательно количество функций, инвариантных относительно фиксированной подстановки g, не превосходит $k^{3k^n/4}$, так как функция должна принимать одинаковые значения на элементах каждой орбиты.

3. Несложно видеть, что порядок группы $AL(n,k)$ не превышает k^{n^2+n}. Таким образом, используя предыдущие рассуждения, получаем, что число функций, инвариантных относительно хотя бы одной нетождественной подстановки из $AL(n,k)$, не превышает $k^{3k^n/4+n^2+n}$, и

$$\lim_{n\to\infty} \frac{|T_k(n)|}{|F_k(n)|} = \lim_{n\to\infty} (1 - k^{n^2+n-k^n/4}) = 1. \square$$

3.4. ПЕРЕЧИСЛЕНИЕ g-ТИПОВ. ЛЕММА БЕРНСАЙДА

Использование отношений эквивалентности относительно рассмотренных групп преобразований позволяет осуществить разбиение множества всех функций на классы (G-типы), состоящие из функций со сходными свойствами. Для осуществления полной классификации необходимо построить список представителей G-типов: $f^{(1)}, \ldots, f^{(l(G))}$, где $l(G)$ — общее число G-типов.

Основной метод решения этой задачи состоит в следующем. Вначале выбирается какой-либо способ построения последовательности $f_1, f_2, f_3 \ldots$ функций из $F_k(n)$. Как правило, сразу не удается добиться того, чтобы все функции лежали в различных G-типах, поэтому способ выработки последовательности может быть основан на различных принципах и выбран из соображений удобства (например, перебор функций по

различным значениям весов, начиная с малых, и др.). На первом шаге полагаем $f^{(1)} = f_1$ и подсчитываем порядок группы инерции $I_G(f^{(1)})$. Далее проверяем, эквивалентна ли функция f_2 функции $f^{(1)}$. Если нет, то полагаем $f^{(2)} = f_2$ и подсчитываем порядок группы инерции $I_G(f^{(2)})$, в противном случае переходим к функции f_3 и т. д. Таким образом, для каждой новой функции f_i из исходной последовательности проверяем, эквивалентна ли она одной из уже отобранных функций $f^{(1)}, \ldots, f^{(m)}$. Если да, переходим к f_{i+1}, если нет, полагаем $f^{(m+1)} = f_i$ и вычисляем порядок группы инерции f_i. Процедура завершается, если становится выполненным равенство

$$\sum_{j=1}^{m+1} \frac{|G|}{|I_G(f^{(j)})|} = k^{k^n}.$$

В этом случае, очевидно, $m+1 = l(G)$. Таким образом, для реализации указанной процедуры необходимо уметь вычислять порядки групп инерции и проверять эквивалентность функций относительно заданной группы.

Заметим, что изначальное знание параметра $l(G)$ упрощает изложенный метод, поскольку отпадает необходимость вычисления порядков групп инерции и процедура заканчивается, как только количество представителей различных классов достигнет $l(G)$. Задача поиска этого параметра носит название задачи перечисления классов эквивалентности (G-типов). Рассмотрим ее решение на примере булевых функций, используя в качестве группы G группу сдвигов. Проведем вначале некоторые полезные рассуждения достаточно общего характера.

Напомним, что при определении G-эквивалентности функций, мы считали, что G — группа подстановок множества Ω_k^n. Введем в рассмотрение группу $\widehat{G}$, действующую на множестве $F_k(n)$. Для этого каждой подстановке $g \in G < S(\Omega_k^n)$ сопоставим подстановку $\widehat{g} \in S(F_k(n))$, определяемую равенством $\widehat{g}(f(\mathbf{x})) = f(g(\mathbf{x}))$. Обозначим $\widehat{G} = \{\widehat{g} : g \in G\}$. Несложно видеть, что отображение $g \to \widehat{g}$ задает изоморфизм групп G и $\widehat{G}$. Покажем, как в новых терминах можно интерпретировать ранее введенные понятия. Пусть $f \in F_k(n)$. Тогда орбита

группы $\widehat{G}$ — это множество

$$\Delta(f) = \{h \in F_k(n) | \exists \widehat{g} \in \widehat{G} : \widehat{g}(f) = h\} = \\ = \{h \in F_k(n) | \exists g \in G : (f(g(\mathbf{x})) = h(\mathbf{x})\},$$

которое, очевидно, совпадает с G-типом $[f]_G$. Стабилизатором элемента f является подгруппа

$$\widehat{G}_f = \{\widehat{g} \in \widehat{G} | \widehat{g}(f) = f\} = \{\widehat{g} \in \widehat{G} | f(g(\mathbf{x})) = f(\mathbf{x})\} = \widehat{I_G(f)},$$

где $I_G(f)$ — группа инерции функции f в группе G.

Ниже мы докажем утверждение, в теории групп подстановок известное как лемма Бернсайда. Формулировку и доказательство приведем в терминах теории дискретных функций.

Обозначим $\pi(\widehat{g})$ число единичных циклов подстановки $\widehat{g}$, т. е.

$$\pi(\widehat{g}) = |\{f \in F_k(n) | \widehat{g}(f) = f\}|.$$

Теорема 3.8. *Пусть* $\Delta_1, \ldots, \Delta_m$ *— орбиты группы* $\widehat{G}$. *Тогда*

$$m = \frac{1}{|\widehat{G}|} \sum_{\widehat{g} \in \widehat{G}} \pi(\widehat{g}).$$

□ Введем обозначение:

$$\delta_{f,\widehat{g}} = \begin{cases} 1 & , \quad \widehat{g}(f) = f \\ 0 & , \quad \widehat{g}(f) \neq f \end{cases}.$$

Рассмотрим сумму из правой части равенства:

$$\sum_{\widehat{g} \in \widehat{G}} \pi(\widehat{g}) = \sum_{\widehat{g} \in \widehat{G}} \sum_{f \in F_k(n)} \delta_{f,\widehat{g}} = \\ = \sum_{f \in F_k(n)} \sum_{\widehat{g} \in \widehat{G}} \delta_{f,\widehat{g}} = \sum_{f \in F_k(n)} |\widehat{G}_f| = \\ = \sum_{f \in F_k(n)} |I_G(f)| = \sum_{f \in F_k(n)} \frac{|G|}{|[f]_G|}.$$

Слагаемые последней суммы сгруппируем так, чтобы в одну группу попали слагаемые, соответствующие функциям из одного G-типа. Поскольку число G-типов равно числу орбит группы $\widehat{G}$, то количество представителей различных G-типов будет равно m. Пусть $f_1, \ldots, f_m$ — такие представители. Тогда

$$\sum_{f \in F_k(n)} \frac{|G|}{|[f]_G|} = \sum_{f \in [f_1]_G} \frac{|G|}{|[f_1]_G|} + \ldots + \sum_{f \in [f_m]_G} \frac{|G|}{|[f_m]_G|} =$$

$$= |G| + \ldots + |G| = m|G|,$$

что и требовалось. □

В качестве примера найдем число G-типов для $G = \Sigma_n$ и $k = 2$. Согласно теореме 3.8 это число равно

$$\frac{1}{|\Sigma_n|} \sum_{\widehat{g} \in \widehat{\Sigma_n}} \pi(\widehat{g}).$$

Число $\pi(\widehat{g})$ единичных циклов подстановки $\widehat{g}$ равно, очевидно, количеству функций, инвариантных относительно подстановки $g \in \Sigma_n$. Аналогично 3.4 функция инвариантна относительно подстановки g тогда и только тогда, когда на элементах каждого ее независимого цикла она принимает одинаковые значения. Несложно видеть, что каждая нетождественная подстановка из Σ_n имеет 2^{n-1} независимых циклов длины 2, следовательно в этом случае $\pi(\widehat{g}) = 2^{2^{n-1}}$. Относительно же тождественной подстановки инвариантны все булевы функции. Таким образом, число G-типов группы G для случая двоичных функций составляет

$$\frac{(2^n - 1)2^{2^{n-1}} + 2^{2^n}}{2^n}.$$

Для остальных рассматриваемых групп число G-типов определяется аналогичными методами, однако соответствующие результаты имеют гораздо более сложный вид. В частности, для группы $AL(n, 2)$ имеем

для $n = 1$ — 3 класса;

для $n = 2$ — 5 классов;

для $n = 3$ — 10 классов;
для $n = 4$ — 32 класса;
для $n = 5$ — 382 класса;
для $n = 6$ — около 16 млн классов.

3.5. ИНВАРИАНТЫ. НАХОЖДЕНИЕ ГРУПП ИНЕРЦИИ И ПРОВЕРКА ЭКВИВАЛЕНТНОСТИ

Как было отмечено выше, построение групп инерции и проверка эквивалентности функции — задачи, имеющие большое значение при проведении классификации дискретных функций. Для решения этих задач используются различные методы. Один из таких методов основан на теории инвариантов. Ниже мы приведем основные определения, необходимые при применении такого подхода.

Пусть G — группа подстановок, действующая на множестве M. Пусть φ — отображение множества M в некоторое множество P (в нашем случае это будет, как правило, числовое поле).

Определение 3.3. *Отображение φ называется инвариантом группы G, если для любой подстановки $g \in G$ и произвольного $m \in M$ справедливо равенство $\varphi(g(m)) = \varphi(m)$.*

Это определение, очевидно, является обобщением определения 3.2 дискретной функции, инвариантной относительно группы. Несложно видеть также (см. утверждение 3.4), что инвариант принимает одинаковые значения на орбитах группы G.

Инвариант φ называется *полным*, если из равенства $\varphi(m_1) = \varphi(m_2)$ следует, что элементы m_1 и m_2 принадлежат одной орбите группы G.

На практике, как правило, используют не один полный инвариант, построение которого весьма затруднительно, а полную систему инвариантов.

Определение 3.4. *Система инвариантов $\varphi_1, \ldots, \varphi_k$ называется полной системой инвариантов группы G, если из того, что одновременно выполняются равенства $\varphi_1(m_1) =$*

$= \varphi_1(m_2), \ldots, \varphi_k(m_1) = \varphi_k(m_2)$, *следует, что элементы* m_1 *и* m_2 *принадлежат одной орбите группы* G.

Принимая во внимание тот факт, что для группы $\widehat{G}$, действующей на множестве $F_k(n)$, орбитами являются G-типы, нетрудно сформулировать правило проверки факта эквивалентности функций f_1 и f_2 из $F_k(n)$.

Пусть φ — инвариант группы $\widehat{G}$, действующей на $F_k(n)$. Если φ — полный инвариант ($\varphi_1, \ldots, \varphi_k$ — полная система инвариантов), то равенство $\varphi(f_1) = \varphi(f_2)$ (одновременное выполнение равенств $\varphi_1(f_1) = \varphi_1(f_2), \ldots \varphi_k(f_1) = \varphi_k(f_2)$) равносильно тому, что функции f_1 и f_2 эквивалентны относительно группы G. Если же φ не является полным инвариантом, равенство $\varphi(f_1) = \varphi(f_2)$ является необходимым условием эквивалентности, т. е. в случае его невыполнения функции не эквивалентны. В случае, когда равенство справедливо, вопрос остается открытым и требуется проводить дополнительное исследование (например, проверять равенство для других инвариантов).

Перечислим некоторые инварианты для различных групп преобразований при $k = 2$ (читателю предоставляется самостоятельно проверить выполнимость определения инварианта в каждом случае):

1) для группы $AGL(n, 2)$ — вес, степень нелинейности функции;

2) для подгрупп группы Q_n — число существенных переменных функции, число простых импликант;

3) для группы S_n — вес функции, число одночленов в многочлене Жегалкина.

Рассмотрим пример. Пусть $f_1(x_1, x_2, x_3) = x_1 \oplus x_2$, $f_2(x_1, x_2, x_3) = x_1 \oplus x_2 \oplus x_3$. Поскольку число существенных переменных этих функций различно, то они не эквивалентны относительно любой подгруппы группы Q_n. Вместе с тем эти функции эквивалентны относительно группы $GL(n, 2)$, поскольку выполняется равенство $f_2(x_1, x_2, x_3) = f_1((x_1, x_2, x_3) \cdot A)$, где A — матрица вида

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Для построения группы инерции заданной функции также применяются методы теории инвариантов. Однако в этом случае задача в некотором смысле является обратной: группа, которая стабилизирует функцию, неизвестна и инварианты применяются для того, чтобы отсеять большое множество подстановок, заведомо не входящих в группу инерции. Затем, если это требуется, производится перебор оставшихся подстановок и проверка их принадлежности искомой группе.

В частности, в случае $k = 2$ используют следующие инварианты:

1) для групп S_n, Σ_n, Q_n, действие которых рассматривается на множестве символов $\{x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n\}$:

— частота встречаемости переменной x_i в многочлене Жегалкина (или в совокупности членов равной степени);

– частота встречаемости x_i и $\bar{x}_i$ в СДНФ (сокращенной ДНФ);

— вес, степень нелинейности подфункций $f_i^{(0)}$ и $f_i^{(1)}$;

— частота встречаемости пар переменных в многочлене Жегалкина и др.;

2) для групп $GL(n, 2)$ и $AGL(n, 2)$, действие которых рассматривается на множестве векторов $\alpha \in \Omega_2^n$:

— степень нелинейности и вес функций $\Delta_\alpha f(\mathbf{x}) = f(\mathbf{x} \oplus \alpha) \oplus f(\mathbf{x})$.

Рассмотрим пример. Найдем группу инерции $I_{S_n}(f)$, где

$$f = 1 \oplus x_1 \oplus x_3 \oplus x_5 \oplus x_6 \oplus x_1x_2 \oplus x_3x_4 \oplus x_3x_5 \oplus \\ \oplus x_4x_5 \oplus x_2x_6 \oplus x_1x_2x_4 \oplus x_2x_3x_5 \oplus x_2x_4x_6 \oplus x_1x_3x_5x_6.$$

Обозначим

$$f_1 = x_1 \oplus x_3 \oplus x_5 \oplus x_6,$$

$$f_2 = x_1x_2 \oplus x_3x_4 \oplus x_3x_5 \oplus x_4x_5 \oplus x_2x_6,$$

$$f_3 = x_1x_2x_4 \oplus x_2x_3x_5 \oplus x_2x_4x_6, \quad f_4 = x_1x_3x_5x_6$$

и составим таблицу (табл. 2.1) количеств вхождений переменных в функции f_1, f_2, f_3, f_4:

Таблица 2.1

	x_1	x_2	x_3	x_4	x_5	x_6
f_1	1	0	1	0	1	1
f_2	1	2	2	2	2	1
f_3	1	3	1	2	1	1
f_4	1	0	1	0	1	1

Отметим, что в этом случае инвариантами группы $I_{S_n}(f)$ являются столбцы этой матрицы. Следовательно подстановка, лежащая в группе инерции, может переставлять только те переменные, для которых столбцы приведенной матрицы совпадают. Одинаковые столбцы соответствуют парам переменным x_1, x_6 и x_3, x_5. Непосредственная проверка позволяет убедиться, что указанные перестановки не меняют многочлен Жегалкина данной функции. Следовательно, группа инерции состоит из 4 подстановок: тождественной, транспозиций $(3,5)$ и $(1,6)$ и их произведения.

В качестве иллюстрации применения методов построения групп инерции докажем теорему, представляющую, кроме того, самостоятельный теоретический интерес:

Теорема 3.9. *Для любого натурального $m \geq 2$ и произвольной группы G порядка m существует натуральное n и булева функция f от n переменных, группа инерции которой в группе S_n перестановок координат изоморфна G.*

□ Если $m = 2$, то, поскольку группа из двух элементов единственна с точностью до изоморфизма, группа G изоморфна S_2. В качестве функции f в данном случае можно выбрать, например, $x_1 x_2$ ($n = m = 2$).

Пусть $m > 2$. Занумеруем элементы группы G произвольным образом, начиная с единицы:

$$G = \{g_1 = e, g_2, \ldots, g_m\}.$$

Выберем множество переменных, от которых будет зависеть искомая функция, следующим образом. Пусть $Y = \{y_i | i \in \overline{1,m}\}$ и $X = \{x_{ij} | i,j \in \overline{1,m}\}$. Построим булеву функцию, существенно зависящую от переменных из множества $X \bigcup Y$, т. е. от $n = m^2 + m$ переменных. Для этого определим вначале на множестве $\overline{1,m}$ операцию $*$ следующим образом:

$i * j = t$, где t определяется соотношением: $g_j^{-1} \cdot g_i = g_t$.

Заметим, что $i * 1 = i$ и для фиксированного $i \in \overline{1,m}$ выполняется $\{i * j | j \in \overline{1,m}\} = \overline{1,m}$.

Рассмотрим теперь булеву функцию f, заданную многочленом Жегалкина следующего вида:

$$\sum_{i=1}^{m}(y_i x_{i1} x_{i2} + x_{i1} x_{i2} x_{i3} + \ldots + x_{i,m-2} x_{i,m-1} x_{im}) + \\ + \sum_{i=1}^{m}(y_{i*2} x_{i2} + y_{i*3} x_{i3} + \ldots + y_{i*m} x_{im}).$$

Покажем, что группа инерции $I_{S_n}(f)$ изоморфна G. Исходя из определения f и свойств операции $*$, составим таблицу (см. табл. 2.2) частот вхождения различных переменных из множества $X \bigcup Y$ в совокупность членов 2-й и 3-й степени.

Таблица 2.2

	степень 3	степень 2
y_i	1	$m-1$
x_{i1}	2	0
x_{i2}	3	1
$\vdots$	$\vdots$	$\vdots$
$x_{i,m-2}$	3	1
$x_{i,m-1}$	2	1
x_{im}	1	1

Используя подход теории инвариантов и то, что многочлен Жегалкина определяет булеву функцию однозначно, несложно

заметить, что если подстановка σ лежит в группе инерции, то $\sigma(Y) \subset Y$, поскольку переменные из множества Y могут переходить только друг в друга.

В дальнейшем доказательство теоремы разобьем на ряд утверждений:

1. *Пусть подстановка σ лежит в $I_{S_n}(f)$ и для некоторого y_i выполняется $\sigma(y_i) = y_i$. Тогда σ — тождественная подстановка.*

Поскольку y_i входит в единственный одночлен третьей степени $y_i x_{i1} x_{i2}$, то $\sigma(\{x_{i1}, x_{i2}\}) = \{x_{i1}, x_{i2}\}$. Из таблицы частот вхождений следует, что x_{i1} не может переходить в x_{i2}, таким образом, $\sigma(x_{i1}) = x_{i1}$ и $\sigma(x_{i2}) = x_{i2}$. Отсюда следует, что $\sigma(\{x_{i1}, x_{i2}, x_{i3}\}) = \{x_{i1}, x_{i2}, x_{i3}\}$, значит, $\sigma(x_{i3}) = x_{i3}$ и т. д. Таким образом, $\sigma(x_{ik}) = x_{ik}$ для всех $k \in \overline{1,m}$.

Теперь покажем, что для любого $s \in \overline{1,m}$, $s \neq i$ также выполняется $\sigma(y_s) = y_s$. Поскольку $\{i * j | j \in \overline{1,m}\} = \overline{1,m}$ и $i * 1 = i$, то существует такое $j \in \overline{2,m}$, что $i * j = s$. Следовательно в множество членов второй степени входит одночлен $y_s x_{ij}$. Так как $\sigma(x_{ij}) = x_{ij}$ и x_{ij} входит только в этот одночлен второй степени, то $\sigma(y_s) = y_s$. Далее, рассуждая аналогично, получаем, что $\sigma(x_{sk}) = x_{sk}$ для всех $k \in \overline{1,m}$. Следовательно, σ — тождественная подстановка.

2. *Для всех $i, j \in \overline{1,m}$ найдется не более одной подстановки из $I_{S_n}(f)$, такой что $\sigma(y_i) = y_j$*

Действительно, предположим, что найдутся две подстановки σ_1 и σ_2, что $\sigma_1(y_i) = y_j$ и $\sigma_2(y_i) = y_j$. Тогда $\sigma_1 \cdot \sigma_2^{-1}(y_i) = \sigma_2^{-1}(y_j) = y_i$, следовательно, по доказанному выше, $\sigma_1 \cdot \sigma_2^{-1}$ — тождественная подстановка и $\sigma_1 = \sigma_2$.

Заметим, что из доказанного вытекает неравенство

$$|I_G(f)| \leq m.$$

3. *Пусть φ — отображение из G в $I_{S_n}(f)$, $\varphi(g_k) = \sigma_k$, причем $\sigma_k(y_i) = y_j$, $\sigma_k(x_{is}) = x_{js}$, $s \in \overline{1,m}$, где j определяется из соотношения $g_i g_k = g_j$. Тогда φ — биекция.*

Покажем, что $\varphi(g_k) = \sigma_k$ лежит в группе $I_{S_n}(f)$. Из определения следует, что подстановка σ_k переводит каждый одночлен степени 3 многочлена Жегалкина функции f в одночлен

степени 3 этого же многочлена. Покажем, что σ_k переводит любой одночлен $y_{i*j}x_{ij}$ степени 2 также в одночлен степени 2 того же многочлена.

Действительно, пусть $\sigma_k(y_{i*j}) = y_r$, $\sigma_k(x_{ij}) = x_{sj}$. Тогда $g_r = g_{i*j}g_k = g_j^{-1}g_ig_k$, $g_s = g_ig_k$. Следовательно, $g_{s*j} = g_j^{-1}g_s = g_j^{-1}g_ig_k = g_r$ и $r = s * j$. Таким образом, одночлен y_rx_{sj} входит в многочлен Жегалкина функции f. Поэтому φ переводит элементы из G в элементы группы инерции.

Очевидно, что для различных k и k' подстановки σ_k и $\sigma_{k'}$ различны. Таким образом, из мощностных соображений вытекает биективность отображения φ.

4. *φ — изоморфизм групп G и $I_{S_n}(f)$.*

Покажем, что из равенства $g_sg_r = g_t$ вытекает равенство $\sigma_s\sigma_r = \sigma_t$. Из пункта 2 следует, что достаточно проверить соотношение $\sigma_s\sigma_r(y_i) = \sigma_t(y_i)$, $i \in \overline{1,m}$.

Пусть $\sigma_s(y_i) = y_j$, $\sigma_r(y_j) = y_k$ и $\sigma_t(y_i) = y_l$. Тогда $g_j = g_ig_s$, $g_jg_r = g_k$, т. е. $g_k = g_ig_sg_r$. С другой стороны, $g_ig_t = g_l$, следовательно $g_l = g_ig_sg_r = g_k$. Таким образом, $\sigma_s\sigma_r(y_i) = \sigma_t(y_i)$ и $\sigma_s\sigma_r = \sigma_t$. Теорема доказана. □

Глава 4

ВЕРОЯТНОСТНЫЕ И КОМБИНАТОРНЫЕ СВОЙСТВА И ПАРАМЕТРЫ ДИСКРЕТНЫХ ФУНКЦИЙ

В этой главе для изучения свойств булевых функций будет применен вероятностный подход, суть которого сводится к следующему. Пусть $x_1, \ldots, x_n$ — случайные величины, распределенные на множестве Ω_2. Если f — некоторая булева функция, то значение $f(x_1, \ldots, x_n)$ также представляет собой случайную величину, некоторые свойства и характеристики которой будут изучены ниже. Кроме того, будут описаны классы функций, представляющие практический интерес, для которых вводимые параметры будут иметь определенные значения.

4.1. ВЕРОЯТНОСТНАЯ ФУНКЦИЯ БУЛЕВОЙ ФУНКЦИИ

Будем считать, что x_i — независимые случайные величины с распределением

$$P(x_i = 1) = p_i > 0,\ P(x_i = 0) = 1 - p_i.$$

Определение 4.1. *Вероятностной функцией булевой функции $f \in F_2(n)$ называют отображение $F_f : [0,1]^n \to [0,1]$, определяемое равенством*

$$F_f(p_1, \ldots, p_n) = P(f(x_1, \ldots, x_n) = 1).$$

Понятие вероятностной функции было введено К. Шенноном [45] и широко используется в теории надежности (см., например, [56]).

Утверждение 4.1. *Справедливо соотношение*

$$F_f(p_1,\dots,p_n)=\sum_{(a_1,\dots,a_n):f(a_1,\dots,a_n)=1}\prod_{i=1}^{n}p_i^{a_i}(1-p_i)^{\overline{a}_i},$$

где $a_i\in\Omega_2$. $\square$

Следующая теорема демонстрирует применение понятия действительного многочлена булевой функции:

Теорема 4.2. *Пусть* D_f *— действительный многочлен булевой функции* f. *Тогда* $F_f(p_1,\dots,p_n)=D_f(p_1,\dots,p_n)$.

$\square$ Доказательство проведем методом математической индукции по n. В случае $n=1$ имеем 4 функции 0, 1, x, $\overline{x}$. Для них равенство легко проверить непосредственно.

Пусть утверждение теоремы справедливо для всех $n<k$. Пусть $k=n$. Для функции f справедливо разложение (см. следствие 1 теоремы 1.2):

$$f(x_1,\dots,x_k)=\overline{x}_1f(0,x_2,\dots,x_k)\vee x_1f(1,x_2,\dots,x_k)=$$
$$=\overline{x}_1f(0,x_2,\dots,x_k)\oplus x_1f(1,x_2,\dots,x_k).$$

Тогда в силу единственности действительного многочлена функции f справедливо равенство

$$D_f(x_1,\dots,x_k)=x_1D_{f_1}(x_2,\dots,x_k)+(1-x_1)D_{f_0}(x_2,\dots,x_k),$$

где $f_0=f(0,x_2,\dots,x_k)$, $f_1=f(1,x_2,\dots,x_k)$.

По предположению индукции,

$$F_{f_0}(p_2,\dots,p_k)=D_{f_0}(p_2,\dots,p_k)\text{ и}$$
$$F_{f_1}(p_2,\dots,p_k)=D_{f_1}(p_2,\dots,p_k).$$

Тогда

$$P(f(\mathbf{x})=1)=P(x_1=1)P(f(\mathbf{x})/x_1=1)+$$
$$+P(x_1=0)P(f(\mathbf{x})/x_1=0)=p_1F_{f_1}(p_2,\dots,p_k)+$$
$$+(1-p_1)F_{f_0}(p_2,\dots,p_k)=p_1D_{f_1}(p_2,\dots,p_k)+$$
$$+(1-p_1)D_{f_0}(p_2,\dots,p_k)=D_f(p_1,\dots,p_k).\ \square$$

Утверждение 4.3. *Если* $p_1 = \ldots = p_n = p$, *то* $F_f(p, \ldots, p) = \sum_{i=0}^{n} a_i p^i (1-p)^{n-i}$, *где* $a_i = |\{\alpha \in \Omega_2^n | f(\alpha) = 1 \text{ и } \|\alpha\| = i\}|$.

В случае $p_1 = \ldots = p_n = p$ *будем вместо* $F_f(p, \ldots, p)$ *писать просто* $F_f(p)$:

$$\square \quad F_f(p) = \sum_{\alpha: f(\alpha)=1} P(\mathbf{x} = \alpha) =$$

$$= \sum_{i=0}^{n} \sum_{\|\alpha\|=i: f(\alpha)=1} P(\mathbf{x} = \alpha) =$$

$$= \sum_{i=0}^{n} a_i p^i (1-p)^{n-i}. \quad \square$$

Обозначим $\delta = p - 1/2$. Тогда

$$F_f(p) = F_f(1/2 + \delta) = \sum_{i=0}^{n} a_i (1/2 + \delta)^i (1/2 - \delta)^{n-i} =$$

$$= \frac{\sum\limits_{i=1}^{n} a_i}{2^n} + d_1\delta + d_2\delta^2 + \ldots + d_n\delta^n =$$

$$= \frac{\|f\|}{2^n} + d_1\delta + d_2\delta^2 + \ldots + d_n\delta^n \quad (4.1)$$

для некоторых d_i. При малых значениях δ, таким образом, вероятность того, что значение f будет равно единице, близка к $\frac{1}{2^n}\|f\|$, т. е. в случае $\|f\| = 2^{n-1}$ распределение близко к равномерному (этим оправдано название «равновероятная функция»). Если при этом несколько коэффициентов $d_1, d_2, \ldots$ из уравнения (4.1) равны нулю, то распределение $f(x_1, \ldots, x_n)$ существенно «ближе» к равномерному, чем распределение x_i.

Опишем свойства вероятностной функции для булевых функций из некоторых предполных классов.

I. Самодвойственные функции

Утверждение 4.4. *Пусть* $f(x_1, \ldots, x_n)$ — *самодвойственная булева функция. Тогда* $F_f(p_1, \ldots, p_n) = 1 - F_f(1 - p_1, \ldots, 1 - p_n)$.

□ Заметим, что

$$P(f(\mathbf{x})=1)=1-P(f(\mathbf{x})=0)=1-P(\overline{f}(\mathbf{x})=1)$$

и

$$P(\overline{f}(\mathbf{x})=1)=\sum_{\alpha:\overline{f}(\alpha)=1} p_i^{a_i}(1-p_i)^{a_i}.$$

Заменяя суммирование по наборам $\alpha=(a_1,\ldots,a_n)$ суммированием по наборам $\beta=(b_1,\ldots,b_n)$, $\beta=\overline{\alpha}$, получим

$$\sum_{\beta:\overline{f}(\overline{\beta})=1} p^{1-b_i}(1-p_i)^{b_i}=$$

$$=\sum_{\beta:f(\beta)=1} p^{1-b_i}(1-p_i)^{b_i}=F_f(1-p_1,\ldots,1-p_n).\ \square$$

Заметим, что при $p_1=p_2=\ldots=p_n=p$ имеем

$$1-F_f(1-p)=F_f(p),$$

так что график вероятностной функции симметричен относительно точки $(1/2,1/2)$.

II. Аффинные функции

Докажем вначале общее утверждение:

Теорема 4.5. *Если* $f(x_1,\ldots,x_n)=f_1(x_1,\ldots,x_k)\oplus f_2(x_{k+1},\ldots,x_n)$, $1\le k<n$, *то*

$$\frac{1}{2}-F_f(p_1,\ldots,p_n)=$$

$$=2(\frac{1}{2}-F_{f_1}(p_1,\ldots,p_k))(\frac{1}{2}-F_{f_2}(p_{k+1},\ldots,p_n)).$$

□ Из очевидного равенства $f_1\oplus f_2=f_1+f_2-2f_1f_2$ (+ и − из поля $\mathbb{R}$) следует, что при $x_i\in\Omega_2$ имеет место равенство значений действительных многочленов:

$$D_f(x_1\ldots,x_n)=D_{f_1}(x_1,\ldots,x_k)+D_{f_2}(x_{k+1},\ldots,x_n)-$$
$$-2D_{f_1}(x_1,\ldots,x_k)D_{f_2}(x_{k+1},\ldots,x_n). \quad (4.2)$$

В правой части этого равенства находится, очевидно, многочлен, в каждый одночлен которого каждая переменная входит в степени не выше первой. Тогда в силу однозначности представления функции действительным многочленом выражения в левой и правой части (возможно, после приведения подобных слагаемых) будут совпадать. Следовательно,

$$\frac{1}{2} - F_f(p_1, \ldots, p_n) = \\ = 2(\frac{1}{2} - F_{f_1}(p_1, \ldots, p_k))(\frac{1}{2} - F_{f_2}(p_{k+1}, \ldots, p_n)). \square$$

Следствие 1. *Пусть* $f(x_1, \ldots, x_n) = x_1 \oplus \ldots \oplus x_k \oplus a_0$ *и* $p_i = 1/2 + \delta_i$, $i \in \overline{1,k}$. *Тогда*

$$F_f(p_1 \ldots, p_n) = 1/2 + (-1)^{a_0}(-2)^{k-1}\delta_1 \ldots \delta_k.$$

$\square$ Докажем следствие индукцией по k. В случае $k = 1$ имеем

$$F_f(p_1) = \begin{cases} p_1, & a_0 = 0 \\ 1 - p_1, & a_0 = 1 \end{cases} = \\ = \begin{cases} 1/2 + \delta_1, & a_0 = 0 \\ 1/2 - \delta_1, & a_0 = 1 \end{cases} = 1/2 + (-1)^{a_0}\delta_1,$$

и утверждение справедливо.

Шаг индукции. Функцию $f(x_1, \ldots, x_n) = x_1 \oplus \ldots x_k \oplus a_0$ представим в виде суммы функций x_1 и $x_2 \oplus \ldots x_k \oplus a_0$. По предположению индукции имеем

$$F_f(p_1, \ldots, p_n) = p_1 + 1/2 + (-1)^{a_0}(-2)^{k-2}\delta_2 \ldots \delta_k - p_1 - \\ - 2p_1(-1)^{a_0}(-2)^{k-2}\delta_2 \ldots \delta_k = \\ = 1/2 - (-1)^{a_0}\delta_2 \ldots \delta_k(-(-2)^{k-2} + 2(1/2 + \delta_1)(-2^{k-2})) = \\ = 1/2 + (-1)^{a_0}(-2)^{k-1}\delta_1 \ldots \delta_k,$$

и следствие доказано. $\square$

III. Монотонные функции

Лемма 4.6. *Для любой монотонной функции* $f_1(x_1, \ldots, x_n)$ *и для любых* $p_i \in (0,1)$, $i \in \overline{2,n}$ *справедливо неравенство*

$$F_{f_0}(p_2, \ldots, p_n) \le F_{f_1}(p_2, \ldots, p_n), \tag{4.3}$$

где $f_a(x_2, \ldots, x_n) = f(a, x_2, \ldots, x_n)$, $a \in \{0,1\}$. *Неравенство строгое тогда и только тогда, когда* x_i *— существенная переменная.*

□ Из определения монотонной функции следует, что для любого набора $(a_2, \ldots, a_n) \in \Omega_2^n$ выполняется

$$f_0(a_2, \ldots, a_n) \le f_1(a_2, \ldots, a_n).$$

При этом

$$F_{f_i}(p_2, \ldots, p_n) = \sum_{\substack{(a_2, \ldots, a_n) \in \Omega_2^n \\ f_i(a_2, \ldots, a_n) = 1}} P(x_2 = a_2, \ldots, x_n = a_n).$$

Очевидно, что все слагаемые, входящие в выражение для f_0, входят и в выражение для f_1. Отсюда следует неравенство (4.3). Переменная x_1 является существенной для f тогда и только тогда, когда существует набор $a_2, \ldots, a_n$, такой что $f(0, a_2, \ldots, a_n) < f(1, a_2, \ldots, a_n)$, а это означает, что в сумме для f_1 найдутся такие положительные слагаемые, которые не входят в сумму для f_0, т. е. неравенство (4.3) является строгим. □

Приведем без доказательства еще одно свойство вероятностной функции для монотонных функций, относящееся к случаю $p_1 = \ldots = p_n = p$.

Теорема 4.7 (Шеннон). *Если* $0 < p < 1$, *то для вероятностной функции* $F_f(p) = F(p)$ *монотонной булевой функции* f *справедливо неравенство*

$$p(1-p)F'(p) \ge F(p)(1 - F(p)),$$

где F' *— производная функции* F. *При этом равенство достигается только в случаях* $F(p) = 0, 1, p$. □

4.2. ЛИНЕЙНЫЕ ПРИБЛИЖЕНИЯ (АППРОКСИМАЦИИ) БУЛЕВЫХ ФУНКЦИЙ

В этом параграфе мы будем рассматривать ситуацию, в которой все переменные x_i являются независимыми равномерно распределенными случайными величинами. При этом, очевидно, $P(f(x_1,\ldots,x_n)=1)=\frac{\|f\|}{2^n}$.

Определение 4.2. *Функцию $g(\mathbf{x})$ будем называть приближением (аппроксимацией) функции $f(\mathbf{x})$, если $P(f(\mathbf{x}) = g(\mathbf{x}))>1/2$.*

Заметим, что если $P(f(\mathbf{x}) = g(\mathbf{x}))<1/2$, то приближением функции $f(\mathbf{x})$ будет являться $\overline{g(\mathbf{x})}$.

Введенное понятие широко изучается в практических приложениях, в частности решается задача о поиске приближения данной функции в различных классах функций. Один из таких классов — линейные функции, т. е. функции вида $a_1x_1\oplus\ldots\oplus a_nx_n =<\alpha,\mathbf{x}>$, $\alpha=(a_1,\ldots,a_n)$. Ниже мы рассмотрим приближение функциями именно из этого класса.

Для произвольной булевой функции $f(\mathbf{x})$ и набора $\alpha\in\Omega_2^n$ определим параметр Δ_α^f следующим образом:

$$P(f(\mathbf{x})=<\alpha,\mathbf{x}>)=\frac{1}{2}+\frac{\Delta_\alpha^f}{2^n}$$

или

$$\Delta_\alpha^f=2^n(P(f(\mathbf{x})=<\alpha,\mathbf{x}>)-\frac{1}{2}). \tag{4.4}$$

Утверждение 4.8.

$$\Delta_\alpha^f=2^{n-1}-\|f(\mathbf{x})\oplus<\alpha,\mathbf{x}>\|.$$

□ Справедливы равенства

$$\begin{aligned}P(f(\mathbf{x})=<\alpha,\mathbf{x}>)&=P(f(\mathbf{x})\oplus<\alpha,\mathbf{x}>=0)=\\&=1-P(f(\mathbf{x})\oplus<\alpha,\mathbf{x}>=1)=\\&=1-\frac{\|f(\mathbf{x})\oplus<\alpha,\mathbf{x}>\|}{2^n}=\frac{1}{2}+\frac{\Delta_\alpha^f}{2^n}.\end{aligned}$$

Отсюда легко получить требуемое соотношение. □

Следствие 1. *Коэффициенты Δ_α^f следующим образом связаны с коэффициентами Уолша–Адамара второго рода:*

$$W_\alpha^f = \frac{1}{2^{n-1}}\Delta_\alpha^f$$

и с коэффициентами Фурье:

$$\Delta_\alpha^f = -2^n C_\alpha^f,\ \alpha \neq \theta,$$

$$\Delta_\theta^f = 2^{n-1} - 2^n C_\theta^f.$$

□ Доказательство следует из утверждения 2.7 и следствия к нему. □

Легко проверяется также равенство:

$$1/2 - f(\mathbf{x}) = \frac{1}{2^n}\sum_{\alpha\in\Omega_2^n}\Delta_\alpha^f(-1)^{<\alpha,\mathbf{x}>}. \qquad (4.5)$$

Следствие 2. *Булева функция однозначно определяется набором коэффициентов Δ_α^f.* □

Свойства коэффициентов Δ_α^f

1^0. *При $n \geq 2$ все коэффициенты Δ_α^f являются целыми числами одинаковой четности.*

То, что коэффициенты являются целыми числами, следует из утверждения 4.8. Четность веса функции определяется ее степенью. А именно, справедливо утверждение (проверьте в качестве упражнения):

Если $l \geq 1$, то функция от l переменных имеет нечетный вес тогда и только тогда, когда ее степень равна l.

В случае $n \leq 2$ прибавление к функции f любой линейной функции не приводит к изменению ее степени. Следовательно, все Δ_α^f либо четны, либо нечетны одновременно.

2^0. $\Delta_\alpha^{\overline{f}} = -\Delta_\alpha^f$.

Несложно видеть, что

$$P(\overline{f(\mathbf{x})} =< \alpha, \mathbf{x} >) = P(f(\mathbf{x}) \neq < \alpha, \mathbf{x} >) = 1 - \frac{1}{2} - \frac{\Delta_\alpha^f}{2^n}.$$

Требуемое равенство следует теперь из определения коэффициента Δ_α^f.

3^0.

$$\sum_\alpha \Delta_\alpha^f = 2^{n-1}(-1)^{f(0,\ldots,0)},\ \sum_\alpha \left(\Delta_\alpha^f\right)^2 = 2^{2(n-1)},$$

$$2^{n/2-1} \le \max_{\alpha\in\Omega_2^n} |\Delta_\alpha^f| \le 2^{n-1}.$$

Соотношения следуют из утверждения 2.8 и первого следствия к утверждению 4.8.

Таким образом, из последнего неравенства вытекает, что у любой булевой функции найдется аффинное приближение $<\alpha,\mathbf{x}>\oplus a_0$, такое что вероятность его совпадения с данной функцией будет не менее, чем $\frac{1}{2}+\frac{1}{2^{n/2+1}}$.

4^0. *Пусть функция* $f(x_1,\ldots,x_n)$ *зависит несущественно от переменных* $x_{k+1},\ldots,x_n$, *т. е.* $f(x_1,\ldots,x_n)=g(x_1,\ldots,x_k)$. *Тогда*

$$\Delta^f_{(a_1,\ldots,a_n)} = \begin{cases} 2^{n-k}\Delta^g_{(a_1,\ldots,a_k)}, & \text{если } a_{k+1}=\ldots=a_n=0 \\ 0, & \text{в противном случае.}\end{cases}$$

Пусть $a_{k+1}=\ldots=a_n=0$. Тогда

$$\Delta^f_{(a_1,\ldots,a_k,0,\ldots,0)} = 2^{n-1} - \|f(\mathbf{x})\oplus a_1x_1\oplus\ldots\oplus a_kx_k\| =$$
$$= 2^{n-1} - 2^{n-k}\|g(x-1,\ldots,x_k)\oplus a_1x_1\oplus\ldots\oplus a_kx_k\| =$$
$$= 2^{n-k}(2^{k-1} - \|g(x_1,\ldots,x_k)\oplus$$
$$\oplus <(a_1,\ldots,a_k),(x_1,\ldots,x_k)>\|) = 2^{n-k}\Delta^g_{(a_1,\ldots,a_k)}.$$

Пусть для некоторого s, $k+1\le s\le n$ верно $a_s\ne 0$. Тогда в многочлен Жегалкина функции $f(\mathbf{x})\oplus<\alpha,\mathbf{x}>$ переменная x_s входит линейно (т. е. только в единственный член — непосредственно x_s). Тогда нетрудно видеть, что данная функция на всех наборах, соседних по переменной x_s, принимает различные значения, а следовательно, ее вес равен 2^{n-1}. Таким образом, $\Delta^f_{(a_1,\ldots,a_n)}=0$ по утверждению 4.8.

5^0. *Пусть* f_1 *и* f_2 *— булевы функции от* n *переменных. Тогда*

$$\Delta_\alpha^{f_1\oplus f_2} = \frac{1}{2^{n-1}}\sum_\beta \Delta_\beta^{f_1}\Delta_{\alpha\oplus\beta}^{f_2}.$$

Суммирование здесь и далее, если не оговаривается иного, будет проводиться по всем элементам множества Ω_2^n.

Доказательство проводится цепочкой равносильных преобразований:

$$\Delta_\alpha^{f_1\oplus f_2} = 2^{n-1}W_\alpha^{f_1\oplus f_2} = \frac{1}{2}\sum_{\mathbf{x}}(-1)^{f_1(\mathbf{x})\oplus f_2(\mathbf{x})\oplus<\alpha,\mathbf{x}>} =$$

$$= \frac{1}{2}\sum_{\mathbf{x}}(-1)^{f_1(\mathbf{x})\oplus<\alpha,\mathbf{x}>}(-1)^{f_2(\mathbf{x})} =$$

$$= \frac{1}{2}\sum_{\mathbf{x}}(-1)^{f_2(\mathbf{x})\oplus<\alpha,\mathbf{x}>}\sum_{\beta}W_\beta^{f_1}(-1)^{<\beta,\mathbf{x}>} =$$

$$= 2^{n-1}\sum_{\beta}W_\beta^{f_1}\left(\frac{1}{2^n}\sum_{\mathbf{x}}(-1)^{f_2(\mathbf{x})\oplus<\alpha\oplus\beta,\mathbf{x}>}\right) =$$

$$= 2^{n-1}\sum_{\beta}W_\beta^{f_1}W_{\alpha\oplus\beta}^{f_2} = \frac{1}{2^{n-1}}\sum_{\beta}\Delta_\beta^{f_1}\Delta_{\alpha\oplus\beta}^{f_2}.$$

Из доказанного свойства легко получить следующее полезное соотношение:

$$\sum_{\beta}\Delta_\beta^f\Delta_{\alpha\oplus\beta}^f = 2^{2(n-1)}\delta_{\alpha,\theta}.$$

Здесь $\delta_{\alpha,\theta}$ принимает значение 1 в том случае, когда $\alpha=\theta$, и значение 0 — в противном случае.

6^0. *Пусть* $f(x_1,\ldots,x_n) = f_1(x_1,\ldots,x_k)\oplus f_2(x_{k+1},\ldots,x_n)$ — *булева функция от* n *переменных. Тогда* $\Delta_\alpha^f = 2\Delta_{\alpha_1}^{f_1}\Delta_{\alpha_2}^{f_2}$, где $\alpha=(a_1,\ldots,a_n)$, $\alpha_1=(a_1,\ldots,a_k)$, $\alpha_2=(a_{k+1},\ldots,a_n)$.

По свойству 4^0 в сумме $\sum_\beta\Delta_\beta^{f_1}\Delta_{\alpha\oplus\beta}^{f_2}$ не равны нулю только те слагаемые, у которых $\beta=(b_1,\ldots,b_k,0,\ldots,0)$ и $\alpha\oplus\beta=(0,\ldots,0,c_{k+1},\ldots,c_n)$, откуда имеем $a_1=b_1,\ldots,a_k=b_k$ и $a_{k+1}=c_{k+1},\ldots,a_n=c_n$. Следовательно, отличное от нуля слагаемое единственно, и по свойству 5^0:

$$\Delta_\alpha^f = \frac{1}{2^{n-1}}(2^{n-k}\Delta_{\alpha_1}^{f_1})(2^k\Delta_{\alpha_2}^{f_2}) = 2\Delta_{\alpha_1}^{f_1}\Delta_{\alpha_2}^{f_2}.$$

7^0. *Для произвольных булевых функций f_1 и f_2 от n переменных справедливо равенство*

$$P(f_1 \oplus f_2 = 1) = \frac{1}{2^{2n}} \sum_{\beta} \left(\Delta_\beta^{f_1} - \Delta_\beta^{f_2}\right)^2.$$

Обозначим левую часть этого равенства буквой A. Используя (4.5), получим

$$P(f_1 \oplus f_2 = 1) = \frac{\|f_1 \oplus f_2\|}{2^{2n}} = \frac{1}{2^n} \sum_{\mathbf{x}} (f_1(\mathbf{x}) \oplus f_2(\mathbf{x})) =$$

$$= \frac{1}{2^n} \sum_{\mathbf{x}} \left(\frac{1}{2} - \frac{1}{2^n} \sum_{\alpha} \Delta_\alpha^{f_1 \oplus f_2} (-1)^{<\alpha, \mathbf{x}>}\right) =$$

$$= \frac{1}{2^n} \sum_{\mathbf{x}} \left(\frac{1}{2} - \frac{1}{2^{2n-1}} \sum_{\alpha} \sum_{\beta} \Delta_\beta^{f_1} \Delta_{\alpha\oplus\beta}^{f_2} (-1)^{<\alpha, \mathbf{x}>}\right) =$$

$$= \frac{1}{2^n} \left(2^{n-1} - \frac{1}{2^{2n-1}} \sum_{\beta} \Delta_\beta^{f_1} \sum_{\alpha} \Delta_{\alpha\oplus\beta}^{f_2} \sum_{\mathbf{x}} (-1)^{<\alpha, \mathbf{x}>}\right).$$

Поскольку внутренняя сумма равна 0 при $\alpha \neq \theta$, а при $\alpha = \theta$ равна 2^n, то

$$A = \frac{1}{2^n} \left(2^{n-1} - \frac{1}{2^{n-1}} \sum_{\beta} \Delta_\beta^{f_1} \Delta_\beta^{f_2}\right) =$$

$$= \frac{1}{2^{2n}} \left(2^{2n-1} - 2 \sum_{\beta} \Delta_\beta^{f_1} \Delta_\beta^{f_2}\right).$$

По свойству 3^0

$$A = \frac{1}{2^{2n}} \left(\sum_{\beta} (\Delta_\beta^{f_1})^2 + \sum_{\beta} (\Delta_\beta^{f_2})^2 - 2 \sum_{\beta} \Delta_\beta^{f_1} \Delta_\beta^{f_2}\right) =$$

$$= \frac{1}{2^{2n}} \sum_{\beta} \left(\Delta_\beta^{f_1} - \Delta_\beta^{f_2}\right)^2.$$

8^0. *Пусть f — функция от n переменных, $f_0 = f_1^0,\ f_1 = f_1^1$ — ее подфункции и $\alpha' = (a_2, \ldots, a_n) \in \Omega_2^{n-1}$. Тогда*

$$\Delta^f_{(0,\alpha')} = \Delta^{f_0}_{\alpha'} + \Delta^{f_1}_{\alpha'},$$

$$\Delta^f_{(1,\alpha')} = \Delta^{f_0}_{\alpha'} - \Delta^{f_1}_{\alpha'}.$$

Справедливы равенства:

$$\Delta^f_{(0,\alpha')} = 2^{n-1} - \|f(\mathbf{x}) \oplus < (0,\alpha'), \mathbf{x} > \| = \\ = 2^{n-1} - \|f(\mathbf{x}) \oplus < \alpha', \mathbf{x}' > \|,$$

где $\mathbf{x}' = (x_2, \ldots, x_n)$. Переходя к подфункциям, имеем

$$2^{n-1} - \|f(\mathbf{x}) \oplus < \alpha', \mathbf{x}' > \| = 2^{n-1} - \\ - \|f_0(\mathbf{x}') \oplus < \alpha', \mathbf{x}' > \| - \|f_1(\mathbf{x}') \oplus < \alpha', \mathbf{x}' > \| = \\ = 2^{n-2} - \|f_0(\mathbf{x}') \oplus < \alpha', \mathbf{x}' > \| + 2^{n-2} - \\ - \|f_1(\mathbf{x}') \oplus < \alpha', \mathbf{x}' > \| = \Delta^{f_0}_{\alpha'} + \Delta^{f_1}_{\alpha'}.$$

Во втором случае имеем

$$\Delta^f_{(1,\alpha')} = 2^{n-1} - \|f(\mathbf{x}) \oplus < (1,\alpha'), \mathbf{x} > \| = 2^{n-1} - \\ - \|f(\mathbf{x}) \oplus x_1 \oplus < \alpha', \mathbf{x}' > \| = \\ = 2^{n-1} - \|f_0(\mathbf{x}') \oplus < \alpha', \mathbf{x}' > \| - \|f_1(\mathbf{x}') \oplus 1 \oplus < \alpha', \mathbf{x}' > \| = \\ = \|f_1(\mathbf{x}') \oplus < \alpha', \mathbf{x}' > \| - \|f_0(\mathbf{x}') \oplus < \alpha', \mathbf{x}' > \| = \\ = \Delta^{f_0}_{\alpha'} - \Delta^{f_1}_{\alpha'}.$$

9^0. *Пусть A — невырожденная матрица размера $n \times n$ над полем $GF(2)$, $\beta \in \Omega_2^n$. Тогда коэффициенты Δ^f_α и Δ^g_α функций $f(\mathbf{x})$ и $g(\mathbf{x}) = f(\mathbf{x}A \oplus \beta)$ связаны соотношениями:*

$$\Delta^g_{\alpha A^T} = (-1)^{<\alpha,\beta>} \Delta^f_\alpha,\ \alpha \in \Omega_2^n.$$

Доказательством служит цепочка равенств, в которой $\mathbf{y} = \mathbf{x}A \oplus \beta$:

$$\begin{aligned}\Delta^g_{\alpha A^T} &= 2^{n-1} - \|g(\mathbf{x})\oplus < \alpha A^T, \mathbf{x} > \| = \\ &= 2^{n-1} - \|g(\mathbf{x})\oplus < \alpha, \mathbf{x}A > \| = \\ &= 2^{n-1} - \|f(\mathbf{x}A \oplus \beta)\oplus < \alpha, \mathbf{x}A > \| = \\ &= 2^{n-1} - \|f(\mathbf{y})\oplus < \alpha, \mathbf{y} \oplus \beta > \| = \\ &= 2^{n-1} - \|f(\mathbf{y})\oplus < \alpha, \mathbf{y} > \oplus < \alpha, \beta > \| = \\ &= \Delta^{f\oplus<\alpha,\beta>}_{\alpha} = (-1)^{<\alpha,\beta>}\Delta^f_\alpha.\end{aligned}$$

10^0. *Пусть* $g(\mathbf{x}) = f(\mathbf{x})\oplus < \gamma, \mathbf{x} > \oplus d_0$, *где* $\gamma \in \Omega^n_2$, $d_0 \in \omega_2$. *Тогда* $\Delta^g_{\alpha\oplus\gamma} = (-1)^{d_0}\Delta^f_\alpha$.

Действительно,

$$\begin{aligned}\Delta^g_{\alpha\oplus\gamma} &= 2^{n-1} - \|g(\mathbf{x})\oplus < \alpha \oplus \gamma, \mathbf{x} > \| = \\ &= \Delta^{g(\mathbf{x})\oplus<\gamma,\mathbf{x}>}_{\alpha} = \Delta^{f(\mathbf{x})+d_0}_{\alpha} = (-1)^{d_0}\Delta^f_\alpha.\end{aligned}$$

4.3. КОЭФФИЦИЕНТ АДДИТИВНОСТИ ДИСКРЕТНЫХ ФУНКЦИЙ

Ниже мы покажем, каким образом методы изучения свойств булевых функций можно обобщить на более широкий класс дискретных функций, в частности на функции в абелевых группах.

Одним из естественных способов сравнения характеристик дискретных функций f, определенных на абелевой группе $(G, +)$ со значениями в абелевой группе $(H, +)$, является подход, основанный на свойстве

$$f(a + b) = f(a) + f(b), \tag{4.6}$$

определяющем гомоморфизм групп. Заметим, что функции над конечным простым полем, обладающие свойством (4.6) при любых $a, b \in G$, являются линейными. Вместе с тем вероятность выполнения соотношения (4.6) при случайном и равновероятном выборе элементов $a, b \in G$ можно рассматривать

как меру близости функции f к линейной функции (сравните с коэффициентами Δ_α^f).

Введем следующие обозначения:

H^G — множество всех функций из G в H;

$Lin(G,H)$ — множество всех линейных функций (гомоморфизмов) из H^G;

$Dist(f,g) = P\{f(u) \neq g(u)\}$ — «расстояние» (вероятность несовпадения) между функциями f, g;

$Dist(f) = \min_{g \in Lin(G,H)} Dist(f,g)$ — «расстояние» от f до класса $Lin(G,H)$;

$Err(f) = P\{f(u) + f(v) \neq f(u+v)\}$ — «отклонение» f от линейных функций;

$$REJ_{G,H}(x) = \min\{Err(f) : f : G \to H, Dist(f) = x\}.$$

Здесь и ниже при использовании вероятностной терминологии предполагается, что на группе G задано равномерное распределение, и аргументы u, v выбираются независимо.

При изучении функций из H^G обычно используются комплексные характеры группы G. Выше мы определили (см. параграф 2) комплексные характеры аддитивной группы конечного поля. В этом параграфе мы будем рассматривать группы $G = (\mathbb{Z}/m)^n$, $H = (\mathbb{Z}/m)^k$ — прямые суммы аддитивной группы кольца $\mathbb{Z}/m$. Для них характеры определяются аналогично, разница лишь в том, что вместо простого p используется произвольное натуральное m. По функции $f \in H^G$ и элементу $a \in H$ определяют комплекснозначную функцию $f_a : G \to C$, положив $f_a(x) = \chi_a(f(x))$, где χ_a — характер группы H, соответствующий элементу a. Обозначим через η_b характер группы G, соответствующий элементу b, и разложим функцию f_a в ряд Фурье по характерам группы G:

$$f_a(x) = \sum_{b \in G} C_b^{f_a} \eta_b(x). \tag{4.7}$$

Систему комплексных чисел $C_b^{f_a}$, где $a \in H, b \in G$, называют спектром функции f_a. Из ортогональности характеров следует, что

$$C_b^{f_a} = \frac{1}{|G|} \sum_{x \in G} f_a(x) \bar{\eta}_b^G(x), \tag{4.8}$$

где $\bar{\eta}_b^G$ — характер, комплексно-сопряженный к η_b^G. Числа $C_b^{f_a}$ при любых $a, b \in G$ обладают следующими свойствами:

$$|C_b^{f_a}| \leqslant 1, \tag{4.9}$$

$$\sum_{b \in G} |C_b^{f_a}|^2 = 1, \tag{4.10}$$

$$\max |C_b^{f_a}| \geqslant \frac{1}{\sqrt{|G|}}. \tag{4.11}$$

Свойство (4.9) получается из (4.8), если учесть, что значениями функций f_a, χ_b^G являются комплексные корни некоторых степеней из 1, и их модули равны 1. Свойство (4.10) следует из соотношений ортогональности характеров. Свойство (4.11) следует из (4.10).

Определение 4.3. *Функция $f : G \to H$ называется бент-функцией, если числа $|C_b^{f_a}|$ равны между собой при любых $b \in G, a \in H \backslash \{0\}$.*

Из равенства (4.10) следует, что для бент-функции $f \in H^G$ выполняются равенства $|C_b^{f_a}| = |G|^{-\frac{1}{2}}$.

Определение 4.4. *Коэффициентом аддитивности функции $f \in H^G$ назовем число*

$$Ad(f) = \frac{|\{(x, y) \in G^2 : f(x + y) = f(x) + f(y)\}|}{|G|^2}.$$

Теорема 4.9. *Пусть $f : G \to H$, где $G = (\mathbb{Z}/m)^n$, $H = (\mathbb{Z}/m)^k$ — прямые суммы аддитивной группы $\mathbb{Z}/m$. Тогда имеет место равенство*

$$Ad(f) = \frac{1}{|H|} \sum_{a \in H} \sum_{b \in G} C_b^{f_a} |C_b^{f_a}|^2 = \\ = \frac{1}{|H|} \sum_{a \in H} \sum_{b \in G} C_b^{f_a} C_b^{f_a} \bar{C}_b^{f_a}. \tag{4.12}$$

□ Обозначим правую часть равенства (4.12) буквой X и вычислим ее. Так как группа Z/m — циклическая порядка m, то ее характер φ_u, соответствующий элементу u, определяется равенствами $\varphi_u(x) = e_1^{ux}$, где e_1 — фиксированный первообразный корень степени m из 1. Не теряя общности, можно считать, что для всех прямых слагаемых групп G, H используется один и тот же корень $e^{\frac{2\pi i}{m}}$. Тогда из (4.7), (4.8) получаем

$$C_b^{f_a} = \frac{1}{|G|}\sum_{x\in G} f_a(x)\bar{\eta}_b^G(x) =$$

$$= \frac{1}{|G|}\sum_{x\in G}\chi_a(f(x))\bar{\eta}_b^G(x) = \frac{1}{|G|}\sum_{x\in G} e^{\frac{2\pi i}{m}(<a,f(x)>-<b,x>)},$$

где для элементов $u = (u_1,\ldots,u_r), v = (v_1,\ldots,v_r) \in (Z/m)^r$ под (u,v) понимается число $u_1v_1 + \ldots + u_rv_r(\mathrm{mod} m)$. Отсюда, учитывая то, что элемент, сопряженный с корнем из 1, совпадает с обратным элементом к этому корню, получим

$$X = \frac{1}{|H||G|^3}\times$$

$$\times\sum_{a\in H}\sum_{b\in G}\left(\sum_{x,y,z\in G} e^{\frac{2\pi i}{m}(<a,f(x)+f(y)-f(z)>-<b,x+y-z>)}\right) =$$

$$= \frac{1}{|H||G|^3}\sum_{a\in H}\sum_{x,y,z\in G} e^{\frac{2\pi i}{m}<a,f(x)+f(y)-f(z)>}\times$$

$$\times\sum_{b\in G} e^{-\frac{2\pi i}{m}<b,x+y-z>}. \quad (4.13)$$

Легко проверить, что для любого $u_1 \in \{1,...,m-1\}$ выполняется равенство

$$e^{\frac{2\pi i}{m}u_10} + e^{\frac{2\pi i}{m}u_11} + ... + e^{\frac{2\pi i}{m}u_1(m-1)} = 0.$$

Следовательно, последняя сумма в (4.13) равна 0 при $x+y \neq z$ и равна $|G|$ при $x+y=z$. Приняв во внимание этот факт и поменяв в (4.13) порядок суммирования, получим

$$Ad(f) = \frac{1}{|H||G|^2}\sum_{x,y\in G}\sum_{a\in H} e^{\frac{2\pi i}{m}<a,f(x)+f(y)-f(x+y)>}. \quad (4.14)$$

По тем же соображениям, что и выше, последняя сумма в (4.14) равна 0, если $f(x)+f(y)\neq f(x+y)$, и равна $|H|$ в противном случае. Значит,

$$X=\frac{1}{|G|^2}\sum_{\substack{x,\,y\,\in\,G,\\ f(x+y)\,=\,f(x)+f(y)}}1=Ad(f),$$

и теорема доказана. □

В том случае, когда f — булева функция от n переменных, формула (4.12) несколько упрощается, поскольку в этом случае $C_b^{f_0}=0$ при любом $b\in G$. Однако и при таком упрощении для вычисления коэффициента $Ad(f)$ необходимо знание всех 2^n коэффициентов $C_b^{f_1}$. В связи с этим естественно возникает вопрос о наличии меньшего числа параметров, определяющих величину $Ad(f)$ для булевой функции f. Докажем, что такая система параметров существует.

Теорема 5.10 (М. М. Глухов). *Пусть* $f\in F_2(n)$, $\|f\|=k$, $N_f=\{\gamma_1,\ldots,\gamma_k\}$, M *— множество из* m *неупорядоченных троек* $\{\alpha,\beta,\gamma\}$ *с ненулевыми компонентами из* N_f, *удовлетворяющими условию*

$$\alpha\oplus\beta\oplus\gamma=0. \tag{4.15}$$

Тогда

$$Ad(f)=1-\frac{6}{2^{2n}}((2^{n-1}-k)k+4m),\ \textit{если}\ f(0)=0, \tag{4.16}$$

и

$$\begin{aligned}&Ad(f)=1-\tfrac{6}{2^{2n}}(2^{n-1}k-(k-1)^2+4m-\tfrac{1}{3}),\\ &\textit{если}\ f(0)=1, m=|M|.\end{aligned} \tag{4.17}$$

□ Обозначим

$$U_i=\{(\alpha,\beta):\alpha,\beta\in\Omega_2^n,\ |\{\alpha,\beta,\alpha\oplus\beta\}\cap N_f|=i\},\ i\in\overline{0,3}.$$

В итоге множество $(\Omega_2^n)^2$ всех упорядоченных пар (α,β) элементов из Ω_2^n разбивается на 4 попарно непересекающихся подмножества $U_i, i\in\overline{0,3}$.

Очевидно, что для пар $(\alpha, \beta) \in U_0 \cup U_2$ и для пар $(\alpha, \beta) \in \in U_1 \cup U_3$ выполняются равенства

$$f(\alpha) \oplus f(\beta) = f(\alpha \oplus \beta).$$

Следовательно

$$Ad(f) = \frac{|U_0 \cup U_2|}{2^{2n}} = 1 - \frac{|U_1 \cup U_3|}{2^{2n}}. \qquad (4.18)$$

Таким образом, для вычисления $Ad(f)$ достаточно найти числа $|U_1|$ и $|U_3|$.

Рассмотрим сначала случай, когда $f(0) = 0$, т. е. множество N_f не содержит нулевого вектора.

Вычисление $|U_3|$.

Из определения $|U_3|$ следует, что элементы пары $(\alpha, \beta) \in U_3$ содержатся в некоторой тройке из M. Так как N_f не содержит нулевого вектора, то компоненты каждой из таких троек попарно различны. Поэтому из каждой тройки (α, β, γ) получается 6 пар множества U_3:

$$(\alpha, \beta), (\beta, \alpha), (\alpha, \gamma), (\gamma, \alpha), (\beta, \gamma), (\gamma, \beta).$$

Следовательно $|U_3| = 6m$.

Вычисление $|U_1|$.

Найдем сначала число N неупорядоченных троек (α, β, γ) с различными ненулевыми компонентами из Ω_2^n с нулевой суммой компонент, в которых ровно одна компонента содержится в N_f.

Замечаем, что для каждого $\gamma_t \in N_f$ существует ровно 2^{n-1} неупорядоченных пар $\{\alpha, \beta\}$, таких что $\alpha \oplus \beta = \gamma_t$ (это пары вида $\{\alpha, \alpha \oplus \gamma_t\}$). Обозначим множество всех таких пар через M_t. Нас интересуют лишь те из них, в которых $\alpha, \beta \notin N_f$. В связи с этим найдем сначала число пар $\{\alpha, \beta\}$, в которых $\alpha, \beta \in N_f$. Обозначим через d_t число соотношений вида (4.15), в которые входит γ_t. Каждое такое соотношение дает одну

неупорядоченную пару $\{\gamma_r, \gamma_s\} \in M_t$. Всего же получим d_t таких пар. Удалив из M_t эти пары, а также пару $\{0, \gamma_t\}$, получим $2^{n-1} - (d_t + 1)$ пар, а значит, и столько же неупорядоченных троек $\{\alpha, \beta, \gamma_t\}$, в которых $\alpha \neq 0, \beta \neq 0$ и $\alpha \notin N_f$ или $\beta \notin N_f$. Суммируя числа $2^{n-1} - (d_t + 1)$ по $t = 1, \ldots, k$, получим число $(2^{n-1} - 1)k - (d_1 + \ldots + d_k)$. Сумма $d_1 + \ldots + d_k$ равна числу вхождений всех букв из N_f в тройки из M и, значит, равна $3m$. Таким образом, получено множество T из $(2^{n-1} - 1)k - 3m$ троек вида $\{\alpha, \beta, \gamma_i\}$. Из его построения видно, что оно содержит все тройки с попарно различными компонентами и с нулевой суммой компонент, в которых γ_i пробегает множество N_f, α, β отличны от 0 и либо оба вектора α, β не принадлежат N_f, либо только один из них принадлежит N_f, причем в последнем случае каждая тройка получена дважды. Действительно, если тройка $\{\alpha, \gamma_i, \gamma_j\} \in T$, то она получена один раз при $t = i$ и один раз при $t = j$. Подсчитаем теперь число троек в T с ровно двумя различными компонентами из N_f. Заметим, что при любых двух компонентах $\gamma_i, \gamma_j \in N_f$ тройки с нулевой суммой компонент ее третья компонента определяется однозначно. Однако она может принадлежать или не принадлежать N_f. Так как пара векторов γ_i, γ_j может входить лишь в одну тройку из M, то при всех возможных парах различных индексов i, j мы всего получим m различных троек с тремя компонентами из N_f, причем каждая такая тройка будет получена трижды, так как тройка $\{\gamma_r, \gamma_i, \gamma_j\}$ будет получена при трех различных сочетаниях из чисел r, i, j по 2. Отсюда следует, что число различных троек в M равно $C_k^2 - 3m$, и

$$N = (2n - 1 - 1)k - 3m - 2(C_k^2 - 3m) = \\ = (2^{n-1} - 1)k + 3m - k(k-1) = (2^{n-1} - k)k + 3m.$$

Заметим теперь, что из каждой полученной тройки $\{\alpha, \beta, \gamma\}$ получается 6 пар из множества U_1. Следовательно

$$|U_1| = 6((2^{n-1} - k)k + 3m). \tag{4.19}$$

Из полученных результатов находим

$$|U_1| + |U_3| = 6((2^{n-1} - k)k + 4m).$$

Теперь из (4.18) следует (4.16), и теорема в рассматриваемом случае доказана.

Пусть теперь $f(0) = 1$, т. е. N_f содержит нулевой вектор 0. Тогда все пары из U_3 без нулевых компонент строятся по тройкам из M, а пары из U_3, содержащие нулевые компоненты, — по тройкам $(\gamma_i, \gamma_i, 0), i \in \overline{1,k}$. При этом из каждой тройки без нулей получится 6 пар, из тройки $(\gamma_i, \gamma_i, 0)$ — 3 пары при $\gamma_i \neq 0$ и одна пара при $\gamma_i = 0$. В итоге имеем

$$|U_3| = 6m + 3(k-1) + 1.$$

Для нахождения числа $|U_1|$ можно поступить следующим образом. Сначала, как и в рассмотренном выше случае, найдем все тройки из U_1, содержащие по одному ненулевому вектору из N_f. Так как число таких элементов равно $k-1$, то в соответствии с формулой (4.19) число полученных троек равно $6((2^{n-1} - (k-1))(k-1) + 3m)$. Теперь найдем все тройки из U_1, содержащие нулевые компоненты. Они исчерпываются $2^n - k$ тройками $\{\alpha, \alpha, 0\}$, где $\alpha \neq \gamma_i, i \in \overline{1,k}$. Из каждой такой тройки получается 3 пары из множества U_1:

$$(\alpha, \alpha), (\alpha, 0), (0, \alpha).$$

В итоге получаем $|U_1| = 6((2^{n-1} - k + 1)(k-1) + 3m) + 3(2^n - k)$, и $|U_1 \cup U_3| = 6((2^{n-1} - k + 1)(k-1) + 3m) + 3(2^n - k) + 6m + 3(k-1) + 1 = (2^{n-1}k - (k-1)^2 + 4m - \frac{1}{3})$.

Отсюда и из (4.18) следует утверждение теоремы и в этом случае. В итоге теорема полностью доказана. □

Из формул (4.16), (4.17) видно, что показатель аффинности функции f определяется тремя ее параметрами: числом переменных n, весом k и числом m соотношений вида (4.15) между векторами из N_f. Заметим, что число m можно рассматривать так же, как число подпространств размерности 2, содержащихся в N_f.

Глава 5

НЕКОТОРЫЕ СПЕЦИАЛЬНЫЕ КЛАССЫ ДИСКРЕТНЫХ ФУНКЦИЙ

Здесь мы рассмотрим отдельные классы булевых функций, для которых рассмотренные параметры представляют интерес для практических приложений.

5.1. КОРРЕЛЯЦИОННО-ИММУННЫЕ ФУНКЦИИ

Пусть снова x_i — независимые равномерно распределенные случайные величины, принимающие значения из множества Ω_2. Для n-местной булевой функциии f и вектора $\alpha = (a_1, \ldots, a_n) \in \Omega_2^n, \alpha \neq \theta$, определим случайные величины $\xi = f(x_1, \ldots, x_n)$ и $\eta =< \alpha, \mathbf{x} >= a_1 x_1 \oplus \ldots \oplus a_n x_n$. Справедлива

Лемма 5.1. *Случайные величины ξ и η независимы тогда и только тогда, когда $C_\alpha^f = 0$.*

□ Две дискретные случайные величины ξ и η независимы тогда и только тогда, когда $P(\xi = a/\eta = b) = P(\xi = a)$ для любых a, b из множеств значений этих случайных величин. Поскольку $f(\mathbf{x})$ и $< \alpha, \mathbf{x} >$ — булевы функции, то необходимо и достаточно доказать, например, равенство: $P(f(\mathbf{x}) = = 1/ < \alpha, \mathbf{x} >= 0) = P(f(\mathbf{x}) = 1/ < \alpha, \mathbf{x} >= 1)$. Имеем

$$P(f(\mathbf{x}) = 1/ < \alpha, \mathbf{x} >= 0) =$$
$$= \frac{P(f(\mathbf{x}) = 1 \text{ и } < \alpha, \mathbf{x} >= 0)}{P(< \alpha, \mathbf{x} >= 0)} = 2 \cdot \frac{1}{2^n} \sum_{\mathbf{x}:<\alpha,\mathbf{x}>=0} f(\mathbf{x}).$$

Аналогично,

$$P(f(\mathbf{x}) = 1/ < \alpha, \mathbf{x} >= 1) = \sum_{\mathbf{x}:<\alpha,\mathbf{x}>=1} f(\mathbf{x}),$$

тогда разность этих двух условных вероятностей равна

$$\frac{1}{2^{n-1}} \left(\sum_{\mathbf{x}:<\alpha,\mathbf{x}>=0} f(\mathbf{x}) - \sum_{\mathbf{x}:<\alpha,\mathbf{x}>=1} f(\mathbf{x}) \right) = \\ = \frac{2}{2^n} \sum_{\mathbf{x}} f(\mathbf{x})(-1)^{<\alpha,\mathbf{x}>} = 2C_\alpha^f.$$

Таким образом, эти вероятности равны тогда и только тогда, когда $C_\alpha^f = 0$. □

Справедливо следующее утверждение, которое в теории вероятностей носит название «линейно-комбинационная лемма». Мы приводим его без доказательства, которое можно найти, например, в [57].

Лемма 5.2. *Пусть $\xi_1, \ldots, \xi_k$ — независимые в совокупности случайные величины со значениями в Ω_2, η — случайная величина со значениями в том же множестве. Тогда случайные величины $\xi_1, \ldots, \xi_k, \eta$ независимы в совокупности тогда и только тогда, когда для любого ненулевого набора $(a_1, \ldots, a_k) \in \Omega_2^k$ случайные величины η и $\zeta = a_1\xi_1 \oplus \ldots \oplus a_k\xi_k$ независимы.* □

Определение 5.1. *Булева функция $f(x_1, \ldots, x_n)$ называется корреляционно-иммунной порядка k, если для независимых в совокупности равномерно распределенных случайных величин $x_1, \ldots, x_n$, принимающих значения из Ω_2, для любого сочетания $j_1, \ldots, j_k$, $1 \le j_1 < \ldots < j_k \le n$ случайная величина $\xi = f(x_1, \ldots, x_n)$ не зависит от случайных величин $x_{j_1}, \ldots, x_{j_k}$, т. е. для любых $a, c_1, \ldots, c_k \in \Omega_2$:*

$$P(f(x_1, ..., x_n) = a / x_{j_1} = c_1, ..., x_{j_k} = c_k) = \\ = P(f(x_1, ..., x_n) = a).$$

Понятие корреляционной иммунности булевой функции отражает ее способность противостоять корреляционному методу анализа в криптографии. Не вдаваясь в детали данного криптографического метода, можно сказать, что это свойство означает отсутствие какой-либо информации в известном значении булевой функции о значениях некоторого подмножества ее аргументов или о значениях некоторых функций ее аргументов. В противном случае, т. е. если такая информация есть, можно пытаться использовать эту информацию для нахождения неизвестного ключа.

Справедлива

Теорема 5.3. *Булева функция $f(x_1, \ldots, x_n)$ является корреляционно-иммунной порядка k тогда и только тогда, когда для всех $\alpha \in \Omega_2^n$, $1 \le \|\alpha\| \le k$ выполняется условие $C_\alpha^f = 0$.*

□ Доказательство легко следует из определения 5.1 и лемм 5.1, 5.2. □

В заключение заметим, что свойство корреляционной иммунности любого порядка не влечет за собой свойства устойчивости (равновероятности) функции. Например, функция

$$f(x_1, x_2, x_3, x_4) = x_1x_3 \oplus x_1x_4 \oplus \\ \oplus x_2x_3 \oplus x_2x_4 \oplus x_1 \oplus x_2 \oplus x_3 \oplus x_4$$

является корреляционно-иммунной порядка 1, но вес ее равен 12.

5.2. k-УСТОЙЧИВЫЕ ФУНКЦИИ

Определение 5.2. *Булева функция $f(x_1, \ldots, x_n)$ называется k-устойчивой или эластичной порядка k, если все ее подфункции, полученные произвольными фиксациями произвольных k ее переменных ($1 \le k \le n-1$), являются устойчивыми. Множество k-устойчивых функций от n переменных будем обозначать $R(n, k)$. Множество $R(n, 0)$ всех устойчивых функций обозначим $R(n)$.*

Отметим, что в некоторых источниках k-устойчивые функции называют *эластичными (resilient) порядка k*.

Лемма 5.4. *Справедливы включения:*

$$R(n) = R(n,0) \supset R(n,1) \supset \ldots \supset R(n,n-1).$$

□ Покажем, что $R(n,k-1) \supset R(n,k)$ при $k \geq 2$. Имеем

$$\|f^{a_{i_1}\ldots a_{i_{k-1}}}_{i_1\ldots i_{k-1}}\| = \|f^{a_{i_1}\ldots a_{i_{k-1}},0}_{i_1\ldots i_{k-1},i_k}\| + \|f^{a_{i_1}\ldots a_{i_{k-1}},1}_{i_1\ldots i_{k-1},i_k}\|.$$

Если веса подфункций в правой части равны 2^{n-k-1}, то вес подфункции в левой части равен 2^{n-k} и f – k-устойчива.

Для доказательства первого включения в цепочке ($k=1$) воспользуемся равенством $\|f\| = \|f^1_1\| + \|f^0_1\|$. □

Из доказанной леммы следует, что k-устойчивая функция будет m-устойчивой при любом $m \leq k$. Таким образом, в любой ее весовой структуре $\|f^{a_{i_1}\ldots a_{i_m}}_{i_1\ldots i_m}\| = 2^{n-m-1}$.

Изучим теперь особенности коэффициентов Δ^f_α k-устойчивой функции f.

Теорема 5.5. *Для того чтобы функция $f(x_1,\ldots,x_n)$ являлась k-устойчивой, необходимо и достаточно, чтобы для любого α из Ω^n_2, такого что $\|\alpha\| \leq k$, выполнялось условие $\Delta^f_\alpha = 0$.*

□ Необходимость. Пусть f — k-устойчива и $\|\alpha\| = k$, $k \geq 1$, причем единицы в наборе α находятся на местах с номерами $i_1,\ldots,i_k$. Тогда

$$\begin{aligned}\Delta^f_\alpha &= 2^{n-1} - \|f \oplus x_{i_1} \oplus \ldots \oplus x_{i_k}\| = \\ &= 2^{n-1} - \sum_{(a_1,\ldots,a_k)\in\Omega^k_2} \|f^{a_1\ldots a_k}_{i_1\ldots i_k} \oplus a_1 \oplus \ldots \oplus a_k\| = \\ &= 2^{n-1} - 2^k 2^{n-k-1} = 0.\end{aligned}$$

Для случая $k=0$ имеем $\Delta^f_\theta = 2^{n-1} - \|f\| = 0$.

Так как k-устойчивая функция по лемме 5.4 является m-устойчивой для любого $m < k$, то $\Delta^f_\alpha = 0$ и для всех α веса, меньшего чем k.

Достаточность. Пусть для любого α из Ω^n_2, такого что $\|\alpha\| \leq k$, выполняется $\Delta^f_\alpha = 0$. Покажем, что f k-устойчива.

Применим метод математической индукции по k. Для $k = 0$ имеем $\|f\| = 2^{n-1}$, и все доказано.

Пусть утверждение верно для всех $m < k$. По свойству 8^0 коэффициентов Δ^f_α:

$$0 = \Delta^f_{(0,\alpha')} = \Delta^{f_0}_{\alpha'} + \Delta^{f_1}_{\alpha'},\ 0 = \Delta^f_{(1,\alpha')} = \Delta^{f_0}_{\alpha'} - \Delta^{f_1}_{\alpha'}$$

для всех α', таких что $\|\alpha'\| < k$. Отсюда следует, что для таких α' имеем $\Delta^{f_0}_{\alpha'} = \Delta^{f_1}_{\alpha'} = 0$. Тогда, по предположению индукции, функции f_0 и f_1, полученные из f фиксацией первой переменной 0 и 1, будут $k-1$-устойчивы. Поскольку аналогичные рассуждения можно провести для любой другой переменной, то функция f по определению является k-устойчивой.□

Заметим, что из теорем 5.3 и 5.5, а также из следствия утверждения 4.8 следует, что k-устойчивая функция является устойчивой корреляционно-иммунной функцией порядка k. Доказанная теорема также показывает, что у k-устойчивой функции не существует аффинных приближений, существенно зависящих от не более чем k переменных. Справедливо и более сильное утверждение:

Теорема 5.6. *Функция является k-устойчивой тогда и только тогда, когда у нее не существует приближений среди всех функций, существенно зависящих от $m \le k$ переменных.*

□ Справедливость обратного утверждения очевидным образом следует из теоремы 5.5. Докажем прямое утверждение. Пусть $f(x_1, \ldots, x_n)$ — k-устойчива. Рассмотрим произвольную функцию $g(x_{i_1}, \ldots, x_{i_k})$, $\{x_{i_1}, \ldots, x_{i_k}\} \subset \{x_1, \ldots, x_n\}$. Тогда вероятность совпадения значений функций f и g равна:

$$P(f = g) = 1 - \frac{1}{2^n}\|f(x_1, \ldots, x_n) \oplus g(x_{i_1}, \ldots, x_{i_k})\| =$$

$$= 1 - \frac{1}{2^n} \sum_{(a_1,\ldots,a_k)\in\Omega_2^k} \|f^{a_1\ldots a_k}_{i_1\ldots i_{k-1}} \oplus g(a_1, \ldots, a_k)\| =$$

$$= 1 - \frac{1}{2^n} 2^k 2^{n-k-1} = \frac{1}{2}. \square$$

5.3. БЕНТ-ФУНКЦИИ

Как было отмечено в предыдущем параграфе, k-устойчивые функции не имеют аффинных приближений среди функций, существенно зависящих от не более чем k переменных. Однако, как показывают следующие рассуждения, такие функции могут иметь приближения, зависящие от большего количества переменных, имеющие бóльшую вероятность совпадения с данной функцией. Действительно, пусть нам известно, что N коэффициентов Δ_α^f функции f равны 0. По свойству 3^0:

$$\sum_\alpha \left(\Delta_\alpha^f\right)^2 = 2^{2(n-1)},$$

и поэтому

$$\max_\alpha |\Delta_\alpha^f| \geq \sqrt{\frac{2^{2(n-1)}}{2^n - N}}.$$

Для k-устойчивых функции $N = \binom{n}{0} + \ldots + \binom{n}{k}$.

В связи с этим возникает вопрос о существовании функций, для которых не будет существовать аффинных приближений, вероятность совпадения которых с данной функцией не будет превышать предельного значения, определяемого свойством 3^0 коэффициентов Δ_α^f.

Определение 5.3. *Булева функция $f(x_1, \ldots, x_n)$ называется бент-функцией, если для любой линейной функции $< \alpha, \mathbf{x} >$ справедливо равенство*

$$|P(f(\mathbf{x}) =< \alpha, \mathbf{x} >) - \frac{1}{2}| = \frac{1}{2^{n/2+1}}.$$

Таким образом, для бент-функции при любом наборе α имеем $|\Delta_\alpha^f| = 2^{n/2-1}$ и $|W_\alpha^f| = \frac{1}{2^{n/2}}$.

Пример. Пусть $n = 2$. Функция $f(x_1, x_2) = x_1 x_2$ является бент-функцией. Этот факт несложно проверить, используя утверждение 4.8.

Опишем некоторые свойства бент-функций.

Теорема 5.7. *Пусть функция $f(x_1, \ldots, x_n)$ является бент-функцией. Тогда $n = 2k$ для некоторого k. При этом, если*

$k > 1$, то степень многочлена Жегалкина функции f не больше, чем k.

□ 1. Так как коэффициенты Δ_α^f являются целыми числами, то по определению бент-функции n должно быть четным целым числом.

2. Пусть $n > m > k$. Покажем, что многочлен Жегалкина функции f не содержит члена $x_1 \dots x_m$. Для этого рассмотрим функцию $g(x_1, \dots, x_m) = f(x_1, \dots, x_m, 0, \dots, 0)$, зависящую от m переменных.

Докажем, что коэффициент Δ_θ^g функции g, соответствующий нулевому набору, равен

$$\frac{1}{2^{n-m}} \sum_\alpha \Delta_\alpha^f,$$

где суммирование ведется по наборам α вида $\alpha = (0, \dots, 0, a_{m+1}, \dots, a_n)$. Действительно,

$$\sum_\alpha \Delta_\alpha^f = 2^{n-1} \sum_\alpha W_\alpha^f = \frac{2^{n-1}}{2^n} \sum_\alpha \sum_{\mathbf{x} \in \Omega_2^n} (-1)^{f(\mathbf{x}) \oplus <\alpha, \mathbf{x}>} =$$

$$= \frac{1}{2} \sum_{\mathbf{x} \in \Omega_2^n} (-1)^{f(\mathbf{x})} \sum_\alpha (-1)^{<\alpha, \mathbf{x}>} =$$

$$= \frac{1}{2} \sum_{\mathbf{x}} (-1)^{f(\mathbf{x})} \sum_{(a_{m+1}, \dots, a_n)} (-1)^{a_{m+1} x_{m+1} \oplus \dots \oplus a_n x_n} =$$

$$= 2^{n-m-1} \sum_{\mathbf{x} = (x_1, \dots, x_m, 0, \dots, 0)} (-1)^{f(\mathbf{x})} =$$

$$= 2^{n-m-1} \sum_{(x_1, \dots, x_m)} (-1)^{g(x_1, \dots, x_m) \oplus <(0, \dots, 0), (x_1, \dots, x_m)>} =$$

$$= 2^{n-1} W_{(0, \dots, 0)}^{g(x_1, \dots, x_m)} = 2^{n-m} \Delta_{(0, \dots, 0)}^{g(x_1, \dots, x_m)}.$$

Тогда, поскольку f является бент-функцией, то $\Delta_\alpha^f = \pm 2^{k-1}$ для всех α, и следовательно:

$$\Delta_{(0, \dots, 0)}^{g(x_1, \dots, x_m)} = \frac{1}{2^{n-m}} \sum \Delta_\alpha^f = \frac{2^{k-1}}{2^{n-m}} \sum \pm 1.$$

Последняя сумма содержит 2^{n-m} слагаемых, равных ± 1, поэтому при $n > m$ является четным числом, скажем $2t$. Тогда

$$\Delta_{(0,\ldots,0)}^{g(x_1,\ldots,x_m)} = 2^{m-k}t.$$

Отсюда

$$\|g\| = 2^{m-1} - \Delta_{(0,\ldots,0)}^{g(x_1,\ldots,x_m)} = 2^{m-1} - 2^{m-k}t = 2^{m-k}(2^{k-1} - t),$$

и так как $m - k > 0$, то $\|g\|$ — четное число, и следовательно, степень g меньше m и одночлен $x_1 \ldots x_m$ не входит в многочлены Жегалкина функций g и f.

Остается рассмотреть случай $m = n$. В этом случае $\|f\| = 2^{n-1} - \Delta_\theta^f = 2^{n-1} \pm 2^{k-1}$, что при $k > 1$ является четным числом. Следовательно, $\deg f < n$. □

Утверждение 5.8. *Пусть $A \in P_{n\times n}^*$ над $GF(2)$, $\beta \in \Omega_2^n$. Если $f(\mathbf{x})$ — бент-функция от n переменных, то $f(\mathbf{x}A \oplus \beta)$ — также бент-функция.*

□ По свойству 9^0 коэффициентов Δ_α^f:

$$\Delta_{\alpha A^T}^g = (-1)^{<\alpha,\beta>}\Delta_\alpha^f = \pm\frac{1}{2^{n/2-1}}. \square$$

Утверждение 5.9. *Пусть*

$$f(x_1,\ldots,x_n) = f_1(x_1,\ldots,x_l) \oplus f_2(x_{l+1},\ldots,x_n).$$

Тогда f — бент-функция тогда и только тогда, когда f_1 и f_2 — бент-функции.

□ По свойству 6^0 коэффициентов Δ_α^f: $\Delta_\alpha^f = 2\Delta_{\alpha_1}^{f_1}\Delta_{\alpha_2}^{f_2}$, где $\alpha = (a_1,\ldots,a_n), \alpha_1 = (a_1,\ldots,a_l), \alpha_2 = (a_{l+1},\ldots,a_n)$.

Если f_1 и f_2 — бент-функции, то $|\Delta_\alpha^f| = 2\cdot 2^{\frac{l}{2}-1}\cdot 2^{\frac{n-l}{2}-1} = 2^{\frac{n}{2}-1}$. Следовательно, f — бент-функция.

Пусть f — бент-функция. Зафиксируем $\alpha_1 \in \Omega_2^l$. Тогда для любого $\alpha_2 \in \Omega_2^{n-l}$ имеем

$$|\Delta_{\alpha_2}^{f_2}| = \frac{|\Delta_{(\alpha_1\alpha_2)}^f|}{2|\Delta_{\alpha_1}^{f_1}|} = \frac{2^{n/2-2}}{|\Delta_{\alpha_1}^{f_1}|}.$$

Таким образом, все коэффициенты $\Delta_{\alpha_2}^{f_2}$ равны по модулю и f_2 — бент-функция. Аналогично показывается, что f_1 также бент-функция. □

Из трех предыдущих утверждений следует

Теорема 5.10. *Пусть $n = 2k$, $f(x_1, \ldots, x_n)$ — бент-функция степени $k \geq 3$. Тогда для любой функции g, аффинно эквивалентной функции f (т. е. принадлежащей $[f]_{AL}$), g не может быть представлена в виде суммы двух функций от непересекающихся множеств переменных.*

□ Из утверждения 5.8 следует, что g является бент-функцией. Предположим, что есть $g = g_1(x_1, \ldots, x_s) \oplus \oplus\ g_2(x_{s+1}, \ldots, x_n)$. При этом g_1 и g_2 по утверждению 5.9 являются бент-функциями. Так как $\deg g = k \geq 3$, то хотя бы одна из функций, предположим g_1, имеет степень k. По теореме 5.7, $s = 2k'$. Тогда, с одной стороны, $\deg g_1 \leq k'$ (вновь по теореме 5.7, так как $2k' \geq 3$) и, следовательно, $k \leq k'$. С другой стороны, $k' = \frac{s}{2} < \frac{n}{2}$, так как $s < n$, и значит, $k' < k$. Противоречие. □

Докажем критерий того, что функция является бент-функцией. Для этого введем понятие, которое используется при изучении криптографических свойств функции:

Определение 5.4. *Автокорреляционной функцией булевой функции $f(x_1, \ldots, x_n)$ называется отображение $Cor_f : \Omega_2^n \to \mathbb{R}$, определяемое равенством*

$$Cor_f(\alpha) = \frac{1}{2^n} \sum_{\mathbf{x} \in \Omega_2^n} (-1)^{f(\mathbf{x}) \oplus f(\mathbf{x} \oplus \alpha)}.$$

Теорема 5.11 (Ротхауз). *Булева функция $f(x_1, \ldots, x_n)$ является бент-функцией тогда и только тогда, когда для всякого α выполняется равенство $Cor_f(\alpha) = \delta_{\alpha,\theta}$.*

□ 1. Пусть $f(x_1, \ldots, x_n)$ — бент-функция. Тогда для любого $\alpha \in \Omega_2^n$ имеем $W_\alpha^f = \pm\frac{1}{2^{n/2}}$. Рассмотрим булеву функцию g, определяемую равенством

$$g(\beta) = \begin{cases} 0, & W_\beta^f > 0 \\ 1, & W_\beta^f < 0 \end{cases}, \ \beta \in \Omega_2^n.$$

Несложно видеть, что $\frac{1}{2^{n/2}}(-1)^{g(\alpha)} = W_\alpha^f$. Тогда

$$W_\alpha^g = \frac{1}{2^n}\sum_{\mathbf{x}}(-1)^{g(\mathbf{x})\oplus<\alpha,\mathbf{x}>} =$$
$$= \frac{1}{2^{n/2}}\sum_{\mathbf{x}}(-1)^{<\alpha,\mathbf{x}>}\frac{1}{2^n}\sum_{\mathbf{y}}(-1)^{f(\mathbf{y})\oplus<\mathbf{x},\mathbf{y}>} =$$
$$= \frac{1}{2^{3n/2}}\sum_{\mathbf{y}}(-1)^{f(\mathbf{y})}\sum_{\mathbf{x}}(-1)^{<\alpha\oplus\mathbf{y},\mathbf{x}>} = \frac{1}{2^{n/2}}(-1)^{f(\alpha)}.$$

Заметим (хотя это и не имеет отношения к доказательству), что из последнего соотношения следует, что g — также бент-функция.

По свойству 5^0 коэффициентов Δ_α^f:

$$\delta_{\beta,\theta} = \sum_\alpha W_\alpha^g W_{\alpha\oplus\beta}^g = \frac{1}{2^n} = \sum_\alpha (-1)^{f(\alpha)\oplus f(\alpha\oplus\beta)} = Cor_f(\beta).$$

2. Пусть $Cor_f(\alpha) = \delta_{\alpha,\theta}$. Покажем, что разность $(W_\alpha^f)^2 - (W_\beta^f)^2$ равна нулю:

$$(W_\alpha^f)^2 - (W_\beta^f)^2 =$$
$$= \frac{1}{2^{2n}}\left(\sum_{\mathbf{x},\mathbf{y}}(-1)^{f(\mathbf{x})\oplus f(\mathbf{y})\oplus<\mathbf{x},\alpha>\oplus<\mathbf{y},\alpha>} - \right.$$
$$\left. - \sum_{\mathbf{x},\mathbf{y}}(-1)^{f(\mathbf{x})\oplus f(\mathbf{y})\oplus<\mathbf{x},\beta>\oplus<\mathbf{y},\beta>}\right) =$$
$$= \frac{1}{2^{2n}}\sum_{\mathbf{x},\mathbf{y}}\left((-1)^{<\mathbf{x}\oplus\mathbf{y},\alpha>} - (-1)^{<\mathbf{x}\oplus\mathbf{y},\beta>}\right)(-1)^{f(\mathbf{x})\oplus f(\mathbf{y})}.$$

Выполнив замену $\mathbf{z} = \mathbf{x} \oplus \mathbf{y}$, получим

$$\frac{1}{2^{2n}} \sum_{\mathbf{x},\mathbf{z}} \left((-1)^{<\mathbf{z},\alpha>} - (-1)^{<\mathbf{z},\beta>}\right) (-1)^{f(\mathbf{x})\oplus f(\mathbf{x}\oplus\mathbf{z})} =$$

$$= \frac{1}{2^n} \sum_{\mathbf{z}} \left((-1)^{<\mathbf{z},\alpha>} - (-1)^{<\mathbf{z},\beta>}\right) \frac{1}{2^n} \sum_{\mathbf{x}} (-1)^{f(\mathbf{x})\oplus f(\mathbf{x}\oplus\mathbf{z})} =$$

$$= \frac{1}{2^n} \sum_{\mathbf{z}} \left((-1)^{<\mathbf{z},\alpha>} - (-1)^{<\mathbf{z},\beta>}\right) \delta_{\mathbf{z},\theta} = 0. \ \square$$

Следствие 1. *$f(\mathbf{x})$ — бент-функция тогда и только тогда, когда $f(\mathbf{x}) \oplus f(\mathbf{x} \oplus \alpha)$ устойчива для всех α.* $\square$

Заметим, что для бент-функций в качестве следствия из теоремы 4.9 можно получить достаточно простые формулы вычисления коэффициента аддитивности (см. 4.3).

Теорема 5.12. *Если функция f из теоремы 4.9 является бент-функцией, то*

$$Ad(f) = \begin{cases} \frac{1}{|H|} + \frac{|H|-1}{|G||H|}, & \text{если} \quad f(0) = 0, \\ \frac{1}{|H|} - \frac{1}{|G||H|}, & \text{если} \quad f(0) \neq 0. \end{cases} \tag{5.1}$$

$\square$ Обозначим: $H \backslash \{0\} = H^*$. Согласно равенству (4.12)

$$Ad(f) = \frac{1}{|H|} \sum_{a\in H} \sum_{b\in G} C_b^{f_a} |C_b^{f_a}|^2 =$$

$$= \frac{1}{|H|} \sum_{a\in H^*} \sum_{b\in G} C_b^{f_a} |G|^{-1} + \frac{1}{|H|} \sum_{b\in G} C_b^{f_0} |C_b^{f_0}|^2. \tag{5.2}$$

Из определения функции f_a следует, что $C_b^{f_0} = 0$, если $b \neq 0$ и $C_b^{f_0} = 1$, если $b = 0$.

Тогда

$$\sum_{b\in G} C_b^{f_0} |C_b^{f_0}|^2 = 1,$$

и потому

$$\sum_{a\in H*} \sum_{b\in G} C_b^{f_a} = \sum_{a\in H} \sum_{b\in G} C_b^{f_a} - 1.$$

Далее:

$$\sum_{a\in H}\sum_{b\in G} C_b^{f_a} = \sum_{a\in H}\sum_{b\in G}\sum_{x\in G} e^{\frac{2\pi i}{m}((a,f(x))-(b,x))} =$$
$$= \sum_{x\in G}\sum_{a\in H} e^{\frac{2\pi i}{m}(a,f(x))}\sum_{b\in G} e^{-\frac{2\pi i}{m}(b,x)} = |G|\sum_{a\in H} e^{\frac{2\pi i}{m}(a,f(0))}.$$

Остается заметить, что

$$\sum_{a\in H} e^{\frac{2\pi i}{m}(a,f(0))} = 0 \text{ при } f(0)\neq 0$$

и

$$\sum_{a\in H} e^{\frac{2\pi i}{m}(a,f(0))} = |H| \text{ при } f(0)=0,$$

и подставить найденные величины в (5.2).

В заключение опишем некоторые способы построения бент-функций.

Утверждение 5.13. *Пусть $g(x_1,\ldots,x_k)$ — произвольная булева функция. Тогда*

$$f(y_1,\ldots,y_k,x_1,\ldots,x_k) = x_1y_1\oplus\ldots\oplus x_ky_k\oplus g(x_1,\ldots,x_k)$$

— бент-функция от $n=2k$ переменных.

□ Рассмотрим функцию

$$h_{\alpha,\beta}(\mathbf{x},\mathbf{y}) = f(\mathbf{x},\mathbf{y})\oplus <\alpha,\mathbf{x}>\oplus <\beta,\mathbf{y}> =$$
$$= g(\mathbf{x})\oplus <\alpha,\mathbf{x}>\oplus <\mathbf{x}\oplus\beta,\mathbf{y}>,$$

$$\mathbf{x}=(x_1,\ldots,x_k),\ \mathbf{y}=(y_1,\ldots,y_k),$$

и найдем ее вес. Для этого найдем вес ее подфункции, получаемой фиксацией набора переменных $\mathbf{x}$ набором значений $\gamma\in\Omega_2^n$, и просуммируем по всем γ. Рассмотрим два случая:
1. $\gamma\neq\beta$. Тогда

$$h_{\alpha,\beta}(\gamma,\mathbf{y}) = g(\gamma)\oplus <\alpha,\gamma>\oplus <\gamma\oplus\beta,\mathbf{y}>$$

— линейная функция относительно переменных $\mathbf{y}$, и следовательно,

$$\|h_{\alpha,\beta}(\gamma,\mathbf{y})\| = 2^{k-1}.$$

2. $\gamma = \beta$.

$$h_{\alpha,\beta}(\beta,\mathbf{y}) = g(\beta)\oplus < \alpha,\beta > \oplus < \beta \oplus \beta,\mathbf{y} >= \\ = g(\beta)\oplus < \alpha,\beta >= const.$$

Тогда $\|h_{\alpha,\beta}(\beta,\mathbf{y})\| = 2^k \cdot C$, где $C \in \{0,1\}$.

Отсюда

$$\|h_{\alpha,\beta}(\mathbf{x},\mathbf{y})\| = (2^k-1)2^{k-1}2^k C = \\ = 2^{n-1} + 2^k(C - \frac{1}{2}) = 2^{n-1} \pm 2^{\frac{n}{2}-1}$$

и

$$\Delta^f_{\alpha,\beta} = 2^{n-1} - \|h_{\alpha,\beta}(\mathbf{x},\mathbf{y})\| = \pm 2^{\frac{n}{2}-1}.$$

Следовательно, f — бент-функция. □

Утверждение 5.14. *Пусть $f_1(\mathbf{x})$, $f_2(\mathbf{x})$, $f_3(\mathbf{x})$ — бент-функции, причем такие, что $f_1(\mathbf{x}) \oplus f_2(\mathbf{x}) \oplus f_3(\mathbf{x})$ — также бент-функция ($\mathbf{x} = (x_1,\ldots,x_n)$, $n = 2k$). Тогда функция*

$$f(x_1,\ldots,x_n,x_{n+1},x_{n+2}) = \\ = f_1(\mathbf{x})f_2(\mathbf{x}) \oplus f_2(\mathbf{x})f_3(\mathbf{x}) \oplus f_1(\mathbf{x})f_3(\mathbf{x}) \oplus \\ \oplus (f_1(\mathbf{x}) \oplus f_2(\mathbf{x}))x_{n+1} \oplus (f_1(\mathbf{x}) \oplus f_3(\mathbf{x}))x_{n+2} \oplus x_{n+1}x_{n+2}$$

есть бент-функция от $n+2$ переменных.

□ Рассмотрим функцию

$$M_{\alpha,a,b}(\mathbf{x},x_{n+1},x_{n+2}) = \\ = f(\mathbf{x},x_{n+1},x_{n+2})\oplus < \alpha,\mathbf{x} > \oplus a x_{n+1} \oplus b x_{n+2}$$

и найдем ее вес.

$$\|M_{\alpha,a,b}\| = \sum_{(c_1,\ldots,c_{n+1})\in\Omega_2^{n+1}} \|(M_{\alpha,a,b})^{c_1\ldots c_{n+1}}_{1\ldots n+1}\|.$$

Последнюю сумму разобьем на две подсуммы S_1 и S_2 по подмножествам Ω_2^{n+1}:

$$C' = \{(c_1, \ldots, c_{n+1}) | f_1(c_1, \ldots, c_n) \oplus f_3(c_1, \ldots, c_n) \oplus c_{n+1} \oplus b = 1\},$$

$$C'' = \{(c_1, \ldots, c_{n+1}) | f_1(c_1, \ldots, c_n) \oplus f_3(c_1, \ldots, c_n) \oplus c_{n+1} \oplus b = 0\}$$

соответственно. Подсчитаем теперь веса подфункций, получаемых фиксацией переменных $x_1, \ldots, x_{n+1}$ наборами значений из C' и C''.

1. $(c_1, \ldots, c_{n+1}) \in C'$. В этом случае

$$\begin{aligned}(M_{\alpha,a,b})_{1\ldots n+1}^{c_1 \ldots c_{n+1}} = \\ = f_1(c_1, ..., c_n) f_2(c_1, ..., c_n) \oplus f_1(c_1, ..., c_n) f_3(c_1, ..., c_n) \oplus \\ \oplus f_2(c_1, ..., c_n) f_3(c_1, ..., c_n) \oplus \\ \oplus (f_1(c_1, ..., c_n) \oplus f_2(c_1, ..., c_n)) c_{n+1} \oplus \\ \oplus a c_{n+1} \oplus < \alpha, (c_1, ..., c_n) > \oplus x_{n+2}.\end{aligned}$$

Следовательно,

$$\|(M_{\alpha,a,b})_{1\ldots n+1}^{c_1 \ldots c_{n+1}}\| = 1$$

и $S_1 = |C'| = 2^n$.

2. $(c_1, \ldots, c_{n+1}) \in C''$. В этом случае

$$\begin{aligned}(M_{\alpha,a,b})_{1\ldots n+1}^{c_1 \ldots c_{n+1}} = \\ = f_1(c_1, \ldots, c_n)(1 \oplus a \oplus b) \oplus f_2(c_1, \ldots, c_n) b \oplus \\ \oplus f_3(c_1, \ldots, c_n) a \oplus ab.\end{aligned}$$

Таким образом, $(M_{\alpha,a,b})_{1\ldots n+1}^{c_1 \ldots c_{n+1}}$ в этом случае есть константа. Поэтому сумма

$$\sum_{(c_1, \ldots, c_{n+1}) \in C''} \|(M_{\alpha,a,b})_{1\ldots n+1}^{c_1 \ldots c_{n+1}}\|$$

будет равна весу функции

$$\begin{aligned}F(x_1, \ldots, x_n, x_{n+2}) = \\ = f_1(x_1, ..., x_n)(1 \oplus a \oplus b) \oplus f_2(x_1, ..., x_n) b \oplus \\ \oplus f_3(x_1, ..., x_n) a \oplus ab.\end{aligned}$$

Заметим, что от переменной x_{n+2} функция F зависит несущественно.

Подсчитаем вес F в различных случаях, в зависимости от значений a и b, принимая во внимание то, что f_1, f_2, f_3 и $f_1 \oplus f_2 \oplus f_3$ — бент-функции от n переменных:
$a = b = 0$. $\|F\| = 2\|f_1\| = 2(2^{n-1} \pm 2^{k-1}) = 2^n \pm 2^k$;
$a = 0,\ b = 1$. $\|F\| = 2\|f_2\| = 2(2^{n-1} \pm 2^{k-1}) = 2^n \pm 2^k$;
$a = 1,\ b = 0$. $\|F\| = 2\|f_3\| = 2(2^{n-1} \pm 2^{k-1}) = 2^n \pm 2^k$;
$a = 0,\ b = 1$. $\|F\| = 2\|f_1 \oplus f_2 \oplus f_3\| = 2(2^{n-1} \pm 2^{k-1}) = 2^n \pm 2^k$.

Таким образом, в любом случае,

$$\|M_{\alpha,a,b}\| = 2^n + 2^n \pm 2^k = 2^{n+1} \pm 2^k = 2^{(n+2)-1} \pm 2^{\frac{n+2}{2}-1},$$

и $f(x_1, \ldots, x_n, x_{n+1}, x_{n+2})$ — бент-функция. □

5.4. БЕНТ-ОТОБРАЖЕНИЯ

Понятие бент-функции можно обобщить на более широкие классы дискретных функций, в частности на отображения из Ω_2^n в Ω_2^m. В данном параграфе мы будем рассматривать эти множества как векторные пространства размерностей n и m над полем из двух элементов, которые будем обозначать соответственно V_n и V_m.

Введем другие необходимые обозначения. Назовем *характеристической функцией* отображения F и будем обозначать через I_F следующую булеву функцию от $m + n$ переменных

$$I_F(\mathbf{x}, \mathbf{y}) = \begin{cases} 1, \text{ если } \mathbf{y} = F(\mathbf{x}) \\ 0, \text{ в противном случае} \end{cases}$$

$$\mathbf{x} = (x_1, \ldots x_n), \quad \mathbf{y} = (y_1, \ldots, y_m).$$

Для описания свойств отображения F важным является изучение множеств вида

$$S_F(\alpha, \beta) = \{\mathbf{x} \in V_n | F(\mathbf{x} \oplus \alpha) \oplus F(\mathbf{x}) = \beta\}$$

для $\alpha \in V_n$, $\beta \in V_m$. Мощность множества $S_F(\alpha, \beta)$ будем обозначать через $s_F(\alpha, \beta)$. Максимальную из мощностей таких множеств будем обозначать через $\mu(F)$:

$$\mu(F) = \max_{\alpha \in V_n \setminus \theta} \max_{\beta \in V_m} s_F(\alpha, \beta).$$

Лемма 5.15. *Для любых* $(\alpha, \beta) \in V_n \times V_m$ *выполнено соотношение*

$$s_F(\alpha, \beta) = \sum_{\mathbf{x} \in V_n} \sum_{\mathbf{y} \in V_m} I_F(\mathbf{x}, \mathbf{y}) I_F(\mathbf{x} \oplus \alpha, \mathbf{y} \oplus \beta).$$

□ Действительно, имеем

$$\sum_{\mathbf{x} \in V_n} \sum_{\mathbf{y} \in V_m} I_F(\mathbf{x}, \mathbf{y}) I(\mathbf{x} \oplus \alpha, \mathbf{y} \oplus \beta) =$$

$$= \sum_{\mathbf{x} \in V_n} I(\mathbf{x} \oplus \alpha, F(\mathbf{x}) \oplus \beta) =$$

$$= |\{\mathbf{x} \in V_n | F(\mathbf{x} \oplus \alpha) = F(\mathbf{x}) \oplus \beta\}| = s_F(\alpha, \beta). \square$$

Теорема 5.16. *Для любого отображения* $F : V_n \longrightarrow V_m$ *выполнено неравенство*

$$\mu(F) \geq 2^{n-m}.$$

□ Пусть α — фиксированный вектор из $V_n \setminus \{\theta\}$. Тогда

$$\sum_{\beta \in V_m} s_F(\alpha, \beta) = 2^n.$$

В этой сумме количество слагаемых равно 2^m и они неотрицательные. Следовательно, $\max_{\beta \in V_m} s_F(\alpha, \beta) \geq 2^{n-m}$ для любого $\alpha \in V_n \setminus \{\theta\}$ и, значит, $\mu(F) \geq 2^{n-m}$. □

Дадим теперь определение бент-отображения.

Определение 5.5. *Отображение* $F : V_n \longrightarrow V_m$ *называется бент-отображением, если для любого* $\beta \in V_m$, $\beta \neq \theta$ *булева функция* $< \beta, F(\mathbf{x}) >$ *является бент-функцией.*

Опишем некоторые свойства бент-отображений.

Лемма 5.17. *1. Пусть* $F : V_n \longrightarrow V_m$ *— бент-отображение. Тогда для любых невырожденных матриц* A_1, A_2 *порядка* n *и* m *соответственно и любых векторов* $\gamma_1 \in V_n$, $\gamma \in V_m$ *отображение* G, $G(\mathbf{x}) = A_2 \cdot F(A_1\mathbf{x} \oplus \gamma_1) \oplus \gamma_2$, *является бент-отображением.*

2. Отображение $F : V_n \longrightarrow V_m$ является бент-отображением тогда и только тогда, когда для коэффициентов Фурье функции I_F справедливы равенства:

$$C_{\alpha,\beta}^{I_F} = \pm 2^{-m-\frac{n}{2}},\ \alpha \in V_n,\ \beta \in V_m \setminus \{\theta\}.$$

□ 1. Обозначим $f_\omega(\mathbf{x}) =< \omega, F(\mathbf{x}) >$, $\omega \in V_m$. По условию, f_ω — бент-функция для любого ω. Тогда согласно теореме 5.8 $f_\omega(A_1\mathbf{x} \oplus \gamma_1)$ — также бент-функция. При этом

$$\begin{aligned} < \beta, G(\mathbf{x}) > &=< \beta, A_2 \cdot F(A_1\mathbf{x} \oplus \gamma_1) \oplus \gamma_2 >= \\ &=< A_2^T\beta, F(A_1\mathbf{x} \oplus \gamma_1) > \oplus < \beta, \gamma_2 >= \\ &= f_{A_2^T\beta}(A_1\mathbf{x} \oplus \gamma_1) \oplus < \beta, \gamma_2 > . \end{aligned}$$

Поскольку $< \beta, \gamma_2 > \in \{0, 1\}$, то $< \beta, G(\mathbf{x}) >$ — бент-функция для любого $\beta \in V_m$.

2. Справедливы равенства:

$$\begin{aligned} C_{\alpha,\beta}^{I_F} &= \frac{1}{2^{m+n}} \sum_{\mathbf{x}\in V_n} \sum_{\mathbf{y}\in V_m} I_F(\mathbf{x}, \mathbf{y})(-1)^{<\mathbf{x},\alpha>\oplus<\mathbf{y},\beta>} = \\ &= \frac{1}{2^{m+n}} \sum_{\mathbf{x}\in V_n} (-1)^{<\mathbf{x},\alpha>\oplus<F(\mathbf{x}),\beta>} = \\ &= \frac{1}{2^{m+n}} \sum_{\mathbf{x}\in V_n} (-1)^{<\mathbf{x},\alpha>\oplus f_\beta(\mathbf{x})} = \frac{1}{2^m} W_\alpha^{f_\beta}. \end{aligned}$$

По определению F — бент-отображение, если f_β — бент-функция для любого $\beta \neq \theta$. Это выполняется тогда и только тогда, когда $W_\alpha^{f_\beta} = \pm\frac{1}{2^{n/2}}$ для любого $\alpha \in V_n$, т. е. когда $C_{\alpha,\beta}^{I_F} = \pm\frac{1}{2^{m+\frac{n}{2}}}$. □

Определение 5.6. *Отображение $F : V_n \longrightarrow V_m$ называется совершенно нелинейным, если $\mu(F) = 2^{n-m}$. Заметим, что последнее равенство возможно лишь при $n \geq m$.*

Теорема 5.18. *Отображение $F : V_n \longrightarrow V_m$ является совершенно нелинейным тогда и только тогда, когда F — бент-отображение.*

Эта теорема является обобщением критерия Ротхауза (5.11).

□ Пусть F — совершенно нелинейное отображение. Тогда $\mu(F) = 2^{n-m}$ и для любых $\alpha \in V_n \setminus \{\theta\}$, $\beta \in V_m$ имеем $s_F(\alpha, \beta) = 2^{n-m}$ (см. теорему 5.16). Кроме того, $s_F(\theta, \theta) = 2^n$ и для $\beta \neq \theta$ выполняется $s_F(\theta, \beta) = 0$. Следовательно,

$$\left(C_{\alpha,\beta}^{I_F}\right)^2 = \frac{1}{2^{2(m+n)}} \sum_{\mathbf{x}\in V_n} \sum_{\mathbf{y}\in V_m} \sum_{\mathbf{x}'\in V_n} \sum_{\mathbf{y}'\in V_m} \Big[I_F(\mathbf{x}, \mathbf{y}) \cdot$$

$$\cdot\, I_F(\mathbf{x}', \mathbf{y}')(-1)^{<\mathbf{x}\oplus\mathbf{x}', \alpha>\oplus<\mathbf{y}\oplus\mathbf{y}', \beta>} \Big] =$$

$$= \frac{1}{2^{2(m+n)}} \sum_{\mathbf{x}\in V_n} \sum_{\mathbf{x}'\in V_n} (-1)^{<\mathbf{x}\oplus\mathbf{x}', \alpha>\oplus<F(\mathbf{x})\oplus F(\mathbf{x}'), \beta>} =$$

$$= \frac{1}{2^{2(m+n)}} \sum_{\mathbf{x}\in V_n} \sum_{\mathbf{z}=\mathbf{x}\oplus\mathbf{x}'\in V_n} (-1)^{<\mathbf{z}, \alpha>\oplus<F(\mathbf{x})\oplus F(\mathbf{x}\oplus\mathbf{z}), \beta>} =$$

$$= \frac{1}{2^{2(m+n)}} \sum_{\mathbf{z}\in V_n} (-1)^{<\mathbf{z}, \alpha>} \sum_{\mathbf{x}\in V_n} (-1)^{<F(\mathbf{x})\oplus F(\mathbf{x}\oplus\mathbf{z}), \beta>} =$$

$$= \frac{1}{2^{2(m+n)}} \sum_{\mathbf{z}\in V_n} (-1)^{<\mathbf{z}, \alpha>} \sum_{\gamma\in V_m} s_F(\mathbf{z}, \gamma)(-1)^{<\gamma, \beta>} =$$

$$= \frac{1}{2^{2(m+n)}} \Bigg[\sum_{\gamma\in V_m} s_F(\theta, \gamma)(-1)^{<\gamma, \beta>} +$$

$$+ \sum_{\mathbf{z}\in V_n\setminus\{\theta\}} (-1)^{<\mathbf{z}, \alpha>} \sum_{\gamma\in V_m} 2^{n-m}(-1)^{<\gamma, \beta>} \Bigg] =$$

$$= \frac{1}{2^{2(m+n)}} \Bigg[s_F(\theta, \theta)(-1)^{<\theta, \beta>} +$$

$$+ \sum_{\mathbf{z}\in V_n\setminus\{\theta\}} (-1)^{<\mathbf{z}, \alpha>} \sum_{\gamma\in V_m} 2^{n-m}(-1)^{<\gamma, \beta>} \Bigg].$$

Значение последнего выражения подсчитаем в различных случаях:

а) $\beta \neq \theta$. В этом случае $\sum_{\gamma \in V_m} 2^{n-m}(-1)^{<\gamma,\beta>} = 0$ и

$$\left(C_{\alpha,\beta}^{I_F}\right)^2 = \frac{1}{2^{2(m+n)}} \cdot 2^n = \frac{1}{2^{2m+n}};$$

б) $\beta = \theta$ и $\alpha \neq \theta$. В этом случае $\sum_{\gamma \in V_m} 2^{n-m}(-1)^{<\gamma,\beta>} = 2^n$ и

$$\left(C_{\alpha,\beta}^{I_F}\right)^2 = \frac{1}{2^{2(m+n)}} \left[2^n + 2^n \sum_{\mathbf{z} \in V_n \setminus \{\theta\}} (-1)^{<\mathbf{z},\alpha>}\right] =$$
$$= \frac{1}{2^{2(m+n)}} \cdot 2^n \sum_{\mathbf{z} \in V_n} (-1)^{<\mathbf{z},\alpha>} = 0;$$

в) $\beta = \alpha = \theta$. В этом случае $\sum_{\mathbf{z} \in V_n} (-1)^{<\mathbf{z},\alpha>} = 2^n$ и

$$\left(C_{\alpha,\beta}^{I_F}\right)^2 = \frac{1}{2^{2m}}.$$

То, что F — бент-отображение следует теперь из *а)* и леммы 5.17.

Докажем обратное. Если F — бент-отображение, то $C_{\alpha,\beta}^{I_F} = \pm 2^{-m-\frac{n}{2}}$ для всех $\alpha \in V_n,\ \beta \in V_m \setminus \{\theta\}$. По лемме 5.15

$$s_F(\alpha,\beta) = \sum_{\mathbf{x} \in V_n} \sum_{\mathbf{y} \in V_m} I_F(\mathbf{x},\mathbf{y}) I_F(\mathbf{x} \oplus \alpha, \mathbf{y} \oplus \beta).$$

Представим функцию $I_F(\mathbf{x},\mathbf{y})$ ее разложением в ряд Фурье. Справедливы равенства:

$$s_F(\alpha,\beta)=\sum_{\mathbf{x}\in V_n}\sum_{\mathbf{y}\in V_m}\sum_{(\gamma_1,\omega_1)\in V_n\times V_m}\sum_{(\gamma_2,\omega_2)\in V_n\times V_m}\Bigg[C^{I_F}_{\gamma_1,\omega_1}\times$$

$$\times\, C^{I_F}_{\gamma_2,\omega_2}(-1)^{<x,\gamma_1>\oplus<y,\omega_1>}(-1)^{<x\oplus\alpha,\gamma_2>\oplus<y\oplus\beta,\omega_2>}\Bigg]=$$

$$=\sum_{(\gamma_1,\omega_1)\in V_n\times V_m}\sum_{(\gamma_2,\omega_2)\in V_n\times V_m}\Bigg[C^{I_F}_{\gamma_1,\omega_1}C^{I_F}_{\gamma_2,\omega_2}\times$$

$$\times\,(-1)^{<\alpha,\gamma_2>\oplus<\beta,\omega_2>}\Bigg]\sum_{\mathbf{x}\in V_n}(-1)^{<x,\gamma_1\oplus\gamma_2>}\sum_{\mathbf{y}\in V_m}(-1)^{<y,\omega_1\oplus\omega_2>}=$$

$$=2^{m+n}\sum_{(\gamma_1,\omega_1)\in V_n\times V_m}\sum_{(\gamma_2,\omega_2)\in V_n\times V_m}\Bigg[C^{I_F}_{\gamma_1,\omega_1}\cdot$$

$$\cdot\, C^{I_F}_{\gamma_2,\omega_2}(-1)^{<\alpha,\gamma_2>\oplus<\beta,\omega_2>}\delta_{\gamma_1,\gamma_2}\delta_{\omega_1,\omega_2}\Bigg],$$

где δ — символ Кронекера. После упрощения получаем

$$s_F(\alpha,\beta)=$$

$$=2^{m+n}\sum_{(\gamma_1,\omega_1)\in V_n\times V_m}\left(C^{I_F}_{\gamma_1,\omega_1}\right)^2(-1)^{<\alpha,\gamma_1>}(-1)^{<\beta,\omega_1>}=$$

$$=2^{m+n}\Bigg[\sum_{(\gamma_1,\omega_1)\in V_n\times(V_m\setminus\{\theta\})}\left(C^{I_F}_{\gamma_1,\omega_1}\right)^2(-1)^{<\alpha,\gamma_1>\oplus<\beta,\omega_1>}+$$

$$+\sum_{\gamma_1\in V_n}\left(C^{I_F}_{\gamma_1,\theta}\right)^2(-1)^{<\alpha,\gamma_1>}\Bigg].$$

Поскольку F — бент-отображение, то при $\omega_1\neq\theta$ выполняется равенство $\left(C^{I_F}_{\gamma_1,\omega_1}\right)^2=\frac{1}{2^{2m+n}}$. Для $\omega_1=\theta$ имеем

$$C^{I_F}_{\gamma_1,\theta}=\frac{1}{2^{m+n}}\sum_{(\mathbf{x},\mathbf{y})\in V_n\times V_m}I_F(\mathbf{x},\mathbf{y})(-1)^{<\gamma_1,\mathbf{x}>}=$$

$$=\frac{1}{2^{m+n}}\sum_{(\mathbf{x})\in V_n}(-1)^{<\gamma_1,\mathbf{x}>}=\frac{1}{2^m}\delta_{\gamma_1,\theta}.$$

Таким образом,

$$s_F(\alpha,\beta) =$$
$$= 2^{m+n}\left[\frac{1}{2^{2m+n}}\sum_{(\gamma_1,\omega_1)\in V_n\times(V_m\setminus\{\theta\})}(-1)^{<\alpha,\gamma_1>\oplus<\beta,\omega_1>}+\right.$$
$$\left.+\sum_{\gamma_1\in V_n}(C^{I_F}_{\gamma_1,\theta})^2(-1)^{<\alpha,\gamma_1>}\right] =$$
$$= 2^{m+n}\left[0+\frac{1}{2^{2m}}\right] = 2^{n-m}$$

и отображение F — совершено нелинейное. □

Теорема 5.19. *Если $F: V_n \longrightarrow V_m$ является бент-отображением, то n — четное и $n \geq 2m$.*

□ Пусть F — бент-отображение. Согласно лемме 5.17 для любого $\beta \in V_m$, $\beta \neq \theta$ имеем $C^{I_F}_{\alpha,\beta} = \pm 2^{-m-\frac{n}{2}}$. Следовательно n — четное число. Введем обозначение

$$N_0 = |\{\beta \in V_m \setminus \{\theta\} | C^{I_F}_{\theta,\beta} = 2^{-m-\frac{n}{2}}\}|.$$

Тогда число $D = 2^{m+\frac{n}{2}}\sum_{\beta\in V_m,\beta\neq\theta} C^{I_F}_{\theta,\beta} = 2N_0 - 2^m + 1$ — нечетное. Кроме того,

$$\sum_{\beta\in V_m,\beta\neq\theta} C^{I_F}_{\theta,\beta} = \sum_{\beta\in V_m} C^{I_F}_{\theta,\beta} - C^{I_F}_{\theta,\theta} =$$
$$= \sum_{\beta\in V_m} 2^{-m-n}\sum_{\mathbf{x}\in\mathbf{V_n}}(-1)^{<\beta,F(\mathbf{x})>} - 2^{-m} =$$
$$= 2^{-m-n}\sum_{\mathbf{x}\in\mathbf{V_n}}\sum_{\beta\in V_m}(-1)^{<\beta,F(\mathbf{x})>} - 2^{-m} = 2^{-n}N_1 - 2^{-m},$$

где $N_1 = |\{\mathbf{x} \in V_n | F(\mathbf{x}) = \theta\}|$. Тогда

$$D = 2^{m+\frac{n}{2}}(2^{-n}N_1 - 2^{-m}) = 2^{m-\frac{n}{2}}N_1 - 2^{\frac{n}{2}}$$
$$\text{и } N_1 = 2^{\frac{n}{2}-m}(D + 2^{\frac{n}{2}}).$$

Поскольку N_1 — целое, D — нечетное, то $2^{\frac{n}{2}-m}$ должно быть целым. Отсюда $n \geq 2m$.□

Аналогично тому, как в утверждении 5.13 был построен класс бент-функций, можно построить и класс бент-отображений.

Поскольку поле $GF(2^n)$ является векторным пространством над полем $GF(2)$, то при произвольном выборе базиса этого пространства задается однозначное соответствие между элементами поля $GF(2^n)$ и множества Ω_2^n. Справедливо следующее

Утверждение 5.20. *Пусть G — произвольное преобразование поля $GF(2^n)$. Тогда отображение $F : GF(2^n) \times \times GF(2^n) \to GF(2^n)$, заданное равенством*

$$F(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y} + G(\mathbf{y}),$$

где «$\cdot$» и «$+$» — умножение и сложение в поле $GF(2^n)$, является бент-отображением. □

Глава 6

ДЕКОМПОЗИЦИЯ БУЛЕВЫХ ФУНКЦИЙ

Еще в 1949 г. К. Шеннон заметил, что из всех 2^{2^n} булевых функций используется на практике при больших n, как правило, лишь малая часть, а именно те функции, которые допускают простую техническую реализацию и конструктивно выполняются в виде каким-либо образом соединенных блоков. При этом нередко обеспечивается условие независимости входных переменных для отдельных блоков. В настоящее время, несмотря на то что функции, в основном, реализуются программно, тем не менее часто используются функции, также составленные из независимых компонентов. В связи с этим возникает проблема описания суперпозиций функций от непересекающихся наборов переменных.

В этой главе в наших рассуждениях вместо упорядоченных наборов переменных, от которых зависят функции, будет удобно использовать просто множества переменных, которые мы будем обозначать большими латинскими буквами $X_1, X_2, Y_1, \ldots$.

Определение 6.1 (Шеннон). *Булева функция f от множества переменных X называется функционально разделимой, если ее можно представить формулой*

$$f(X) = F(h_1(X_1), \ldots, h_k(X_k)), \tag{6.1}$$

где $X = X_1 \cup \ldots \cup X_k$, $X_i \cap X_j = \emptyset$ при $i \neq j$ (при выполнении таких условий будем говорить, что имеет место разбиение множества X, и писать $X = X_1 \coprod \ldots \coprod X_k$), $X_i \neq \emptyset$, а $F, h_1, \ldots, h_k$ — некоторые булевы функции.

Нам будет удобно использовать и другое понятие:

Определение 6.2. *Говорят, что функция $f(X)$ допускает простую декомпозицию, если существуют такие функции F и h, что*

$$f(X) = F(h(X_1), X_2),\ X_1, X_2 \neq \emptyset,\ X_1 \cap X_2 = \emptyset,\ X_1 \cup X_2 = X.$$

В случае, когда $|X_1| > 1$ и функция f существенно зависит от всех переменных, простую декомпозицию называют нетривиальной.

На самом деле, приведенные определения эквивалентны: если $f(X) = F(h_1(X_1), \ldots, h_k(X_k))$, то рассматриваем функцию

$$\Psi(y, X_2, \ldots, X_k) = F(y, h_2(X_2) \ldots, h_k(X_k))$$

и для нее имеем $f(X) = \Psi(h_1(X_1), X_2, \ldots, X_k)$. Если же $f(X) = F(h(X_1), X_2)$, то полагаем $h_1 = h$, $h_i = x_{j_i}$, где $i \in \overline{2,k}$, а $\{x_{j_2}, \ldots, x_{j_k}\} = X_2$, и получаем равенство (6.1).

Приведем сложившуюся в классификации декомпозиций терминологию:

1) кратная декомпозиция:

$$f(X) = F(h_1(X_1), \ldots, h_k(X_k), X_{k+1});$$

2) итеративная декомпозиция:

$$f(X) = F(h_k(h_{k-1}(\cdots(h_1(X_1), X_2), \cdots), X_k), X_{k+1});$$

3) сложная (комплексная) декомпозиция — комбинация декомпозиций 1го и 2го типа.

Ниже будет показано, что всякая сложная декомпозиция полностью определяется некоторым набором простых декомпозиций. Поэтому задача описания всех декомпозиций сводится к задаче нахождения всех простых декомпозиций данной функции. Это сведение впервые было осуществлено Р. Ашенхерстом в 1959 г.

Таблица 2.1
Представление булевой функции

			x_1	0	$\dots$	a_1	$\dots$	1
			$\vdots$	$\vdots$		$\vdots$		$\vdots$
			x_k	0	$\dots$	a_k	$\dots$	1
x_{k+1}	$\dots$	x_n						
0	$\dots$	0				$\vdots$		
$\vdots$		$\vdots$				$\vdots$		
a_{k+1}	$\dots$	a_n		$\dots$	$\dots$	$f(a_1,\dots,a_n)$		
$\vdots$		$\vdots$						
1	$\dots$	1						

6.1. ПРОСТАЯ ДЕКОМПОЗИЦИЯ

В этом параграфе будут доказаны два критерия наличия у функции простой декомпозиции. Для доказательства первого из них мы воспользуемся следующим способом задания булевой функции $f(x_1,\dots,x_n)$. Рассмотрим таблицу размера $2^{n-k}\times 2^k$, столбцы которой занумерованы наборами из Ω_2^k, строки — наборами из Ω_2^{n-k}, а на пересечении столбца $(a_1,\dots,a_k)$ и строки $(a_{k+1},\dots,a_n)$ стоит значение $f(a_1,\dots,a_n)$ (табл. 2.5).

Несложно заметить, что столбцы этой таблицы являются таблицами подфункций $f_{1\dots k}^{a_1\dots a_k}$, а строки — таблицами подфункций $f_{k+1\dots n}^{a_{k+1}\dots a_n}$.

Теорема 6.1. *Равносильны утверждения:*

1) функция $f(x_1,\dots,x_n)$ допускает простую декомпозицию вида

$$f(x_1,\dots,x_n)=F(h(x_1,\dots,x_k),x_{k+1},\dots,x_n);$$

2) в таблице 2.1 существует такая строка, что остальные либо совпадают с ней, либо являются ее

отрицанием, либо состоят целиком из констант — нулей или единиц;

3) в таблице 2.1 есть не более двух различных столбцов.

□ 1 ⇒ 2. В строках таблицы 2.1 в данном случае стоят таблицы функций

$$f_{k+1\ldots n}^{a_{k+1}\ldots a_n}(x_1,\ldots,x_n) = F(h(x_1,\ldots,x_k),a_{k+1},\ldots,a_n).$$

Рассмотрим функцию $g(y)$, задаваемую формулой $F(y,a_{k+1},\ldots,a_n)$. Имеем четыре варианта:

1. $g(y)\equiv 0$. В этом случае строка, соответствующая набору $a_{k+1},\ldots,a_n$, состоит из одних нулей.
2. $g(y)\equiv 1$. В этом случае строка, соответствующая набору $a_{k+1},\ldots,a_n$, состоит из одних единиц.
3. $g(y)\equiv y$. В этом случае строка, соответствующая набору $a_{k+1},\ldots,a_n$, совпадает с таблицей функции $h(x_1,\ldots,x_k)$.
4. $g(y)\equiv \overline{y}$. В этом случае строка, соответствующая набору $a_{k+1},\ldots,a_n$, совпадает с таблицей функции $\overline{h}(x_1,\ldots,x_k)$.

Отсюда очевидным образом следует утверждение 2.

2 ⇒ 3. Предположим противное: пусть в таблице 2.1 имеется 3 попарно различных столбца, которые соответствуют наборам $\alpha_1,\alpha_2,\alpha_3\in\Omega_2^k$. Допустим, что столбцы α_1 и α_2 различаются в строке, соответствующей набору $\beta=(b_{k+1},\ldots,b_n)\in\in\Omega_2^{n-k}$. Тогда $f(\alpha_1,\beta)=c$, а $f(\alpha_2,\beta)=\overline{c}$, $c\in\Omega_2$. Не ограничивая общности, будем считать, что в столбце α_3 также стоит значение c, т. е. $f(\alpha_1,\beta)=f(\alpha_3,\beta)$. Строка β не является строкой констант, значит, любая другая строка таблицы либо совпадает с ней, либо является ее отрицанием, либо является строкой констант. В любом случае, так как в этой строке в столбцах α_1 и α_3 стоят одинаковые значения, то и во всех остальных строках значения, стоящие в этих столбцах, будут совпадать. Следовательно, столбцы α_1 и α_3 совпадают, что противоречит предположению.

3 ⇒ 1. Пусть есть только два вида столбцов. Тогда есть только 2 различные подфункции, соответствующие фиксациям переменных $x_1,\ldots,x_k$. Обозначим их $F_0(x_{k+1},\ldots,x_n)$ и $F_1(x_{k+1},\ldots,x_n)$. Определим теперь функцию $h(x_1,\ldots,x_k)$ ра-

венствами:

$$h(\alpha) = \begin{cases} 0, & f_{1\ldots k}^{a_1\ldots a_k} = F_0 \\ 1, & f_{1\ldots k}^{a_1\ldots a_k} = F_1 \end{cases}.$$

Пусть функция $F(y, x_{k+1}, \ldots, x_n)$ задана формулой

$$F(y, x_{k+1}, \ldots, x_n) = \overline{y}F_0(x_{k+1}, \ldots, x_n) \vee yF_1(x_{k+1}, \ldots, x_n).$$

Теперь несложно убедиться, что

$$f(x_1, \ldots, x_n) = F(h(x_1, \ldots, x_k), x_{k+1}, \ldots, x_n). \ \square$$

Читателю предлагается самостоятельно доказать еще один несложный критерий наличия у функции простой декомпозиции:

Теорема 6.2. *Функция $f(x_1, \ldots, x_n)$ допускает простую декомпозицию вида*

$$f(x_1, \ldots, x_n) = F(h(x_1, \ldots, x_k), x_{k+1}, \ldots, x_n)$$

тогда и только тогда, когда ее многочлен Жегалкина может быть записан в виде

$$f(x_1, ..., x_n) = h(x_1, ..., x_k)F_1(x_{k+1}, ..., x_n) \oplus \\ \oplus F_2(x_{k+1}, ..., x_n). \ \square$$

6.2. РАЗЛОЖИМЫЕ ФУНКЦИИ

Определение 6.3. *Функция $f(X)$ называется $*$-разложимой, где $* \in \{\oplus, \vee, \wedge\}$, если существует разбиение множества X: $X = X_1 \cup X_2$, $X_1 \cap X_2 = \emptyset$, $X_i \neq \emptyset$, $i = 1, 2$ и функции h_1, h_2, не равные константе, такие что $f(X) = h_1(X_1) * h_2(X_2)$.*

Теорема 6.3. *Пусть $* \in \{\oplus, \vee, \wedge\}$ и $f(X) = h_1(X_1) * \cdots * h_k(X_k)$, $X = X_1 \cup \ldots \cup X_k$ — разбиение X, причем k — максимальное натуральное число с таким свойством.*

Тогда f не является разложимой относительно двух других операций из множества $\{\oplus, \vee, \wedge\}$, и в случае $* \in \{\vee, \wedge\}$ такое разложение единственно с точностью до

перестановки членов. В случае $* = \oplus$ *разложение имеет вид*

$$f(X) = (h_1(X_1) \oplus a_1) \oplus \ldots \oplus (h_k(X_k) \oplus a_k),$$

$$a_i \in \Omega_2, \sum_{i=1}^{k} a_i \equiv 0(\mathrm{mod}2),$$

и однозначно с точностью до перестановки h_i *и выбора* a_i.

□ Докажем теорему в случае, когда $* = \oplus$. Рассмотрим неориентированный граф $\Gamma_f(\oplus) = (X, E)$ с множеством вершин X, причем ребро (x_i, x_j) содержится во множестве ребер E тогда и только тогда, когда найдется член многочлена Жегалкина функции f, в который одновременно входят переменные x_i и x_j. Поскольку многочлен Жегалкина для f определен однозначно, то и граф $\Gamma_f(\oplus)$ однозначно задается функцией f.

Пусть (X_i, E_i) — компонента связности графа $\Gamma_f(\oplus)$, $i \in \in \overline{1,k}$. Тогда функция f, очевидно, представима в виде суммы:

$$f(X) = h_1(X_1) \oplus \ldots \oplus h_k(X_k),$$

где h_i — функции, соответствующие членам многочлена Жегалкина, содержащим переменные из компоненты X_i. Предположим, что имеется еще одно разложение

$$f(X) = \tilde{h}_1(Y_1) \oplus \ldots \oplus \tilde{h}_m(Y_m).$$

Очевидно, что все одночлены, переменные которых составляют одну компоненту X_i, входят в многочлен Жегалкина одной из функций $\tilde{h}_j(Y_j)$. Таким образом, каждое множество Y_j есть объединение некоторых множеств X_i. Следовательно,

$$\tilde{h}_j(Y_j) = h_{j_1}(X_{j_1}) \oplus \ldots \oplus h_{j_s}(X_{j_s}).$$

Поскольку в многочленах $h_{j_t}(X_{j_t})$ не может быть одинаковых членов, кроме свободных, то суммы свободных членов в различных разложениях должны быть равны. Утверждение теоремы следует теперь из максимальности k.

В случаях, когда $* \in \{\vee, \wedge\}$, рассуждения аналогичны, но вместо многочлена Жегалкина надо использовать представление функции сокращенной ДНФ (также однозначное) и граф $\Gamma_f(\vee)$.

Покажем, что в случае разложимости функции по одной из операций, она неразложима по двум другим.

Действительно, предположим

$$f(X) = f_1(X_1) \oplus f_2(X_2) = h_1(X_1) \cdot h_2(X_2).$$

Тогда, очевидно, граф $\Gamma_{h_1 \cdot h_2}(\oplus)$ связен, а $\Gamma_{f_1 \oplus f_2}(\oplus)$ — нет. Противоречие. Если

$$f(X) = f_1(X_1) \wedge f_2(X_2) = h_1(X_1) \vee h_2(X_2),$$

то граф $\Gamma_{f_1 \cdot f_2}(\vee)$ связен, а $\Gamma_{h_1 \vee h_2}(\vee)$ — нет. Если

$$f(X) = f_1(X_1) \vee f_2(X_2) = h_1(X_1) \oplus h_2(X_2),$$

то $\overline{f}(X) = \overline{f}_1(X_1) \wedge \overline{f}_2(X_2) = h_1(X_1) \oplus \overline{h}_2(X_2)$, и противоречие получается так же, как и в первом случае. □

6.3. СЛОЖНЫЕ ДЕКОМПОЗИЦИИ

В этом параграфе мы проведем классификацию всех возможных декомпозиций. Вначале без доказательства приведем следующее утверждение (доказательство не слишком сложно, но весьма объемно):

Теорема 6.4. *Пусть выполняются тождества:*

$$f(X) \equiv \Psi_1(h_1(X_0, X_1), X_2, X_3) \equiv \Psi_2(h_2(X_0, X_2), X_1, X_3),$$

где $X = X_0 \bigsqcup X_1 \bigsqcup X_2 \bigsqcup X_3$ *— разбиение, причем* $X_0, X_1, X_2 \neq \emptyset$. *Тогда существуют функции* $\Psi, \tilde{h}_0, \tilde{h}_1, \tilde{h}_2$, *такие что*

$$f(X) = \Psi(\tilde{h}_0(X_0) * \tilde{h}_1(X_1) * \tilde{h}_2(X_2), X_3),$$

где $* \in \{\oplus, \wedge, \vee\}$. □

Пусть функция $f(x_1, \ldots, x_n)$ существенно зависит от всех своих переменных. На множестве X введем отношение $\overset{f}{\sim}$ следующим образом: переменные x_i и x_j находятся в отношении $\overset{f}{\sim}$ (пишут: $x_i \overset{f}{\sim} x_j$) тогда и только тогда, когда существует простая декомпозиция $f(X) = \Psi(h(X_1), X_2)$, $X_1 \cap X_2 = \emptyset$, $X_2 \neq \emptyset$, $x_i, x_j \in X_1$.

Утверждение 6.5. *Отношение* $\overset{f}{\sim}$ *является отношением эквивалентности.*

□ Рефлексивность и симметричность данного отношения очевидны. Проверим транзитивность. Пусть $x_i \overset{f}{\sim} x_j$ и $x_j \overset{f}{\sim} x_k$. Тогда

$$f(X) \equiv \Psi_1(h_1(Y_1), Y_2) \equiv \Psi_2(h_2(Z_1), Z_2),$$

$x_i, x_j \in Y_1$, $x_j, x_k \in Z_1$. Если $x_k \in Y_1$ или $x_i \in Z_1$, то все доказано. Пусть это не так. Обозначим $X_0 = Y_1 \cap Z_1$, $X_1 = Y_1 \setminus X_0$, $X_2 = Z_1 \setminus X_0$, $X_3 = X \setminus (X_0 \cup X_1 \cup X_2) = Z_2 \cap Y_2$. При этом множества X_0, X_1, X_2 непусты, так как содержат переменные x_j, x_i, x_k соответственно. Тогда по теореме 6.4 имеем

$$f(X) \equiv \Psi(\tilde{h}_0(X_0) * \tilde{h}_1(X_1) * \tilde{h}_2(X_2), X_3).$$

Если $X_3 \neq 0$, то, обозначая $\tilde{h}_0(X_0) * \tilde{h}_1(X_1) * \tilde{h}_2(X_2) = h(X_0 \cup X_1 \cup X_2)$, получим требуемую декомпозицию. Если же $X_3 = \emptyset$, то

$$f(X) \equiv \Psi(*(\tilde{h}_0(X_0) * \tilde{h}_2(X_2), \tilde{h}_1(X_1))) \equiv \tilde{\Psi}(h'(X_0 \cup X_2), \tilde{h}_1(X_1)),$$

где $h'(X_0 \cup X_2) = \tilde{h}_0(X_0) * \tilde{h}_2(X_2)$, а $\tilde{\Psi}$ — композиция Ψ и $*$. □

Таким образом, множество переменных X разбивается на непересекающиеся классы эквивалентности:

$$X = Z_1 \cup \ldots \cup Z_k.$$

Следствие 1. *Для любого* $i \in \overline{1,k}$ *найдутся функции* Ψ_i *и* h_i, *такие что* $f(X) = \Psi_i(h_i(Z_i), Z_1, \ldots, Z_{i-1}, Z_{i+1}, \ldots, Z_k)$.

□ Докажем следствие для $i = 1$. Пусть $Z_1 = \{x_1, \ldots, x_m\}$. Если $m = 1$, то полагаем $\Psi_1 \equiv f, h_1(x_1) \equiv x_1$ и имеем требуемое представление.

Пусть $m > 1$. Поскольку $x_1 \overset{f}{\sim} x_2$, то существует простая декомпозиция:

$$f(X) = \Psi^{(1)}(h^{(1)}(Y_1), Z_1 \setminus Y_1, Z_2, \ldots, Z_n),$$

где $x_1, x_2 \in Y_1 \subseteq Z_1$. Если $Y_1 = Z_1$, то все доказано. Пусть найдется $x_j \in Z_1 \setminus Y_1$. Так как $x_1 \overset{f}{\sim} x_j$, то найдутся функции $\Psi^{(2)}, h^{(2)}$, такие что

$$\Psi^{(2)}(h^{(2)}(Y_2), Z_1 \setminus Y_2, Z_2, \ldots, Z_n),$$

где $x_1, x_j \in Y_2 \subseteq Z_1$. Полагая $Y_1 \cap Y_2 = X_0$, $Y_1 \setminus Y_2 = X_1$, $Y_2 \setminus Y_1 = X_2$, $X_3 = X \setminus (Y_1 \cup Y_2)$, по теореме 6.4 получим

$$f(X) = \Psi(\tilde{h}_0(X_0) * \tilde{h}_1(X_1) * \tilde{h}_2(X_2), X_3) = \Psi(h(Y_1 \cup Y_2), X_3),$$

где во множество $Y_1 \cup Y_2$ входят переменные x_1, x_2, x_3. Продолжая такой процесс «присоединения переменных», придем к требуемой декомпозиции. □

Прежде чем сформулировать основной результат параграфа, докажем две вспомогательные леммы:

Лемма 6.6 (О кратной декомпозиции). *Пусть функция $f(X)$ допускает две простые декомпозиции:*

$$f(X) = \Psi_1(h_1(X_1), X_2, X_3) = \Psi_2(h_2(X_2), X_1, X_3),$$

где $X = X_1 \bigsqcup X_2 \bigsqcup X_3$ — разбиение, причем $X_1, X_2 \neq \emptyset$. Тогда найдется функция Ψ, такая что

$$f(X) = \Psi(h_1(X_1), h_2(X_2), X_3).$$

□ По теореме 6.2 имеем равенства:

$$f(X) = h_1(X_1)\Psi_{11}(X_2, X_3) \oplus \Psi_{12}(X_2, X_3) = \\ = h_2(X_2)\Psi_{21}(X_1, X_3) \oplus \Psi_{22}(X_1, X_3). \quad (6.2)$$

Так как функция h существенно зависит от всех переменных, то можно зафиксировать все ее переменные, кроме одной (скажем, x_1), так, чтобы полученная подфункция существенно зависела бы от x_1, т. е. была бы равна x_1^a для подходящего a. Выполнив в равенстве (6.2) соответствующую фиксацию, а также заменив x_1^a на z, получим

$$z\Psi_{11}(X_2, X_3) \oplus \Psi_{12}(X_2, X_3) = h_2(X_2)\tilde{\Psi}_{21}(z, X_3) \oplus \tilde{\Psi}_{22}(z, X_3)$$

для некоторых функций $\tilde{\Psi}_{21}(z, X_3)$ и $\tilde{\Psi}_{22}(z, X_3)$. Полагая $z = 0$, получим

$$\Psi_{12}(X_2, X_3) = h_2(X_2)\tilde{\Psi}_{21}(0, X_3) \oplus \tilde{\Psi}_{22}(0, X_3),$$

а при $z = 1$ имеем

$$\Psi_{11}(X_2, X_3) = h_2(X_2)\tilde{\Psi}_{21}(1, X_3) \oplus \tilde{\Psi}_{22}(1, X_3) \oplus \Psi_{12}(X_2, X_3).$$

После подстановки в последнее равенство выражения для $\Psi_{12}(X_2, X_3)$ получим

$$\Psi_{11}(X_2, X_3) = h_2(X_2)(\tilde{\Psi}_{21}(1, X_3) \oplus \Psi_{21}(0, X_3)) \oplus \\ \oplus (\tilde{\Psi}_{22}(1, X_3) \oplus \Psi_{22}(0, X_3)).$$

Тогда из равенства (6.2) следует соотношение

$$f(X) = h_1(X_1)(h_2(X_2)F_1(X_3) \oplus \\ \oplus F_2(X_3)) \oplus h_2(X_2)F_3(X_3) \oplus F_4(X_3)$$

для подходящих F_i, $i \in \overline{1,4}$. Для завершения доказательства остается положить

$$\Psi(x, y, X_3) = x(yF_1(X_3) \oplus F_2(X_3)) \oplus yF_3(X_3) \oplus F_4(X_3). \square$$

Следствие 2. *Пусть множество переменных функции f разбивается на k классов эквивалентности: $X = Z_1 \bigsqcup \ldots \bigsqcup Z_k$. Тогда найдутся функции Φ и h_i, $i \in \overline{1,k}$, такие что*

$$f(X) = \Phi(h_1(Z_1), \ldots, h_k(Z_k)),$$

при этом функция Φ не допускает нетривиальной простой декомпозиции.

$\square$ По лемме 6.6 найдутся такие функции Ψ_{12} и Ψ_{13}, что

$$f(Z_1, \ldots, Z_k) = \Psi_{12}(h_1(Z_1), h_2(Z_2), Z_3, \ldots, Z_k) = \\ = \Psi_{13}(h_1(Z_1), h_3(Z_3), Z_2, Z_4, \ldots, Z_k).$$

Рассмотрим функцию

$$\Psi_1(y, Z_2, \ldots, Z_k) = \Psi_{12}(y, h_2(Z_2), Z_3, \ldots, Z_k) = \\ = \Psi_{13}(y, h_3(Z_3), Z_2, Z_4, \ldots, Z_k).$$

Последнее равенство справедливо в силу того, что h_1 не является константой. Тогда, снова в силу леммы 6.6, найдется функция Ψ_{123}, такая что

$$\Psi_1(y, Z_2, \ldots, Z_k) = \Psi_{123}(y, h_2(Z_2), h_3(Z_3), Z_4, \ldots, Z_k).$$

При этом

$$\Psi_1(h_1(Z_1), Z_2, \ldots, Z_k) = f(Z_1, \ldots, Z_k) = \\ = \Psi_{123}(h_1(Z_1), h_2(Z_2), h_3(Z_3), Z_4, \ldots, Z_k).$$

Таким образом, найдено такое представление, когда уже три класса переменных входят во «внутренние функции». Действуя аналогично, получим представление, когда все классы эквивалентности будут входить во «внутренние функции».

Предположим, что функция Φ допускает нетривиальную простую декомпозицию:

$$\Phi(y_1, \ldots, y_k) = F(G(y_1, \ldots, y_l), y_{l+1}, \ldots, y_k),\ 1 < l < k.$$

Следовательно,

$$f(Z_1, ..., Z_k) = F(G(h_1(Z_1), ..., h_l(Z_l)), h_{l+1}(Z_{l+1}), ..., h_k(Z_k)).$$

Обозначим

$$\tilde{G}(Z_1, \ldots, Z_l) = G(h_1(Z_1), \ldots, h_l(Z_l)) \text{ и} \\ \tilde{F}(y, Z_{l+1}, \ldots, Z_k) = F(y, h_{l+1}(Z_{l+1}), \ldots, h_k(Z_k)).$$

Тогда, очевидно,

$$f(Z_1, \ldots, Z_k) = \tilde{F}(\tilde{G}(Z_1, \ldots, Z_l), Z_{l+1}, \ldots, Z_k),$$

и, следовательно, переменные из множеств $Z_1, \ldots, Z_l$ эквивалентны между собой, что противоречит тому, что классов эквивалентности ровно k. □

Вторая лемма описывает другой тип декомпозиций.

Лемма 6.7 (Об итеративной декомпозиции). *Пусть функция f допускает две декомпозиции вида:*

$$f(X) = \Psi_1(h_1(X_1), X_2, X_3) = \Psi_2(h_2(X_1, X_2), X_3).$$

Тогда существует функция h, такая что

$$f(X) = \Psi_2(h(h_1(X_1), X_2), X_3).$$

Замечание. Из утверждения теоремы следует, что Ψ_1 допускает простую декомпозицию.

□ Так же, как и в лемме 6.7, имеем равенство:

$$f(X) = h_1(X_1)\Psi_{11}(X_2, X_3) \oplus \Psi_{12}(X_2, X_3) = \\ = h_2(X_1, X_2)\Psi_{21}(X_3) \oplus \Psi_{22}(X_3). \quad (6.3)$$

Так же, как и в лемме 6.7, фиксируем все переменные функции $h_1(X_1)$, кроме x_1, и выполняем замену x_1^a на z. В результате имеем

$$z\Psi_{11}(X_2, X_3) \oplus \Psi_{12}(X_2, X_3) = \tilde{h}_2(z, X_2)\Psi_{21}(X_3) \oplus \Psi_{22}(X_3).$$

При $z = 0$ получаем

$$\Psi_{12}(X_2, X_3) = \tilde{h}_2(0, X_2)\Psi_{21}(X_3) \oplus \Psi_{22}(X_3).$$

Используя это равенство и равенство, получающееся при фиксации $z = 1$, будем иметь

$$\Psi_{11}(X_2, X_3) = \tilde{h}_2(1, X_2)\Psi_{21}(X_3) \oplus \tilde{h}_2(0, X_2)\Psi_{21}(X_3).$$

Таким образом,

$$f(X) = \Big(h_1(X_1)(\tilde{h}_2(0, X_2) \oplus \tilde{h}_2(1, X_2)) \oplus \tilde{h}_2(0, X_2)\Big) \cdot \\ \cdot \Psi_{21}(X_3) \oplus \Psi_{22}(X_3) = \Psi_2(h(h_1(X_1), X_2), X_3),$$

где $h(y, X_2) = y(\tilde{h}_2(0, X_2) \oplus \tilde{h}_2(1, X_2)) \oplus \tilde{h}_2(0, X_2)$. □

Теперь мы можем доказать основной результат параграфа.

Теорема 6.8 (Ашенхёрст). *Пусть функция $f(X)$ зависит более чем от двух переменных. Тогда $f(X)$ определяет на множестве X:*

1) не более двух классов эквивалентности переменных тогда и только тогда, когда f $*$-разложима, где $* \in \{\oplus, \vee, \wedge\}$;

2) имеет ровно k классов эквивалентности ($k \geq 3$) тогда и только тогда, когда найдутся функции Ψ и h_i, $i \in \overline{1,s}$, $s \geq 3$, такие что $f(X) = \Psi(h_1(X_1), \ldots, h_s(X_s))$, $X_i \cap X_j = \emptyset$, $i \neq j$. При этом $s = k$ и набор множеств $X_1, \ldots, X_s$ совпадает с набором $Z_1, \ldots, Z_k$ классов эквивалентности с точностью до перестановки.

□ 1) $k \leq 2$. Доказательство в прямую сторону.

Предположим имеется один класс эквивалентности, т. е. все переменные эквивалентны относительно f. Выберем нетривиальную простую декомпозицию $f(X) = \Psi_1(h_1(X_1), X_2)$ так, чтобы множество X_1 имело максимальную мощность среди всех таких декомпозиций. Так как переменные из множеств X_1 и X_2 эквивалентны, то существует еще одна декомпозиция $f(X) = \Psi_2(h_2(Y_1), Y_2)$, где $X_1 \cap Y_1 \neq \emptyset$, $Y_1 \setminus X_1 \neq \emptyset$. Так как X_1 — максимальное, то $|X_1| \geq |Y_1|$, значит, $X_1 \setminus Y_1 \neq 0$ и $|X_1 \cup Y_1| > |X_1|$. По теореме 6.4 функция f представима в виде

$$f(X) = \Psi(h'_1(X_1 \cap Y_1) * h'_2(X_1 \setminus Y_1) * h'_3(Y_1 \setminus X_1), X_2 \cap Y_2).$$

В этом представлении на месте первого аргумента функции Ψ стоит функция от набора переменных $X_1 \cup Y_1$, откуда следует, что $X_2 \cap Y_2 = \emptyset$, иначе имеем противоречие с максимальностью X_1. Тогда Ψ — функция (существенно зависящая!) от одного переменного и следовательно

$$f(X) = h''_1(X_1 \cap Y_1) * h''_2(X_1 \setminus Y_1) * h''_3(Y_1 \setminus X_1)$$

для соответствующих h''_1, h''_2, h''_3. Таким образом, функция f разложима.

Пусть теперь $f(X)$ задает два класса эквивалентности переменных. Тогда по следствию из леммы 6.6 $f(X) = \Psi(h_1(Z_1), h_2(Z_2))$. При этом Ψ существенно зависит от

обеих переменных, поэтому $\Psi(y_1, y_2) = y_1^{a_1} * y_2^{a_2}$, $* \in \{\oplus, \vee, \wedge\}$. Следовательно,

$$f(Z_1, Z_2) = \Psi(h_1(Z_1), h_2(Z_2)) = h_1^{a_1}(Z_1) * h_2^{a_2}(Z_2),$$

и f $*$-разложима.

Доказательство в обратную сторону. Пусть $f(X)$ — $*$-разложима. Возможны два варианта:

а) f разложима на 3 «части», т. е. найдутся f_1, f_2, f_3, такие что $f(X) = f_1(X_1) * f_2(X_2) * f_3(X_3)$. Тогда можно записать:

$$f(X) = (f_1(X_1) * f_2(X_2)) * f_3(X_3) = (f_2(X_2) * f_3(X_3)) * f_1(X_1),$$

откуда следует, что все переменные из множества X эквивалентны относительно функции f, и имеем 1 класс эквивалентности.

б) f разложима только на 2 «части»: $f(X) = f_1(X_1) * * f_2(X_2)$, где f_1 и f_2 неразложимы. Тогда переменные из множеств X_1 и X_2, очевидно, эквивалентны, и число классов не превышает двух (на самом деле, их ровно 2, читателю предлагается показать это самостоятельно).

2) $k \geq 3$. Доказательство в прямую сторону.

Пусть число классов эквивалентности равно k, $k \geq 3$. Тогда по следствию леммы 6.6

$$f(X) = \Phi(h_1(Z_1), \ldots, h_k(Z_k)),$$

и Ψ не допускает нетривиальной простой декомпозиции.

Проведем доказательство в обратную сторону. Пусть

$$f(X) = \Phi(h_1(X_1), \ldots, h_s(X_s)),$$

и Φ не допускает простой декомпозиции. Тогда, очевидно, что все переменные из каждого множества X_j, $j \in \overline{1,s}$, эквивалентны между собой. Поэтому каждый класс Z_i, $i \in \overline{1,k}$, является объединением нескольких множеств X_j. Покажем, что s не может превышать k. Предположим, что это не так. Тогда какой-то из классов эквивалентности представим в виде объединения нескольких множеств X_j, $j \in \overline{1,s}$. Получим

противоречие в случае, когда $Z_1 = X_1 \cup X_2$ (в общем случае рассуждения аналогичны).

По следствию утверждения 6.5 существует простая декомпозиция

$$f(X) = \Psi(C(Z_1), X_3, \ldots, X_s).$$

Тогда по лемме об итеративной декомпозиции (6.7) должны существовать простые декомпозиции:

$$C(Z_1) = B_1(h_1(X_1), X_2),\ C(Z_1) = B_2(X_1, h_2(X_2)).$$

По лемме о кратной декомпозиции (6.6) существуют простые декомпозиции вида

$$\Psi(y, X_3, \ldots, X_s) = D_i(y, X_3, \ldots, h_i(X_i), \ldots, X_s),\ i \in \overline{3, s},$$

при некоторых функциях D_i. Применяя несколько раз лемму о кратной декомпозиции, получаем

$$C_1(Z_1) = F(h_1(X_1), h_2(X_2)),$$
$$\Psi(y, X_3, \ldots, X_s) = F'(y, h_3(X_3), \ldots, h_s(X_s)).$$

Таким образом,

$$\Phi(h_1(X_1), \ldots, h_s(X_s)) = \\ = F'(F(h_1(X_1), h_2(X_2)), h_3(X_3), \ldots, h_s(X_s)),$$

и функция Φ допускает простую декомпозицию, что противоречит условию. $\square$

6.4. ГРУППЫ ИНЕРЦИИ СУПЕРПОЗИЦИИ БУЛЕВЫХ ФУНКЦИЙ В ГРУППАХ Σ_n, S_n, Q_n

Опишем вначале группы инерции суперпозиции функций от непересекающихся множеств переменных в группе Σ_n. Пусть функция $f(X)$ существенно зависит от всех своих переменных и допускает декомпозицию вида:

$$f(X) = \Psi(h_1(X_1), \ldots, h_k(X_k)),\ |X_i| = n_i > 1,\ i \in \overline{1, k}.$$

Теорема 6.9. *Группа инерции функции $f(X)$ в группе Σ_n состоит из тех и только тех преобразований $g(\mathbf{x}) = \mathbf{x} \oplus \alpha$, которые удовлетворяют условию $h_i(X_i \oplus \alpha_i) = h_i(X_i) \oplus \oplus a_i,\ a_i \in \Omega_2, i \in \overline{1,k}$, а преобразование*

$$\overline{g} = \begin{pmatrix} y_1 & \dots & y_k \\ y_1 \oplus a_1 & \dots & y_k \oplus a_k \end{pmatrix}$$

содержится в группе инерции функции Ψ в группе Σ_k. (Под $X_i \oplus \alpha_i$ мы понимаем прибавление ко всем переменным из множества X_i соответствующих координат вектора α.)

□ Пусть g — произвольное преобразование из группы $I_{\Sigma_n}(f)$. Тогда $f(\mathbf{x}) = f(g(\mathbf{x}))$, или

$$\Psi(h_1(X_1), \dots, h_k(X_k)) = \\ = \Psi(h_1(X_1 \oplus \alpha_1), \dots, h_k(X_k \oplus \alpha_k)). \quad (6.4)$$

Докажем несложную лемму, которая будет использоваться в дальнейших рассуждениях.

Лемма 6.10. *Пусть выполняется равенство*

$$h(X)\Psi_1(Y) \oplus \Psi_2(Y) = h'(X)\Psi'_1(Y) \oplus \Psi'_2(Y),$$

где $X \cap Y = \emptyset$. Тогда $h'(X) = h(X) \oplus a,\ a \in \Omega$.

◁ Перепишем равенство в следующем виде:

$$h(X)\Psi_1(Y) \oplus h'(X)\Psi'_1(Y) = \Psi_2(Y) \oplus \Psi'_2(Y).$$

Поскольку в многочлен Жегалкина функции из правой части последнего равенства не входят переменные из множества X, то для каждого члена многочлена Жегалкина функции $h(X)\Psi_1(Y)$, содержащего переменные из множества X, найдется такой же член многочлена Жегалкина функции $h'(X)\Psi'_1(Y)$ и наоборот. Поскольку каждый член многочлена Жегалкина функции $h(X)\Psi_1(Y)$ (функции $h'(X)\Psi'_1(Y)$) является произведением членов многочленов Жегалкина функций $h(X)$ и $\Psi_1(Y)$ ($h'(X)$ и $\Psi'_1(Y)$), то многочлены Жегалкина

функций $h(X)$ и $h'(X)$, ($\Psi_1(Y)$ и $\Psi'_1(Y)$) совпадают с точностью до константы из Ω_2. $\triangleright$

Из равенства 6.4 и теоремы 6.2 следует, что $h_i(X_i \oplus \alpha_i) = h_i(X_i) \oplus a_i$, $a_i \in \Omega_2, i \in \overline{1,k}$. Подставляя значения $h_i(X_i \oplus \oplus \alpha_i)$ в равенство 6.4, убеждаемся, что $\Psi(y_1 \oplus a_1, \ldots, y_k \oplus a_k) = \Psi(y_1, \ldots, y_k)$. Следовательно, $\overline{g} \in I_{\Sigma_n}(\Psi)$. $\square$

Следствие 1. *В условиях теоремы 6.9 выполняется неравенство*

$$|I_{\Sigma_n}(f)| \leq |I_{\Sigma_k}(\Psi)| \prod_{i=1}^{k} |I_{\Sigma_{n_i}}(h_i)|.$$

$\square$ Из доказательства теоремы 6.9 следует, что отображение, сопоставляющее подстановке $g \in I_{\Sigma_n}(f)$ подстановку $\overline{g} \in \in I_{\Sigma_k}(\Psi)$, является гомоморфизмом соответствующих групп инерции. Ядро этого гомоморфизма состоит из подстановок g, оставляющих все функции h_i на месте, т. е. совпадает с прямым произведением групп $I_{\Sigma_{n_i}}(h_i)$. $\square$

Перейдем к изучению групп инерции функций в группах S_n и Q_n. Согласно теореме 6.8 каждую функцию f, существенно зависящую от всех переменных из множества $X = \{x_1, \ldots, x_n\}$, можно представить в виде

$$f(X) = \Psi(h_1(X_1), \ldots, h_k(X_k)), \tag{6.5}$$

где функции $\Psi, h_1, \ldots, h_k$ удовлетворяют одному из двух условий:

а) $\Psi(y_1, \ldots, y_k) = \underline{y_1} * \cdots * y_k$, $* \in \{\oplus, \wedge, \vee\}$, $2 \leq k \leq n$ и h_i $*$-неразложимы, $i \in \overline{1,k}$;

б) Ψ не допускает простой декомпозиции и $3 \leq k \leq n$ (этот случай включает и функционально неразделимые функции $f = \Psi$ при $n = k$).

В представлении (6.5) разбиение $X = \bigsqcup_{i=1}^{k} X_i$ определено однозначно. Воспользуемся этим фактом для описания группы инерции функции f.

Теорема 6.11. *Пусть функция $f(X)$ существенно зависит от всех своих переменных и удовлетворяет условию а)*

или б). Тогда группы инерции функции f в группах S_n и Q_n описываются следующим образом:

$$I_{S_n}(f) = \{g \in S_n | h_i(g(X)) = h_{j_i}(X_{j_i}) \oplus a_i,\ i \in \overline{1,k},$$

$$\overline{g} = \begin{pmatrix} y_1 & \dots & y_k \\ y_{j_1} \oplus a_1 & \dots & y_{j_k} \oplus a_k \end{pmatrix} \in I_{Q_k}(\Psi)\},$$

$$I_{Q_n}(f) = \{g \in Q_n | h_i(g(X)) = h_{j_i}(X_{j_i}) \oplus a_i,\ i \in \overline{1,k},$$

$$\overline{g} = \begin{pmatrix} y_1 & \dots & y_k \\ y_{j_1} \oplus a_1 & \dots & y_{j_k} \oplus a_k \end{pmatrix} \in I_{Q_k}(\Psi)\}.$$

□ Доказательство проводится полностью аналогично доказательству теоремы 6.9 с той лишь разницей, что в правой части равенства (6.4) стоит суперпозиция вида (6.5), у которой, быть может, переставлены множества переменных $X_1, \dots, X_k$ в соответствии с некоторой подстановкой:

$$\begin{pmatrix} 1 & \dots & k \\ j_1 & \dots & j_k \end{pmatrix}. \square$$

Следствие 1. *В условиях теоремы 6.11 справедливы неравенства*

$$|I_{S_n}(f)| \le |I_{Q_k}(\Psi)| \cdot \prod_{i=1}^{k} |I_{S_{n_i}}(h_i)|,$$

$$|I_{Q_n}(f)| \le |I_{Q_k}(\Psi)| \cdot \prod_{i=1}^{k} |I_{Q_{n_i}}(h_i)|,$$

□ Доказательство проводится аналогично следствию из теоремы 6.9. □

Основываясь на теореме 6.11, можно сводить описание групп инерции в группах S_n и Q_n произвольной функционально разделимой функции f к описанию групп инерции функций-компонент h_i и функции Ψ. Точнее, пусть f — функционально разделимая функция. Тогда для нее существует представление вида (6.5), удовлетворяющее условию *а)* и *б)*.

Рассмотрим функции h_i. Если они функционально разделимы, то для них также существуют представления

$$h_i = \Psi_i(h_{i_1}, \ldots, h_{ik_i}),\ i = \overline{1,k},$$

удовлетворяющие условиям *а)* или *б)* и т. д. После того как процесс разложения функции будет закончен, можно воспользоваться теоремой 6.11 для описания групп инерции функций $h_1, \ldots, h_k, h_{11}, \ldots$.

Часть III

Теория алгоритмов

Глава 1

ЭЛЕМЕНТЫ ТЕОРИИ АЛГОРИТМОВ

Слово алгоритм (или алгорифм) происходит от имени арабского математика (из Хорезма) Мохаммеда ибн Мусы Альхваризми (IX в.), из трактата которого Европа в XII в. познакомилась с десятичной позиционной системой счисления и с арифметическими действиями над числами в этой системе. В связи с этим и само понятие алгоритма ассоциировалось вначале с искусством счета. Постепенно понятие алгоритма трансформировалось и к началу XX в. под алгоритмом стали понимать четко определенную процедуру решения некоторого класса задач или преобразования исходных данных (условий задач) в выходные данные (ответы к задачам). Конечно, приведенное пояснение ни в коей мере не может служить определением понятия алгоритма. Однако таким весьма туманным пониманием алгоритма математики довольствовались вплоть до XX в., пока не появилась необходимость в доказательстве отсутствия алгоритмов для решения некоторых классов задач. Дело в том, что к началу XX в. в математике накопилось много задач алгоритмического характера, не поддававшихся решению, несмотря на многочисленные усилия математиков. Примером может служить известная 10-я проблема Гильберта, сформулированная им в докладе «Математические проблемы», произнесенном в 1900 г. на II Международном конгрессе математиков в Париже. Эта проблема заключалась в нахождении способа, позволяющего за конечное число операций установить, разрешимо ли произвольно заданное диофантово уравнение (алгебраическое уравнение с целыми коэффициентами) в целых числах или нет. Наличие

таких задач зарождало у математиков идею о доказательстве отсутствия алгоритмов их решения. Результаты об отсутствии алгоритмов решения задач теми или иными ограниченными средствами к тому времени уже были получены. Например, было известно, что задачи трисекции угла, удвоения куба и т. п. неразрешимы с помощью циркуля и линейки. Однако теперь речь шла об отсутствии алгоритма вообще. Для решения такого рода задач необходим был новый качественный скачок в математике. А именно нужно было дать точное определение понятия алгоритма, поскольку невозможно доказать отсутствие чего-то туманного и расплывчатого. Задача определения алгоритма была решена в 1930-х гг. в работах математиков и логиков Гильберта, Гёделя, Чёрча, Клини, Поста и Тьюринга. Было дано несколько разных определений понятия алгоритма. При этом Гильберт, Гёдель, Чёрч и Клини подошли к понятию алгоритма через вычислимые арифметические функции, а Пост и Тьюринг — через сведение алгоритма к элементарным преобразованиям слов в конечных алфавитах. Второй подход к определению алгоритма был использован в 1940-х гг. и советским математиком А. А. Марковым (1903–1979). Определенные им алгоритмы получили название нормальных алгоритмов Маркова. Прежде чем дать строгое определение алгоритма, математики проанализировали известные примеры алгоритмов и выделили наиболее общие их свойства. Для выявления этих свойств рассмотрим и мы повнимательнее, например, хорошо известный из курса алгебры алгоритм Евклида нахождения наибольшего общего делителя двух натуральных чисел. В нем предписывается делить с остатком 1-е число на 2-е, затем 2-е на полученный (первый) остаток, затем 1-й остаток на 2-й остаток и т. д. до тех пор, пока не получится остаток, равный нулю. Последний, не равный нулю остаток и будет искомым наибольшим общим делителем. В итоге любая пара натуральных чисел a, b преобразуется в число d, равное их наибольшему общему делителю:

$$a, b \to d.$$

Характерными свойствами алгоритма Евклида являются:

1) массовость — алгоритм может быть применен к любой паре натуральных чисел, причем сама схема работы алгоритма не зависит от исходных данных (в любом случае делим 1-е число на 2-е и т. д.);

2) дискретность — весь алгоритм можно разбить на отдельные элементарные операции (шаги алгоритма), которые могут быть занумерованы натуральными числами в порядке их выполнения;

3) детерминированность — каждый шаг алгоритма однозначно определяется его предыдущими шагами;

4) конечная определенность — исходные данные, а также результаты каждого шага алгоритма и ответ записываются в виде конечных последовательностей символов исходного конечного алфавита. Нетрудно видеть, что указанными свойствами обладают и другие известные нам алгоритмы, например алгоритмы умножения целых чисел, многочленов и матриц, алгоритм Гаусса для решения систем линейных уравнений и т. д. На примере алгоритма Евклида просматриваются не только общие черты алгоритма вообще, но и упомянутые выше два подхода к общему определению алгоритма. А именно, с одной стороны, это вычисление значений некоторой числовой функции двух переменных a, b, а с другой — преобразование одной последовательности символов (цифр чисел a и b, разделенных запятой) в другую последовательность (цифр числа d). Конечно, не все алгоритмы должны иметь дело с натуральными числами, однако слова в произвольном конечном (и даже счетном) алфавите можно занумеровать натуральными числами, и потому любой алгоритм можно свести к вычислению арифметической функции. Тем самым строгое определение алгоритма при первом подходе, по существу, сводится к нахождению какого-то строгого и конструктивного описания всех вычислимых арифметических функций. При втором подходе требуется четко определить элементарные преобразования слов и последовательности их выполнения во всех алгоритмах. В каждом из имеющихся в настоящее время определений алгоритма, по существу, описывается некоторый класс алгоритмов и приводится обоснование того, что решение любой массовой задачи осуществляется алгоритмом из

выделенного класса. Обоснование, как правило, заключается в построении большого числа примеров и в доказательстве замкнутости выделенного класса алгоритмов относительно различного рода комбинаций алгоритмов (композиции, объединения, разветвления и т. п.). Наиболее убедительным доводом в указанном обосновании явилось доказательство равносильности всех имеющихся определений алгоритмов. В итоге к настоящему времени в математическом мире сложилось достаточно единодушное мнение о законности имеющихся определений алгоритма, так что если доказывается отсутствие, скажем, нормального алгоритма для решения какого-либо класса задач, то говорится об отсутствии алгоритма вообще. В данной главе мы опишем три различных определения понятия алгоритма и приведем простейшие примеры на доказательство теорем об отсутствии алгоритмов для решения некоторых задач. В заключение отметим, что наличие алгоритмически неразрешимых задач ни в коей мере не противоречит общепризнанному положению диалектического материализма о познаваемости мира. Дело в том, что в теории алгоритмов речь идет об отсутствии общего алгоритма для решения слишком широкого (бесконечного) класса задач, а не какой-либо отдельной задачи.

1.1. НОРМАЛЬНЫЕ АЛГОРИТМЫ

Каждый нормальный алгоритм (НА) является определенным процессом преобразования слов в некотором конечном алфавите и задается набором допустимых элементарных преобразований и правилами, определяющими порядок применения этих преобразований. При этом в качестве элементарного преобразования используется замена одного вхождения подслова в слове некоторым другим (или тем же словом). Всевозможные замены для заданного НА определяются его схемой, а последовательность проведения замен — схемой и некоторыми дополнительными соглашениями. Эти соглашения одни и те же для всех НА, а потому НА, по существу, однозначно определяется алфавитом и схемой.

Определение 1.1. *Схемой нормального алгоритма* $\mathfrak{A}$ *в алфавите* $A = \{a_1, a_2, \ldots, a_n\}$ *называется упорядоченная последовательность*

$$\begin{array}{c} P_1\ \gamma_1\ Q_1 \\ \ldots \\ P_m\ \gamma_m\ Q_m \end{array} \tag{1.1}$$

слов в алфавите $A \cup \{\to, \cdot\}$, *где* P_i, Q_i — *слова в алфавите* A, *а* γ_i *есть слово* $\to$ *или* $\to \cdot$. *При этом слово* $P_i\gamma_i Q_i$ *схемы* (1.1) *называется ее* i*-й формулой с левой частью* P_i *и правой частью* Q_i.

Формула $P_i\gamma_i Q_i$ *называется простой, если* γ_i *есть* $\to$, *и заключительной, если* γ_i *есть* $\to \cdot$.

Действие НА $\mathfrak{A}$ на слово R в алфавите A описывается следующим определением.

Определение 1.2. *Пусть* R — *слово в алфавите* A *и хотя бы одно из слов* $P_1, \ldots, P_m$ *является его подсловом. Элементарным преобразованием слова* R *по нормальному алгоритму* $\mathfrak{A}$ *со схемой* (1.1) *называется замена в* R *первого вхождения слова* P_i *с наименьшим номером* $i \in \overline{1, m}$ *словом* Q_i. *Если результатом указанного элементарного преобразования является слово* R_1, *то пишут* $\mathfrak{A} : R\ \gamma_i\ R_1$, *т. е.* $\mathfrak{A} : R \to R_1$ *или* $\mathfrak{A} : R \to \cdot\, R_1$.

Определение 1.3. *Говорят, что НА* $\mathfrak{A}$ *применим к слову* R *в алфавите* A *и перерабатывает его в слово* R_k, *если существует конечная последовательность слов*

$$R = R_0, R_1, \ldots, R_k, k \geq 0, \tag{1.2}$$

в которой

1) $\mathfrak{A} : R_i \to R_{i+1},\ i \in \overline{0, k-2}$;

2) $\mathfrak{A} : R_{k-1} \to \cdot R_k$ *или* $\mathfrak{A} : R_{k-1} \to R_k$ *и слово* R_k *не содержит подслов* $P_1, \ldots, P_m$.

В противном случае говорят, что алгоритм $\mathfrak{A}$ *не применим к слову* R.

Из приведенных определений естественным образом извлекается описание процесса переработки слова R по нормальному алгоритму $\mathfrak{A}$ (или нормальным алгоритмом $\mathfrak{A}$). А именно

по заданному слову R НА $\mathfrak{A}$ строит последовательность слов. Если R не содержит подслов $P_1, \ldots, P_m$, то эта последовательность одноэлементна и состоит из единственного слова R. Если же R содержит хотя бы одно из подслов $P_1, \ldots, P_m$, то производится элементарное преобразование слова R по НА $\mathfrak{A}$, в результате чего получается вполне определенное слово R_1. Если при переходе от R к R_1 использовалась заключительная формула, то искомая последовательность двухэлементна: R, R_1. В противном случае, так же, как и выше, по слову R_1 строится слово R_2 и т. д. В процессе построения указанной последовательности могут представиться три различные возможности: или на каком-то шаге будет использована заключительная формула, или появится слово, не содержащее подслов $P_1, \ldots, P_m$, или не произойдет ни того, ни другого. В первых двух случаях мы получим конечную последовательность, последнее слово которой называют результатом применения НА к слову R и обозначают $\mathfrak{A}(R)$. В третьем случае процесс преобразования слов по НА $\mathfrak{A}$ будет длиться бесконечно, это и означает, что НА $\mathfrak{A}$ не применим к R.

Таким образом, НА $\mathfrak{A}$ в алфавите A задает частичное отображение множества $W(A)$ всех слов в алфавите A в себя. Выбирая различные схемы, мы будем получать различные НА.

Если B — алфавит, содержащий A, то НА в алфавите B называются нормальными алгоритмами над алфавитом A.

Приведем примеры нормальных алгоритмов. При этом буквой Λ всегда обозначается пустое слово (в любом алфавите).

1. НА в алфавите A со схемой

$$\Lambda \to \cdot\Lambda$$

перерабатывает любое слово $R \in W(A)$ в себя, причем последовательность слов, соответствующая слову R, имеет вид R, R.

2. НА со схемой

$$\Lambda \to \Lambda$$

не применим ни к одному слову. Последовательность, соответствующая слову R, будет бесконечной:

$$R, R, R, \ldots .$$

3. НА в алфавите A со схемой

$$\Lambda \to \cdot Q$$

приписывает к любому слову $R \in W(A)$ слева слово Q:

$$\mathfrak{A}(R) = QR.$$

4. Построим НА $\mathfrak{A}$, приписывающий к любому слову $R \in$ $\in W(A)$ справа фиксированное непустое слово $Q \in W(A)$. Это сделать несколько сложнее, чем приписывание слова Q слева. Для этого удобнее расширить алфавит A, добавив к нему одну, новую букву α, и построить искомый НА в алфавите $A' = A \cup \{\alpha\}$ (т. е. над A). Нетрудно проверить, что нужное нам преобразование будет осуществлять НА со схемой

$$\alpha\ a_1 \to a_1\alpha$$

$$\alpha\ a_2 \to a_2\alpha$$

$$\dots$$

$$\alpha\ a_n \to a_n\alpha$$

$$\alpha \to \cdot\, Q$$

$$\Lambda \to \alpha.$$

Последовательность слов, соответствующая произвольному слову R в алфавите A, для этого НА имеет вид

$$R = a_{i_1}a_{i_2}\dots a_{i_k},\ \alpha a_{i_1}a_{i_2}\dots a_{i_k},\ a_{i_1}\alpha a_{i_2}\dots a_{i_k}, \dots$$

$$\dots, a_{i_1}a_{i_2}\dots a_{i_k}\alpha,\ a_{i_1}a_{i_2}\dots a_{i_k}Q.$$

Очевидно, что то же самое преобразование слов будет осуществлять НА, полученный из $\mathfrak{A}$ любой перестановкой первых n формул его схемы. В связи с этим вместо первых n формул схемы пишут просто

$$\alpha x \to x\alpha,\ x \in A,$$

так что вся схема запишется в виде

$$\alpha x \to x\alpha,\ x \in A,$$

$$\alpha \to \cdot Q,$$

$$\Lambda \to \alpha.$$

Заметим, что, выполняя свою задачу приписывания к словам из $W(A)$ справа слова Q, мы совсем не интересовались переработкой слов в алфавите A', содержащих букву α. Здесь алгоритм $\mathfrak{A}$ будет действовать иначе. А именно слово R в алфавите A', содержащее k вхождений буквы α, он будет перерабатывать в слово

$$R'Q\underbrace{\alpha \dots \alpha}_{k-1},$$

где R' получается из R удалением всех вхождений буквы α (т. е. R' есть проекция R в алфавит A).

5. Построим НА $\mathfrak{A}$, перерабатывающий любое слово $R = a_{i_1} \dots a_{i_k} \in W(A)$ в перевернутое слово $\breve{R} = a_{i_k} \dots a_{i_1}$. Для этого добавим к A две новые буквы α, β и возьмем следующую схему $\mathfrak{A}$ в алфавите $A_2 = A \cup \{\alpha, \beta\}$:

$$\begin{gathered}
\alpha\alpha \to \beta, \\
\beta\alpha \to \beta, \\
\beta x \to x\beta,\ x \in A, \\
\beta \to \cdot\Lambda, \\
\alpha xy \to y\alpha x,\ y \in A, \\
\Lambda \to \alpha.
\end{gathered}$$

Проследите в качестве упражнения, что $\mathfrak{A}(R) = \breve{R}$ для любого слова R в алфавите A.

Приведем еще примеры НА над числами. Условимся натуральное число n записывать в виде n вертикальных палочек:

$$\underbrace{|| \dots |}_{n}.$$

6. НА в алфавите $\{|, +\}$ со схемой

$$+ \ \to \cdot\Lambda$$

осуществляет сложение натуральных чисел.

7. НА в алфавите $\{|, *, \alpha, \beta\}$ со схемой

$$\begin{gathered}\beta| \to |\beta,\\ \alpha| \to |\beta\alpha,\\ \alpha \to \Lambda,\\ |* \to *\alpha,\\ *| \to *,\\ * \to \Lambda,\\ \beta \to |\end{gathered}$$

осуществляет умножение натуральных чисел. Заметим, что рассмотренные нами примеры и большое число других примеров НА можно найти в монографии А. А. Маркова [40]. Там же А. А. Марков формулирует и обосновывает свой принцип нормализации алгоритмов. В формулировке этого принципа участвуют математически определенное понятие нормального алгоритма и не определенное, а лишь интуитивное понятие алгоритма, и потому для изложения принципа нормализации нам придется пользоваться обоими этими понятиями.

1.2. ПРИНЦИП НОРМАЛИЗАЦИИ АЛГОРИТМОВ

Пусть $\mathfrak{A}$ есть алгоритм (в интуитивном смысле), решающий какой-то класс задач. Так как условие и ответ к каждой задаче записываются словами, например, русского языка, то, не теряя общности, можно считать, что алгоритм $\mathfrak{A}$ перерабатывает слова в некотором конечном алфавите A. В этом случае будем говорить, что $\mathfrak{A}$ есть алгоритм над A. Из самого назначения алгоритма $\mathfrak{A}$ следует, что нас интересует его действие лишь на те слова в алфавите A, которые являются записями условий наших задач. Например, в алгоритме сложения натуральных чисел мы интересуемся его действием лишь на слова вида $|\ldots|+|\ldots|$ (с одним вхождением символа $+$), к другим словам он может быть и не применим. В связи с этим естественно считать, что алгоритм $\mathfrak{A}$ к каким-то словам R применим и перерабатывает их в слова $\mathfrak{A}(R)$, а к некоторым — не применим. В крайних случаях множества слов, к которым $\mathfrak{A}$ применим или не применим, могут быть пустыми.

Определение 1.4. *Два алгоритма* $\mathfrak{A}$, $\mathfrak{L}$ *над алфавитом* A *называются эквивалентными относительно* A*, если они применимы к одним и тем же словам из* $W(A)$ *и каждое слово из* $W(A)$*, к которому они применимы, перерабатывают в одно и то же слово. Этот факт будем записывать в виде*

$$\mathfrak{A}(R) \simeq \mathfrak{L}(R),\ R \in W(A).$$

Принцип нормализации алгоритмов, сформулированный А. А. Марковым, заключается в следующем.

Всякий алгоритм над алфавитом A *эквивалентен относительно* A *некоторому нормальному алгоритму над* A.

Короче его формулируют еще и так.

Всякий алгоритм нормализуем.

Таким образом, принцип нормализации, по существу, утверждает, что с точностью до эквивалентности все алгоритмы исчерпываются нормальными алгоритмами.

Подчеркнем, что принцип нормализации алгоритмов не может быть доказан, ибо он устанавливает связь между неопределенным (интуитивным) понятием алгоритма и точно определенным понятием НА.

Этот принцип можно лишь как-то обосновать, подкрепить разумными доводами и после этого принять или не принять. Согласившись с этим принципом, его можно считать точным определением алгоритма (алгоритм — это нормальный алгоритм), и потому для доказательства отсутствия алгоритма для решения какого-то класса задач достаточно доказать отсутствие НА, решающего этот класс задач. Для обоснования своего принципа А. А. Марков приводит большое число примеров нормализации известных алгоритмов, а также доказывает, что различные комбинации НА сами являются НА. Этот довод весьма убедителен, поскольку в практике человека сложные алгоритмы обычно строятся путем различных сочетаний простых алгоритмов. Так, наиболее часто используется композиция алгоритмов, когда к результату действия одного алгоритма применяется другой алгоритм. В связи с этим для примера докажем следующее утверждение.

Теорема 1.1. *Если $\mathfrak{A}_1$ и $\mathfrak{A}_2$ — НА над алфавитом A, то их композиция $\mathfrak{A}_2 \circ \mathfrak{A}_1$ эквивалентна относительно A некоторому НА над A.*

□ Пусть $\mathfrak{A}_1$ и $\mathfrak{A}_2$ — НА соответственно в алфавитах A_1, A_2 и $A_i \supset A$, $i = 1, 2$. Для простоты рассмотрим случай, когда $A_1 = A_2 = A$. Условимся схему алгоритма $\mathfrak{A}_i$ обозначать той же буквой $\mathfrak{A}_i$, $i = 1, 2$. Каждой букве a_i алфавита A_i сопоставим символ $\overline{a}_i$, и множество всех полученных таким образом символов обозначим через $\overline{A}$. Образуем новый алфавит

$$B = A \cup \overline{A} \cup \{\alpha, \beta\}$$

и рассмотрим НА $\mathfrak{L}$ в алфавите B со схемой:

$$x\alpha \to \alpha x,\ x \in A$$
$$\alpha x \to \alpha \overline{x},\ x \in A,$$
$$\overline{x}y \to \overline{xy},\ x, y \in A,$$
$$\overline{x}\beta \to \beta \overline{x},\ x \in A,$$
$$\beta \overline{x} \to \beta x,\ x \in A,$$
$$x\overline{y} \to xy,\ x, y \in A,$$
$$\alpha\beta \to \cdot\Lambda,$$
$$\overline{\mathfrak{A}}_2^{\beta,\alpha},$$
$$\mathfrak{A}_1^{\alpha},$$

где $\mathfrak{A}_1^{\alpha}$ — схема, полученная добавлением к схеме $\mathfrak{A}$ формулы

$$\Lambda \to \cdot\Lambda, \tag{1.3}$$

с последующей заменой всюду точки $\cdot$ символом α, а $\overline{\mathfrak{A}}_2^{\beta,\alpha}$ — схема, полученная добавлением к схеме $\mathfrak{A}_2$ формулы (1.3) с последующей заменой каждой буквы $a_i \in A$ буквой $\overline{a}_i$, каждой точки $\cdot$ буквой β и каждой формулы вида

$$\Lambda \to P$$

формулой

$$\alpha \to \alpha P.$$

Сначала заметим, что добавление к схеме любого НА $\mathfrak{A}$ формулы (1.3) приводит к эквивалентному алгоритму $\mathfrak{A}'$. В применении к слову R разница между $\mathfrak{A}$ и $\mathfrak{A}'$ заключается лишь в условиях окончания работы. Если $\mathfrak{A}(R) = Q$ и в слово Q не входит ни одна из левых частей формул из $\mathfrak{A}$, то алгоритм $\mathfrak{A}'$ сделает после этого еще один шаг по формуле (1.3) и снова выдаст слово Q.

Теперь можно проследить работу НА $\mathfrak{L}$ в применении к слову $R \in W(A)$. Так как в $W(A)$ не входят буквы из $\overline{A} \cup \cup \{\alpha,\ \beta\}$, то к слову R будет вначале применяться алгоритм $\mathfrak{A}_1^{\alpha}$. Если $\mathfrak{A}_1$ не применим к R, то и $\mathfrak{A}_1^{\alpha}$ не применим, а потому и $\mathfrak{L}$ не будет применим к R. Если же $\mathfrak{A}_1$ применим к R и $\mathfrak{A}_1(R) = Q$, то $\mathfrak{A}_1^{\alpha}$ переведет R в слово вида $Q'\alpha Q''$, которое затем формулой $x\alpha \to \alpha x$ переведется в слово αQ. Далее с помощью формул $\alpha x \to \alpha\overline{x}$ и $\overline{x}y \to \overline{xy}$ мы придем к слову $\alpha\overline{Q}$, где $\overline{Q}$ — слово, полученное из Q заменой букв a_i на $\overline{a}_i$. К слову $\alpha\overline{Q}$ применимы лишь формулы схемы $\overline{\mathfrak{A}}_2^{\beta,\alpha}$. При этом если $\mathfrak{A}_2$ не применим к Q, то $\overline{\mathfrak{A}}_2^{\beta,\alpha}$ не будет применим к $\alpha\overline{Q}$. Если же $\mathfrak{A}_2(Q) = P$, то, применяя к $\alpha\overline{Q}$ алгоритм $\overline{\mathfrak{A}}_2^{\beta,\alpha}$, а затем формулы вида $\overline{x}\beta \to \beta\overline{x}$, $\beta\overline{x} \to \beta x$, $x\overline{y} \to xy$, мы придем к слову $\alpha\ \beta\ P$, к которому применима лишь формула $\alpha, \beta \to \cdot\Lambda$.

В итоге получим: $\mathfrak{L}$ применим к R тогда и только тогда, когда $\mathfrak{A}_1$ применим к R и $\mathfrak{A}_2$ применим к $\mathfrak{A}_1(R)$, причем в этом случае

$$\mathfrak{L}(R) = \mathfrak{A}_2(\mathfrak{A}_1(R)). \ \square$$

Используя доказанную теорему и НА сложения, вычитания и умножения натуральных чисел, мы можем утверждать существование НА для вычисления любого многочлена над $\mathbb{N}$.

На практике часто используется также разветвление алгоритмов: к слову R применим алгоритм $\mathfrak{A}$, если оно удовлетворяет некоторому свойству (*), и алгоритм $\mathfrak{L}$ в противном случае. Нетрудно доказать, что если $\mathfrak{A}$ и $\mathfrak{L}$ — нормальные алгоритмы и свойство (*) можно распознавать некоторым НА, то и указанное разветвление можно реализовать подходящим НА.

1.3. МАШИНЫ ТЬЮРИНГА

В 1936 г. независимо одна от другой появились работы английского математика А. Тьюринга и американского математика Э. Поста, в которых были даны уточнения понятия алгоритма или «эффективной процедуры» для решения массовых задач в терминах некоторых идеализированных вычислительных машин, называемых теперь машинами Тьюринга или Тьюринга–Поста. Устройства этих машин у А. Тьюринга и Э. Поста отличаются лишь несущественными деталями, а именно процедура вычислений у Э. Поста раздроблена на более мелкие операции, чем у А. Тьюринга. В настоящее время в литературе описано много различных модификаций таких идеализированных машин. Мы далее рассмотрим одну из них, называя ее машиной Тьюринга (сокращенно — МТ).

Машина Тьюринга, как и нормальный алгоритм, предназначается для переработки слов в некотором конечном алфавите $A = \{a_0, a_1, \ldots, a_n\}$, называемом внешним алфавитом машины. Слова в алфавите А записываются на ленту МТ, разбитую на ячейки, так, что в каждую ячейку записывается какая-либо одна буква алфавита A. Сама лента называется внешней памятью МТ. Предполагается, что лента в процессе работы может наращиваться, т. е. ленту можно считать потенциально бесконечной в обе стороны. Все вновь пристраиваемые ячейки заполняются символом a_0, который называется пустым символом.

Переработка слов, записываемых на ленту МТ, осуществляется в дискретном времени по тактам, занумерованным натуральными числами, под действием так называемого управляющего устройства. Последнее в каждый такт может находиться в одном из конечного множества состояний $Q = \{q_0, q_1, \ldots, q_n\}$, называемых внутренними состояниями.

Управляющее устройство связано с лентой так называемой считывающей головкой, которая в каждый такт находится против одной из ячеек ленты (обозревает одну ячейку). По команде управляющего блока, зависящей от внутреннего состояния и от содержимого обозреваемой ячейки, машина или останавливается, или переходит в новое состояние.

В последнем случае головка заменяет символ обозреваемой ячейки тем же самым или новым символом из A и остается против той же ячейки, или сдвигается на одну ячейку вправо, или сдвигается на одну ячейку влево. При этом если головка обозревала последнюю (первую) ячейку ленты и ей дана команда сдвинуться вправо (влево), то к ленте справа (слева) автоматически добавляется новая ячейка с буквой a_0. Оказавшись в состоянии q_n, МТ прекращает работу (q_n называют стоп-состоянием).

Из приведенного описания видно, что положение МТ в каждый момент времени полностью определяется следующими параметрами:

1) словом, записанным на ленте;
2) внутренним состоянием;
3) номером обозреваемой ячейки.

Если в некотором такте работы МТ на ее ленте записано слово

$$P = a_{i_1} \dots a_{i_k},$$

управляющее устройство находится в состоянии q_j и головка обозревает ячейку с номером r, то всю эту информацию можно записать одним, так называемым машинным словом

$$\hat{P} = a_{i_1} \dots a_{i_{r-1}} q_j a_{i_r} \dots a_{i_k}$$

(символ внутреннего состояния записывается перед буквой, записанной в обозреваемой ячейке). При переходе к новому такту машинное слово меняется согласно системе команд МТ. Изменения зависят от состояния q_j МТ и от содержимого a_j обозреваемой ячейки. Поэтому каждая команда записывается в виде формулы

$$q_j a_i \to q_s a_t \Delta, \tag{1.4}$$

где q_s — состояние, в которое переходит МТ, a_t — буква, на которую заменяется a_i (не исключаются случаи $q_s = q_j$, $a_t = a_i$), Δ — одна из букв H, R, L, которые означают соответственно сохранение положения головки, сдвиг ее на одну ячейку вправо, сдвиг на одну ячейку влево. Слова $q_j a_i$ и $q_s a_t \Delta$ называются соответственно левой и правой частями команды.

Подчеркнем, что система команд МТ удовлетворяет следующему естественному условию детерминизма: каждое из слов вида $q_j a_i$ (если q_j не стоп-состояние) является левой частью ровно одной команды. Это условие позволяет всю систему команд МТ записать в виде таблицы (табл. 3.1).

Таблица 3.1
Система команд машины Тьюринга

	q_0	q_1	$\dots$	q_j	$\dots$	q_{n-1}
a_0	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
$\vdots$						
a_i	$\dots$	$\dots$	$\dots$	$q_S a_i \Delta$	$\dots$	$\dots$
$\vdots$						
a_m	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$

Заметим, что в левых частях команд, а потому и во входной строке таблицы, состояние q_n не участвует, поскольку в состоянии q_n МТ прекращает работу. Опишем теперь процесс преобразования слов в алфавите A описанной выше МТ $\mathfrak{M}$ с системой команд S. В начале работы МТ устанавливается в состояние q_0, на ее ленту записывается исходное слово $P \in \in W(A)$ и считывающая головка помещается против первой (самой левой) ячейки (т. е. обозревает первую букву слова P). Если $P = a_{i_1} \dots a_{i_k}$, то вся информация о начале работы запишется машинным словом

$$\hat{P} = q_0 a_{i_1} a_{i_2} \dots a_{i_k}.$$

Теперь, как и при описании работы НА, сопоставим слову P конечную или бесконечную последовательность машинных слов

$$\hat{P}_0 = \hat{P}, \hat{P}_1, \hat{P}_2, \dots . \tag{1.5}$$

Для этого находим в S команду вида

$$q_0 a_{i_1} \to q_j a_r \Delta,$$

и в зависимости от значения Δ, равного H, R или L, в качестве $\hat{P}_1$ берем соответственно слово

$$q_j a_r a_{i_2} \dots a_{i_k},\ a_r q_j a_{i_2} \dots a_{i_k},\ \text{ или } q_j a_0 a_r a_{i_2} \dots a_{i_k}.$$

Если $q_j = q_n$, то в качестве искомой возьмем двухэлементную последовательность $\hat{P}_0, \hat{P}_1$. В противном случае, как и выше, построим слово P_2 и т. д. Продолжая этот процесс, мы получим искомую последовательность (1.5). Последовательность слов в алфавите A, полученных удалением из слов $\hat{P}_i$ символов-состояний, назовем путем переработки слова P. Если последовательность (1.5) конечна и оканчивается словом $P' q_n P''$, то говорят, что машина $\mathfrak{M}$ применима к слову P и перерабатывает его в слово $P'P'' = Q$. Этот факт записывают в виде

$$\mathfrak{M}(P) = Q.$$

Если же последовательность (1.5) бесконечна, то говорят, что машина не применима к слову P. Таким образом, любая МТ осуществляет частичное отображение множества слов в алфавите A в себя.

Приведем пример МТ, осуществляющей удвоение натуральных чисел. Как и при построении НА, натуральное число k будем изображать в виде последовательности k вертикальных черточек. В качестве внешнего алфавита возьмем множество $A = \{0,\ |,\ \alpha\}$, где 0 — пустой символ, а в качестве множества внутренних состояний

$$Q = \{q_0, q_1, q_2, q_3\}.$$

Систему команд зададим таблицей 3.2. В начале работы на ленту записывается слово $k = ||\dots|$, а вся информация о начале работы определится машинным словом

$$q_0||\dots|.$$

В качестве упражнения постройте последовательность машинных слов, начинающуюся словом $q_0|||$.

Таблица 3.2
Система команд машины Тьюринга, удваивающей натуральное число

	q_0	q_1	q_2
0	$q_1\alpha H$	$q_2 0 L$	$q_3 0 H$
|	$q_0\alpha L$	$q_0\alpha L$	$q_2 \| H$
α	$q_0\alpha L$	$q_1\alpha R$	$q_2 \| L$

Из сравнения определений НА и МТ видно, что в машинах Тьюринга процесс преобразования слов разбит на более мелкие операции — в них в каждый такт заменяется не более одной буквы слова. Кроме того, в них есть и другое сильное ограничение — расширение слова может происходить только за счет приписывания новых символов слева или справа от слова (т. е. в начале его или в конце). В связи с этим строить НА для решения задач, как правило, проще, чем МТ. Так задачу удвоения натурального числа можно решить НА со схемой

$$\alpha| \to ||\alpha,$$

$$\alpha \to \cdot\Lambda,$$

$$| \to \alpha|.$$

В нем всего лишь три команды вместо девяти в построенной выше машине, да и число тактов работы по удвоению слова существенно меньше. На первый взгляд может показаться, что и в принципе МТ имеют меньше возможностей по сравнению с НА. Однако А. Тьюрингом и Э. Постом была выдвинута гипотеза (или тезис) о том, что любой алгоритм над алфавитом A эквивалентен относительно A алгоритму, осуществляемому подходящей машиной Тьюринга (соответственно Поста). Как и тезис А. А. Маркова о нормализации алгоритмов, эта гипотеза не может быть доказана, ее можно лишь подкреплять какими-либо разумными доводами. Наиболее важным из них является известная к настоящему времени теорема о том, что всякий НА осуществим на МТ.

1.4. НУМЕРАЦИЯ СЛОВ И АРИФМЕТИЗАЦИЯ АЛГОРИТМОВ

Всякий алгоритм $\mathfrak{A}$ в алфавите A определяет частичное отображение множества слов $W(A)$ в алфавите A в себя:

$$\mathfrak{A} : W(A) \to W(A).$$

Если все слова в алфавите A занумеровать натуральными числами так, чтобы разные слова имели разные номера, то по алгоритму $\mathfrak{A}$ можно будет определить частичную арифметическую функцию

$$f_{\mathfrak{A}} : \mathbb{N}_0 \to \mathbb{N}_0,$$

положив $f_{\mathfrak{A}}(n) = m$ в том случае, когда n, m есть номера соответственно слов P, Q и $\mathfrak{A}(P) = Q$, и полагая функцию $f_{\mathfrak{A}}$ неопределенной в других случаях. Если при этом нумерация слов осуществляется эффективно, т. е. существуют алгоритмы нахождения номеров слов и слов по номерам, то с помощью алгоритма $\mathfrak{A}$ можно вычислять значения функции $f_{\mathfrak{A}}(x)$ при всех значениях x, в которых она определена, т. е. функция $f_{\mathfrak{A}}(x)$ будет вычислимой. И наоборот, если мы каким-то образом умеем вычислять значения функции $f_{\mathfrak{A}}(x)$, то, по существу, мы сможем выяснить, в какие слова перерабатывает алгоритм $\mathfrak{A}$ все слова в алфавите A, к которым он применим.

Таким образом, при существовании эффективной нумерации слов в алфавите A описание алгоритмов в алфавите A, а тем самым и всех алгоритмов, сводится к описанию класса всех вычислимых функций.

Опишем некоторые способы нумерации слов в конечном или счетном алфавите.

1) Канторовская нумерация. Занумеруем сначала пары целых неотрицательных чисел, т. е. множество $\mathbb{N}_0^2$. Для этого упорядочим множество $\mathbb{N}_0^2$, положив для пар $(x, y), (u, v) \in \in \mathbb{N}_0^2$:

$$(x, y) \prec (u, v) \Leftrightarrow x + y < u + v$$

или

$$x + y = u + v, x < u.$$

Легко видеть, что таким образом определенное отношение $\prec$ есть отношение линейного порядка на $\mathbb{N}_0^2$ и существует единственное биективное отображение

$$c_2 : \mathbb{N}_0^2 \to \mathbb{N}_0,$$

такое что

$$c_2(0,0) = 0,$$

$$c_2(x,y) < c_2(u,v) \Leftrightarrow (x,y) \prec (u,v).$$

Нетрудно подсчитать, что

$$c_2(x,y) = \frac{(x+y)(x+y+1)}{2} + x, \tag{1.6}$$

и потому существует алгоритм нахождения номера любой заданной пары (x,y).

Выясним теперь, как по номеру неизвестной пары $c_2(x,y) = n$ найти ее компоненты x и y. Обозначим их как функции от n:

$$x = l_2^1(n),\ y = l_2^2(n)$$

или, подробнее,

$$x = l_2^1(c_2(x,y)),\ y = l_2^2(c_2(x,y)). \tag{1.7}$$

Исходя из равенств (1.6) и $c_2(x,y) = n$, непосредственной проверкой убеждаемся, что

$$(2x+2y+1)^2 \le 8n+1 < (2x+2y+3)^2.$$

Отсюда имеем

$$x+y \le \frac{\sqrt{8n+1}-1}{2} < x+y+1.$$

Следовательно,

$$x+y = \left[\frac{\sqrt{8n+1}-1}{2}\right].$$

А так как

$$n = \frac{(x+y)(x+y+1)}{2} + x,$$

то отсюда нетрудно найти x и y, т. е. $l_2^1(n)$ и $l_2^2(n)$:

$$l_2^1(n) = x = n - \frac{1}{2}\left[\frac{\sqrt{8n+1}-1}{2}\right]\left[\frac{\sqrt{8n+1}+1}{2}\right];$$

$$l_2^2(n) = y = \left[\frac{\sqrt{8n+1}-1}{2}\right]\left(1 + \frac{1}{2}\left[\frac{\sqrt{8n+1}+1}{2}\right]\right) - n.$$

Отсюда видно, что по номеру n пары можно эффективно найти компоненты этой пары.

Отметим еще, что наряду с соотношениями (1.7) функции c_2, $l_2^1(n)$, $l_2^2(n)$ связаны также равенством

$$c_2(l_2^1(z),\ l_2^2(z)) = z$$

при любом $z \in \mathbb{N}_0$.

Теперь индуктивно можно построить нумерации для наборов чисел любой фиксированной длины.

Допустим, что для некоторого $n \geq 2$ определены вычислимые функции

$$c_n : \mathbb{N}_0^n \to \mathbb{N}_0,\ l_n^i : \mathbb{N}_0 \to \mathbb{N}_0,$$

такие что:

1) c_n — биективна;
2) $c_n(l_n^1(z), \ldots, l_n^n(z)) = z$;
3) $l_n^i(c_n(x_1, \ldots, x_n)) = x_i,\ i \in \overline{1,n}$.

Занумеруем элементы множества $\mathbb{N}^{n+1}$, положив

$$c_{n+1}(x_1, \ldots, x_{n+1}) = c_n(c_2(x_1, x_2), x_3, \ldots, x_{n+1}).$$

Так как c_2 и c_n биективны, то биективна и функция c_{n+1}. При этом если

$$c_{n+1}(x_1, \ldots, x_{n+1}) = z,$$

то

$$x_1 = l_2^1(l_n^1(z)),\ x_2 = l_2^2(l_n^1(z)),\ x_i = l_n^{i-1}(z),\ i \in \overline{3, n+1}.$$

Определив теперь функции $l^i_{n+1},\ i \in \overline{1, n+1}$, равенствами

$$l^1_{n+1}(z) = l^1_2(l^1_n(z)),\ l^2_{n+1}(z) = l^2_2(l^1_n(z)),$$
$$l^i_{n+1}(z) = l^{i-1}_n(z),\ i \in \overline{3, n+1},$$

мы выразим компоненты $x_1, \ldots, x_{n+1}$ через номер z набора $(x_1, \ldots, x_{n+1})$. Так как функции $l^i_2,\ l^i_n$ вычислимы, то вычислимы и функции l^i_{n+1}. Кроме того, имеют место тождества:

$$c_{n+1}(l^1_{n+1}(z), \ldots, l^{n+1}_{n+1}(z)) = z,$$

$$l^i_{n+1}(c_{n+1}(x_1, \ldots, x_{n+1})) = x_i,\ i \in \overline{1, n+1}.$$

Исходя из нумераций множеств $\mathbb{N}^n_0$, можно установить нумерацию наборов чисел из $\mathbb{N}_0$ всевозможных длин, т. е. нумерацию множества

$$M = \mathbb{N}_0 \cup \mathbb{N}^2_0 \cup \mathbb{N}^3_0 \cup \ldots.$$

Определим на M функцию c, положив для любого $n \in \mathbb{N}$:

$$c(x_1, \ldots, x_n) = c_2(c_n(x_1, \ldots, x_n), n-1),$$

при этом предполагается, что $c_1(x_1) = x_1$. Очевидно, что c — биективное отображение множества M на $\mathbb{N}_0$. При этом если $c(x_1, \ldots, x_n) = z$, то

$$c_n(x_1, \ldots, x_n) = l^1_2(z),$$

где

$$n = l^2_2(z) + 1 \text{ и } x_i = l^i_n(l^1_2(z)).$$

Таким образом, функция c вычислима и осуществляет биективную нумерацию множества M.

Пусть теперь требуется занумеровать множество $W(A)$ в некотором счетном алфавите

$$A = \{a_0, a_1, a_2, \ldots\}.$$

Определим функцию $\nu : W(A) \to \mathbb{N}_0$ следующим образом через функцию c:

$$\nu(a_{i_1}, \ldots, a_{i_s}) = c(i_1, \ldots, i_s) + 1,\ s \geq 1,$$

$$\nu(\Lambda) = 0,$$

для пустого слова Λ. Из биективности функции c легко следует, что ν — биективное отображение $W(A)$ на $\mathbb{N}_0$. Ясно также, что значение функции ν эффективно вычислимо для любого слова

$$a_{i_1}\ a_{i_2}\ \ldots\ a_{i_s} \in W(A)$$

и, наоборот, любое слово однозначно и эффективно восстанавливается по его номеру.

Можно указать также способ нумерации не всех слов в алфавите A, а какого-либо подмножества $B \subset W(A)$ специально устроенных слов при условии, что имеется эффективная процедура выделения нумеруемых слов из всех слов. Для этого занумеруем сначала все слова в алфавите A. Теперь будем перебирать все слова из B в порядке возрастания их номеров и приписывать им новые номера $0,\ 1,\ 2, \ldots$.

В итоге для каждого слова из B найдется вполне определенный номер и несложно построить эффективный алгоритм нахождения слова из B по его номеру в полученной нумерации слов из B.

2) Гёделева нумерация. Пусть

$$A = \{a_0, a_1, a_2, \ldots\}$$

— конечный или счетный алфавит. Сопоставим каждой букве a_i нечетное число $\varphi(a_i) = 2i + 3$ и занумеруем простые числа в порядке их возрастания:

$$p_0,\ p_1,\ p_2, \ldots$$

так, что $p_0 = 2,\ p_1 = 3,\ p_2 = 5,\ \ldots$.

Гёделевым номером непустого слова $P = a_{i_0}\ a_{i_1}\ \ldots\ a_{i_k}$ назовем число

$$\Gamma(P) = p_0^{\varphi(a_{i_0})}\ p_1^{\varphi(a_{i_2})}\ \ldots\ p_k^{\varphi(a_{i_k})}.$$

Гёделевым номером пустого слова Λ назовем число 0. Из однозначности разложения натуральных чисел на простые множители легко следует, что различные слова в алфавите A будут иметь разные гёделевы номера. Кроме того, по слову мы эффективно найдем его гёделев номер, а по любому натуральному числу n сможем узнать, является оно гёделевым номером какого-либо слова или нет, и если является, то найдем это слово.

1.5. РЕКУРСИВНЫЕ ФУНКЦИИ

Как было показано выше, определение алгоритма в алфавите A можно свести к точному определению класса вычислимых арифметических функций. Такое определение было дано в 1936 г. Американский математик А. Чёрч выдвинул гипотезу о том, что класс всех вычислимых функций совпадает с классом так называемых рекурсивных функций, определенных в работах Д. Гильберта, К. Гёделя и самого А. Чёрча (работы 1930–1936 гг.). В 1936 г. соотечественник А. Чёрча С. Клини расширил класс рекурсивных функций до класса частично рекурсивных функций и сформулировал тезис о том, что *класс частично рекурсивных функций совпадает с классом всех частично определенных вычислимых арифметических функций*. Это утверждение называют теперь тезисом Чёрча. Оно, как и принцип Маркова о нормализации алгоритмов или тезис Тьюринга–Поста, не может быть доказано, поскольку в его формулировке присутствует неопределенное понятие вычислимой функции. Тезис Чёрча можно лишь подкреплять различными разумными доводами. Важнейшим из них является доказательство того, что любая частичная арифметическая функция, вычислимая с помощью нормального алгоритма или с помощью машины Тьюринга–Поста, является частично рекурсивной.

Класс частично рекурсивных функций определяется путем указания некоторых исходных функций и операций, позволяющих из одних функций получать другие функции.

В качестве исходных берутся функции:

1) нуль-местная функция-константа 0;

2) одноместная функция $s(x) = x + 1$;

3) $J_m^n(x_1, \ldots, x_n) = x_m, n \in \mathbb{N},\ m \in \overline{1, n}$.

В качестве основных операций выбираются следующие три операции:

1. Операция суперпозиции, или подстановки, позволяющая из n-местной функции (т. е. функции от n переменных) φ и функции $f_1, \ldots, f_n$ получить функцию $f = \varphi(f_1, \ldots, f_n)$, которую обозначают также в виде $S(\varphi, f_1, \ldots, f_n)$. Функция f будет определена при некоторых натуральных значениях ее аргументов, если при этих значениях определены функции $f_1, \ldots, f_n$ и φ определена в точке, заданной соответствующими значениями функций $f_1, \ldots, f_n$. В частности, если функции $\varphi, f_l, \ldots, f_n$ определены всюду (т. е. при любых значениях переменных), то тем же свойством будет обладать и функция f.

2. Операция примитивной рекурсии, позволяющая из n-местной функции g и $(n + 2)$-местной функции h получить $(n + 1)$-местную функцию f, определенную равенствами:

$$f(x_1, \ldots, x_n, 0) = g(x_1, \ldots, x_n),$$

$$f(x_1, \ldots, x_n, y + 1) = h(x_1, \ldots, x_n, y, f(x_1, \ldots, x_n, y)).$$

Функцию f обозначают через $R(g, h)$. Она будет всюду определенной, если всюду определены g и h.

3. Операция минимизации, позволяющая из $(n+1)$-местной функции φ получить n-местную функцию f, определяемую следующим образом. Для натуральных чисел $a_1, \ldots, a_n$ полагаем

$$f(a_1, \ldots, a_n) = b,$$

если $\varphi(a_1, \ldots, a_n, y)$, как функция от y, определена при $y = 0,\ 1,\ \ldots,\ b$, причем

$$\varphi(a_1, \ldots, a_n, y) \neq 0 \text{ при } y = 0,\ 1,\ \ldots,\ b - 1,$$

$$\varphi(a_1, \ldots, a_n, b) = 0.$$

Во всех остальных случаях функция f в точке $(a_l, \ldots, a_n)$ считается не определенной. Функцию f обозначают в виде

$$f(x_1, \ldots, x_n) = \mu_y(\varphi(x_1, \ldots, x_n, y) = 0)$$

или короче — $\mu(\varphi)$.

Определение 1.5. *Минимальный класс частичных арифметических функций, содержащий функции* $0, s(x), J_m^n(x_1, \ldots, x_n)$, $n \in N$, $1 \leq m \leq n$ *и замкнутый относительно операций подстановки, примитивной рекурсии и минимизации, называется классом частично рекурсивных функций и обозначается* $K_{ч.р.}$.

Если в определении $K_{ч.р.}$ *оставить лишь операции подстановки и примитивной рекурсии, то получим определение класса* $K_{п.р.}$ *примитивно рекурсивных функций.*

Класс $K_{о.р.}$ *всюду определенных функций из* $K_{ч.р.}$ *называется классом общерекурсивных функций.*

Функции из классов $K_{п.р.}$, $K_{о.р.}$, $K_{ч.р.}$ *называются соответственно примитивно рекурсивными, общерекурсивными и частично рекурсивными функциями.*

Примеры. 1. Функция $f(x_1, x_2) = x_1 + x_2$ примитивно рекурсивна, так как однозначно определяется равенствами

$$x_1 + 0 = x_1,$$

$$x_1 + (y + 1) = (x_1 + y) + 1.$$

Подробнее, через исходные функции и операции S, R функция $f(x_1, x_2) = x_1 + x_2$ может быть выражена в виде

$$f(x_1, x_2) = R(J_1^1(x_1), S(J_3^3(x_1, x_2, x_3), J_1^1(x_1), J_2^2(x_1, x_2), S(s(x_1), J_3^3(x_1, x_2, x_3))))$$

или короче:

$$f = R(J_1^1, S(J_3^3, J_1^1, J_2^2, S(s, J_3^3))).$$

2. Определим функцию $x_1 \dot{-} x_2$, называемую арифметической (усеченной) разностью:

$$x_1 \dot{-} x_2 = \begin{cases} 0, & \text{если } x_1 \leq x_2 \\ x_1 - x_2, & \text{если } x_1 > x_2 \end{cases}.$$

Функции $x_1 \dot{-} 1$ и $x_1 \dot{-} x_2$ определяются соответственно системами равенств

$$\begin{cases} 0 \dot{-} 1 = 0 \\ (x_1 + 1) \dot{-} 1 = x_1 \end{cases}, \quad \begin{cases} x_1 \dot{-} 0 = x_1 \\ x_1 \dot{-} (x_2 + 1) = (x_1 \dot{-} x_2) \dot{-} 1 \end{cases}.$$

Отсюда следует, что функция $x_1 \dot{-} x_2$ примитивно рекурсивна.

3. Из построенных примеров следует примитивная рекурсивность функций

$$|x_1 - x_2|,\ \min\{x-1, x_2\},\ \max\{x-1, x_2\},$$

поскольку:

$$|x_1 - x_2| = (x_1 \dot{-} x_2) + (x_2 \dot{-} x_1),$$

$$\min\{x-1, x_2\} = x_2 \dot{-} (x_2 \dot{-} x_1),$$

$$\max\{x-1, x_2\} = (x_1 + x_2) \dot{-} \min\{x-1, x_2\}.$$

4. Легко показать, что функции $x_1 x_2$, $x_1^{x_2}$, $x!$, а также

$$sg(x) = \begin{cases} 0, & \text{если } x = 0 \\ 1, & \text{если } x > 0 \end{cases} \quad \text{и} \quad \overline{sg}(x) = 1 - sg(x)$$

примитивно рекурсивны. Читателю предлагается сделать это самостоятельно.

5. Рассмотрим частичную функцию $h(x)$, не определенную при $x > 0$ и равную нулю при $x = 0$. Она частично рекурсивна, поскольку

$$h(x) = \mu_y(x + y = 0).$$

6. Из приведенных выше примеров следует, что частично рекурсивной является функция «разность» $x - y$, которая считается не определенной при $x < y$. Действительно, легко проверить, что

$$x - y = x \dot{-} (h(y \dot{-} x) + 1),$$

где h — функция примера 5. Можно показать, что примитивно рекурсивными будут функции:

$\left[\frac{x}{y}\right]$ — целая часть дроби $\frac{x}{y}$;
$\pi(x)$ — число простых чисел, не превосходящих x;
$p(x)$ — простое число с номером x;
$d(x, y)$ — наибольший общий делитель чисел x, y;
$\nu(x)$ — число различных делителей числа x;
$\sigma(x)$ — сумма делителей числа x;
$r(x, y)$ — остаток от деления x на y и т. д. (см. [38]).

Отметим некоторые простейшие факты о функциях классов $K_{\text{п.р.}}$, $K_{\text{о.р.}}$, $K_{\text{ч.р.}}$.

1. Каждая примитивно рекурсивная функция является всюду определенной. Это следует из того, что свойством «быть всюду определенной» обладает каждая из исходных функций $0, s(x), J_m^n$, а операции подстановки и примитивной рекурсии сохраняют указанное свойство.

2. $K_{\text{п.р.}} \subset K_{\text{о.р.}} \subset K_{\text{ч.р.}}$ (очевидно).

3. Любая частично рекурсивная функция f вычислима (в интуитивном смысле), т. е. в каждой из точек, в которой функция f определена, ее значение может быть эффективно вычислено. Действительно, свойством «быть вычислимой» обладают исходные функции $0, s(x), J_m^n$, и легко видеть, что операции подстановки, примитивной рекурсии и минимизации сохраняют указанное свойство.

4. Существует всюду определенная вычислимая, но не примитивно рекурсивная функция.

Первый пример такой функции был приведен учеником Д. Гильберта В. Аккерманом в 1928 г. Он построил последовательность вычислимых функций от двух переменных:

$$P_0(x,y),\ P_1(x,y),\ P_2(x,y),\ldots,$$

в которой $P_0(x,y) = x+y$, $P_1(x,y) = x\cdot y$, а $P_n(x,y)$ при $n \geq 1$ определяется равенствами

$$P_{n+1}(x,0) = 1,\ P_{n+1}(x,y+1) = P_n(x, P_{n+1}(x,y)),$$

и доказал, что функция $A(x) = P_x(2,x)$ растет быстрее любой одноместной примитивно рекурсивной функции, т. е.

$$\forall f(x) \in K_{\text{п.р.}},\ \exists a \in \mathbb{N}_0,\ \forall x \geq a : f(x) < A(x).$$

Вычислимая, но не примитивно рекурсивная функция $A(x)$ называется теперь функцией Аккермана. Несколько позднее (1935–1936) существование вычислимых, но не примитивно рекурсивных функций другим путем доказала венгерский математик Р. Петер, воспользовавшись теоремой американского математика Р. Робинсона о том, что класс одноместных

примитивно рекурсивных функций совпадает с минимальным классом одноместных функций, содержащим функции

$$s(x) = x + 1, q(x) = x \dot{-} [\sqrt{x}]^2,$$

где квадратные скобки означают взятие целой части числа, и замкнутым относительно операций сложения, композиции и итерации. Из этой теоремы Робинсона следует, что все примитивно рекурсивные функции от одного переменного x можно эффективно занумеровать (возможно, с повторениями):

$$f_0(x), f_1(x), f_2(x), \ldots .$$

Функция $f(x, y) = f_x(y)$, очевидно, вычислима. Если бы она была примитивно рекурсивной, то таковой же была бы и функция $\varphi(x) = f(x, x) + 1$ и потому нашлось бы такое k, что при всех $x \in \mathbb{N}_0$

$$f(x, x) + 1 = f_k(x).$$

Отсюда при $x = k$ мы получили бы противоречивое равенство

$$f(k, k) + 1 = f(k, k).$$

Определенная выше функция $f(x, y)$ называется универсальной функцией для класса одноместных примитивно рекурсивных функций, а метод построения функции $f(x) = f_x(x) + 1$ — методом диагонализации. Он широко используется не только в теории рекурсивных функций. Можно доказать, что построенная универсальная функция общерекурсивна.

1.6. ПРИМЕРЫ АЛГОРИТМИЧЕСКИ НЕРАЗРЕШИМЫХ ПРОБЛЕМ

1. Проблема распознавания самоприменимости алгоритмов

Первые простейшие примеры на доказательство отсутствия алгоритма для решения того или иного класса задач были получены в самой теории алгоритмов. Проиллюстрируем это на проблеме распознавания самоприменимости алгоритмов. Эту

проблему можно ставить при всех рассмотренных выше подходах к определению алгоритма. Начнем с нормальных алгоритмов. Пусть $\mathfrak{A}$ — НА в алфавите $A = \{a_1, \ldots, a_n\}$. Для простоты ограничимся случаем $n \geq 4$. Выпишем в строчку все формулы схемы НА $\mathfrak{A}$ в их естественном порядке, заменяя в них $\to$ на a_{n+1}, $\cdot$ на a_{n+2} и отделяя формулу от формулы буквой a_{n+3}. В итоге мы получим слово в алфавите

$$B = A \cup \{a_{n+1}, a_{n+2}, a_{n+3}\},$$

называемое изображением НА $\mathfrak{A}$. Обозначим его через $\mathfrak{A}^m$. Теперь определим отображение τ множества $W(B)$ в $W(A_0)$, где $A_0 = \{a_1, a_2\}$, положив:

$$\tau(a_i) = a_1 a_2^i a_1,\ i \in \overline{1, n+3},\ \tau(\Lambda) = \Lambda;$$

$$\tau(a_{i_1} \ldots a_{i_k}) = \tau(a_{i_1}) \ldots \tau(a_{i_k}),\ k \geq 1.$$

Тогда $\tau(\mathfrak{A}^m)$ будет словом в алфавите A_0. Его называют записью НА $\mathfrak{A}$, обозначим его в виде $\mathfrak{A}$. Так как $A_0 \subset A$, то можно ставить вопрос: применим алгоритм $\mathfrak{A}$ к своей записи или нет?

Определение 1.6. *Нормальный алгоритм $\mathfrak{A}$ в алфавите A называется самоприменимым, если он применим к своей записи. В противном случае $\mathfrak{A}$ называется несамоприменимым.*

Так как существуют НА в алфавите A, применимые ко всем словам в A и не применимые ни к одному из слов, то все НА в алфавите A разбиваются на два непустых класса (самоприменимые и несамоприменимые).

Возникает естественная проблема, называемая проблемой распознавания самоприменимости НА в алфавите A. Ее можно сформулировать так:

существует ли НА $\mathfrak{L}$, который для любого НА $\mathfrak{A}$ в алфавите A распознает, самоприменим $\mathfrak{A}$ или нет.

Ниже мы докажем, что искомого НА $\mathfrak{L}$ не существует. Доказательство таких утверждений (об отсутствии алгоритмов) проводится методом сведения к другим утверждениям такого же рода. В итоге все теоремы об отсутствии алгоритмов

сводятся к одной. Первой такой теоремой об отсутствии НА является

Теорема 1.2. *Не существует НА над алфавитом A, применимого к записям тех и только тех НА в алфавите A, которые несамоприменимы.*

□ Случай 1. Пусть существует НА $\mathfrak{L}$ в алфавите A, такой что для любого НА $\mathfrak{A}$ в алфавите A:

$$\mathfrak{L} \text{ применим к } \mathfrak{A}^r \Leftrightarrow \mathfrak{A} \text{ не применим к } \mathfrak{A}^r. \qquad (1.8)$$

Взяв теперь в качестве $\mathfrak{A}$ алгоритм $\mathfrak{L}$ (ведь $\mathfrak{A}$ — любой в алфавите A и $\mathfrak{L}$ в алфавите A), получим противоречие:

$$\mathfrak{L} \text{ применим к } \mathfrak{L}^r \Leftrightarrow \mathfrak{L} \text{ не применим к } \mathfrak{L}^r.$$

Следовательно, наше допущение неверно, т. е. такого НА $\mathfrak{L}$ не существует.

Случай 2. Пусть $\mathfrak{L}$ есть НА в алфавите $B = A \cup \{b_1, \ldots, b_k\}$, удовлетворяющий условию (1.8). Построим перевод τ_1 слов из $W(B)$ в слова в алфавите $A_1 = \{a_1, a_2, a_3, a_4\}$, положив:

$$\tau_1(\Lambda) = \Lambda,\ \tau_1(a_i) = a_i,\ i \in \{1, \ldots, n\},$$

$$\tau_(b_j) = a_3\, a_4^j\, a_3,\ j \in \overline{1,k}$$

и

$$\tau_1(c_1 c_2 \ldots c_m) = \tau_1(c_1)\tau_1(c_2)\ldots\tau_1(c_m)$$

для любого слова $c_1 c_2 \ldots c_m \in W(B)$.

Если левые и правые части всех формул НА $\mathfrak{L}$ заменить их образами при τ_1 (т. е. переводами), то получим схему НА $\mathfrak{L}'$ в алфавите A_1. Легко видеть, что для любого слова $P \in W(B)$

$$\mathfrak{L}'(\tau_1(P)) \simeq \tau_1(\mathfrak{L}(P)).$$

В частности, для $P \in W(A_0)$ имеем $\tau_1(P) = P$ и потому

$$\mathfrak{L}'(P) \simeq \tau_1(\mathfrak{L}(P)). \qquad (1.9)$$

А так как отображение τ_1 применимо к любому слову $P \in$ $\in W(B)$, то из (1.9) следует, что НА $\mathfrak{L}$ и $\mathfrak{L}'$ применимы к одним и тем же словам из $W(A_0)$. Следовательно,

$$\mathfrak{L}' \text{ применим к } \mathfrak{A}^r \Leftrightarrow \mathfrak{L} \text{ не применим к } \mathfrak{A}^r$$

и потому $\mathfrak{L}'$, как и $\mathfrak{L}$, удовлетворяет условию (1.8). Но $\mathfrak{L}'$ — НА в алфавите A, и мы получаем противоречие с доказанным в случае 1. Значит, НА $\mathfrak{L}'$, а потому и $\mathfrak{L}$, не существуют. □

Теперь можно решить поставленную выше проблему распознавания самоприменимости. Если решающий ее алгоритм существует, то он каким-то образом должен отличать самоприменимые алгоритмы от несамоприменимых. Не теряя общности, можно считать, что он перерабатывает записи самоприменимых алгоритмов в пустое слово Λ, а записи несамоприменимых — в слово, отличное от Λ.

Следствие 1. *Если $A_1 \subset A$, то невозможен НА $\mathfrak{L}$ над A_0, применимый к записям всех НА в алфавите A и такой, что*

$$\mathfrak{L}(\mathfrak{A}^r) = \Lambda \Leftrightarrow \mathfrak{A} \text{ — } \textit{самоприменим.} \tag{1.10}$$

□ Вопреки утверждению теоремы допустим, что НА $\mathfrak{L}$ существует, и пусть B — его алфавит. Рассмотрим два вспомогательных НА $\mathfrak{L}_1$, $\mathfrak{L}_2$:

$$(\mathfrak{L}_1) \quad \alpha x \to \alpha,\ x \in B$$

$$x \to \alpha,\ x \in B$$

$$\alpha \to \cdot\Lambda$$

$$\Lambda \to \cdot a_1,$$

$$(\mathfrak{L}_2) \quad x \to x,\ x \in B,$$

где α — некоторый новый символ.

Из схем видно, что $\mathfrak{L}_2$ применим только к пустому слову Λ, а $\mathfrak{L}$ применим ко всем словам в алфавите B, причем

$\mathfrak{L}(\Lambda) = a_1$ и $\mathfrak{L}(P) = \Lambda$ для любого непустого слова $P \in$ $\in W(B)$.

Возьмем теперь композицию НА

$$\mathfrak{L}' = \mathfrak{L}_2 \circ \mathfrak{L}_1 \circ \mathfrak{L}.$$

По определению композиции имеем

$$\mathfrak{L}'(P) \simeq \mathfrak{L}_2(\mathfrak{L}_1(\mathfrak{L}(P))),\ P \in W(A),$$

и, в частности,

$$\mathfrak{L}'(\mathfrak{A}^r) \simeq \mathfrak{L}_2(\mathfrak{L}_1(\mathfrak{L}(\mathfrak{A}^r))).$$

Из свойств алгоритмов $\mathfrak{L}$, $\mathfrak{L}_1$, $\mathfrak{L}_2$ видно, что

$$\mathfrak{L}' \text{ применим к } \mathfrak{A}^r \Leftrightarrow \mathfrak{A} \text{ не самоприменим.}$$

Однако такого алгоритма не существует по теореме 1.2. Значит, не существует и $\mathfrak{L}$. □

Аналоги понятий самоприменимого и несамоприменимого НА можно ввести и для других определений алгоритма. Покажем, как это сделать для машин Тьюринга и частично рекурсивных функций. Используя идею перевода слов, можно, не теряя общности, ограничиться МТ с внешним алфавитом $\Omega = \{0,\ 1\}$. В качестве внутренних алфавитов таких МТ будем выбирать начальные отрезки счетного множества символов

$$Q = \{q_0, q_1, q_2, \ldots\}.$$

Записывая команды любой МТ $\mathfrak{M}$ в строчку одну за другой и отделяя соседние команды буквой γ, мы получим слово в алфавите

$$A' = A \cup Q \cup \{\gamma\} \cup \{R, L, H\}.$$

В итоге получим счетное множество МТ.

Теперь осуществим какую-либо эффективную нумерацию всех МТ с указанными алфавитами:

$$\mathfrak{M}_0,\ \mathfrak{M}_1,\ \mathfrak{M}_2, \ldots .$$

Назовем МТ $\mathfrak{M}_x$ самоприменимой, если она применима к своему номеру x. В противном случае МТ $\mathfrak{M}_x$ называется несамоприменимой.

Теорема 1.3. *Не существует МТ, применимой к номерам тех и только тех МТ, которые несамоприменимы.*

□ Допустим, такая МТ $\mathfrak{M}$ существует: $\mathfrak{M}$ применима к $x \in \mathbb{N}$ тогда и только тогда, когда $\mathfrak{M}_x$ несамоприменима. Тогда если y — номер МТ $\mathfrak{M}$, то при $x = y$ получаем противоречие: $\mathfrak{M}_y$ самоприменима тогда и только тогда, когда $\mathfrak{M}_y$ несамоприменима.

Следствие 1. *Не существует МТ $\mathfrak{M}$, применимой ко всем номерам всех МТ и такой, что $\mathfrak{M}(x) = 0$ тогда и только тогда, когда $\mathfrak{M}_x$ самоприменима.*

Это утверждение доказывается по той же схеме, что и следствие теоремы 1.2, с использованием двух вспомогательных МТ. Проведите подробное доказательство в качестве упражнения.

Если согласиться с тезисом Тьюринга–Поста, то следствие теоремы 1.3 можно сформулировать следующим образом:

Не существует алгоритма, позволяющего для каждой машины Тьюринга выяснять, самоприменима она или нет.

Аналогично, занумеровав натуральными числами все одноместные частично рекурсивные функции, мы определим самоприменимую (несамоприменимую) функцию как функцию, определенную (не определенную) на своем номере. Точно так же, как и для НА и МТ, можно сформулировать и отрицательно решить проблему распознавания самоприменимости для частично рекурсивных функций. Читателю предлагается проделать это в качестве упражнения.

2. Проблема тождества слов в полугруппе

Пусть даны произвольный алфавит A и множество $S = \{(P_i, Q_i) : i \in I\}$ пар слов в этом алфавите. На множестве $W(A)$ всех слов в алфавите A введем отношение эквивалентности.

Элементарным преобразованием слова P назовем:

а) замену в P любого одного вхождения слова P_i словом Q_i;

б) замену в P любого одного вхождения слова Q_i словом P_i.

Если слово Q получено одним элементарным преобразованием из P, то будем писать $P \to Q$. Слова P, Q из $W(A)$

называются эквивалентными, если существует конечная последовательность слов R_i, такая что

$$P = R_0 \to R_1 \to \ldots \to R_k = Q.$$

Эквивалентность слов P, Q обозначим $P \equiv Q$. Легко видеть, что введенное отношение $\equiv$ является отношением эквивалентности на множестве $W(A)$, и потому $W(A)$ разбивается на непересекающиеся классы эквивалентных слов. Класс слов, содержащий слово P, обозначим через $[P]$. На множестве классов Π введем операцию умножения, положив

$$[P] \cdot [Q] = [PQ].$$

Очевидно, что определение корректно и введенная операция ассоциативна. Следовательно, множество классов Π есть полугруппа. Она называется *полугруппой, заданной системой образующих элементов A и системой определяющих соотношений S*, и обозначается в виде

$$\Pi =< A, S > . \tag{1.11}$$

Заметим, что таким образом с точностью до изоморфизма может быть задана любая полугруппа. Например, в качестве A можно взять множество всех элементов полугруппы, а в качестве S — всю ее таблицу Кэли.

Полугруппа Π называется *конечно порожденной*, если ее можно задать в виде (1.11) с конечной системой образующих A, и *конечно определенной*, если она задается в виде (1.11) с конечными A и S.

Далее мы будем рассматривать только конечно определенные полугруппы.

Задание (1.11) зачастую бывает удобным, поскольку позволяет компактно записывать и хранить всю информацию о полугруппе Π, которая в общем случае может быть и бесконечной. Однако, чтобы задание (1.11) было эффективным и чтобы мы могли пользоваться им для решения каких-либо вопросов о полугруппе Π, необходимо, конечно, в первую очередь уметь эффективно распознавать равенство элементов в

полугруппе П, или, что то же самое, эквивалентность слов P, Q в алфавите A. Эта задача и называется *проблемой равенства (или тождества) слов в* П. Она формулируется следующим образом:

Найти алгоритм, позволяющий для любых двух слов в алфавите A выяснить, эквивалентны эти слова в П *или нет.*

Пример. Рассмотрим полугруппу

$$\Pi =< \{a, b, c\}; \{(ab, ba)\ ,\ (ac, ca)\ ,\ (bc, cb)\} > .$$

Нетрудно видеть, что П есть абелева полугруппа и слова P, Q в ней эквивалентны тогда и только тогда, когда содержат одинаковое число вхождений каждой из букв a, b, c. Следовательно, в полугруппе П проблема равенства слов алгоритмически разрешима. Однако существуют конечно определенные полугруппы с неразрешимой проблемой равенства слов. Такие полугруппы впервые (1946–1947) независимо были указаны А. А. Марковым и Э. Постом. Построенные ими полугруппы имели весьма громоздкие задания в виде (1.11). Позднее были найдены полугруппы с неразрешимой проблемой равенства слов, имеющие более простые задания. Таким примером может служить найденная отечественным математиком Г. С. Цейтиным полугруппа с системой образующих $A = \{a, b, c, d, e\}$ и с системой определяющих соотношений

$$S = \{(ac, ca)\ ,\ (ad, da)\ ,\ (bc, cb)\ ,\ (bd, db)\ ,\ (abac\ ,\ abaee)\ ,\ (eca, ae)\ ,\ (edb, be)\}.$$

Сразу отметим, что А. А. Марков и Э. Пост строили свои полугруппы в явном виде и доказывали соответственно отсутствие нормального алгоритма и отсутствие машины Поста для решения проблемы равенства слов в построенной полугруппе. Их построения и доказательства весьма сложны, и мы не имеем возможности их изложить здесь.

Мы ограничимся лишь доказательством существования полугрупп с неразрешимой проблемой равенства слов, при этом будем существенно использовать тезис Тьюринга–Поста.

Начнем с того, что каждой МТ $\mathfrak{M}$ сопоставим конечно определенную полугруппу $\Pi_{\mathfrak{M}}$.

Пусть

$$Q = \{q_0, q_1, \ldots, q_m\},\ A = \{a_0, a_1, \ldots, a_n\}$$

— внутренний и внешний алфавиты МТ $\mathfrak{M}$. В качестве системы образующих элементов искомой полугруппы $\Pi_{\mathfrak{M}}$ возьмем множество

$$A_1 = Q \cup A \cup \{h\},$$

где h — новая буква, она будет указывать на край ленты МТ $\mathfrak{M}$. Система определяющих соотношений S строится следующим образом. Сначала сопоставим каждой команде МТ подходящую систему соотношений. Команде вида

$$q_i a_j \to q_s a_t H$$

сопоставим соотношение

$$(q_i a_j, q_s a_t), \tag{1.12}$$

команде вида

$$q_i a_j \to q_s a_t L$$

сопоставим систему соотношений

$$\begin{cases} (a_r q_i a_j, q_s a_r a_t),\ r \in \overline{0,n} \\ (h q_i a_j, h q_s a_0 a_t) \end{cases} \tag{1.13}$$

и команде вида

$$q_i a_j \to q_s a_t R$$

сопоставим систему соотношений

$$\begin{cases} (q_i a_j a_r, a_t q_s a_r),\ r \in \overline{0,n} \\ (q_i a_j h, a_t q_s a_0 h) \end{cases}. \tag{1.14}$$

В качестве S возьмем объединение соотношений, сопоставленных всем командам МТ $\mathfrak{M}$. В итоге получим полугруппу

$$\Pi_{\mathfrak{M}} =< A_1, S > .$$

Мы знаем, что МТ осуществляет переработку машинных слов. Каждое машинное слово имеет вид

$$\hat{P} = a_{i_1} \dots a_{i_{r-1}} q_j a_{i_r} \dots a_{i_k}.$$

Сопоставим ему слово

$$\tilde{P} = h\hat{P}h,$$

которое назовем обобщенным машинным словом (отсюда видно, что буква h заменяет края ленты).

Лемма 1.4. *Если $\mathfrak{M}$ переводит P в Q, то $\tilde{P} \equiv \tilde{Q}$ в $\Pi_{\mathfrak{M}}$.*

□ Утверждение леммы очевидно, поскольку все замены (элементарные преобразования) машинных слов, которые может осуществлять МТ $\mathfrak{M}$, могут быть произведены над обобщенными машинными словами в $\Pi_{\mathfrak{M}}$. □

Лемма 1.5. *Если $\tilde{P}$ и $\tilde{Q}$ — обобщенные машинные слова, $\tilde{Q}$ содержит букву q_m и $\tilde{P} \equiv \tilde{Q}$ в $\Pi_{\mathfrak{M}}$, то $\tilde{Q}$ может быть получено из $\tilde{P}$ с использованием лишь элементарных преобразований типа а) (т. е. замен лишь левых частей соотношений на правые части).*

□ Так как $\tilde{P} \equiv \tilde{Q}$ в $\Pi_{\mathfrak{M}}$, то Q может быть получено из P конечной цепочкой элементарных преобразований:

$$\tilde{P} = R_0 \to R_1 \to \dots \to R_k = \tilde{Q}. \tag{1.15}$$

Докажем наше утверждение индукцией по k. Если $k = 0$, то $\tilde{P} = \tilde{Q}$ и утверждение верно. Пусть оно верно для $k \le l$, докажем для $k = l+1$. Так как $\tilde{Q}$ содержит q_m, и q_m не входит в левые части соотношений из S, то при переходе $R_{k-1} \to \to R_k$ использовалось преобразование типа *а)*. Если в цепочке (1.15) есть преобразование типа *б)*, то найдется такое α, что в цепочке

$$R_\alpha \to R_{\alpha+1} \to R_{\alpha+2}$$

используется сначала преобразование типа *б)*, затем преобразование типа *а)*. Пусть переход от $R_{\alpha+1}$ к $R_{\alpha+2}$ осуществлен с помощью соотношения (1.12). Тогда

$$R_{\alpha+1} = R' q_i a_j r'', \; R_{\alpha+2} = r' q_s a_t R'',$$

и при переходе от R_α к $R_{\alpha+1}$ должно использоваться соотношение, левая часть которого содержит подслово $q_i a_j$. Однако из построения S видно, что такое соотношение единственно и совпадает с тем же соотношением (1.12). Отсюда следует, что $R_\alpha = R_{\alpha+2}$, т. е. слово $\tilde{Q}$ можно получить из $\tilde{P}$ с помощью цепочки из $k-2$ элементарных преобразований, и утверждение леммы верно по предположению индукции. Аналогичны рассуждения и в тех случаях, когда при переходе от R_α к $R_{\alpha+1}$ используется одно из соотношений вида (1.13) или (1.14). □

Из лемм 1.4, 1.5 очевидным образом получаем

Следствие 1. *Если $\hat{P}, \hat{Q}$ — машинные слова и Q содержит букву q_m, то*

$$\tilde{P} \sim \tilde{Q} \text{ в } \Pi_{\mathfrak{M}} \Leftrightarrow \mathfrak{M}(P) = Q.$$

Теорема 1.6. *Существует конечно определенная полугруппа с неразрешимой проблемой равенства слов.*

□ Обозначим через B множество номеров самоприменимых машин:

$$B = \{x : \mathfrak{M}_x \text{ самоприменима}\}.$$

Неразрешимость проблемы самоприменимости для МТ означает, что не существует такой МТ $\mathfrak{M}$, которая бы распознавала принадлежность чисел к множеству B или $\overline{B} = \mathbb{N}_0 \setminus B$, т. е. такой, что

$$\mathfrak{M}(x) = \begin{cases} 1, & \text{если } x \in B \\ 0, & \text{если } x \in \bar{B} \end{cases}.$$

Определим частичную арифметическую функцию $f(x)$, полагая ее не определенной при $x \in \bar{B}$ и равной 1 при $x \in B$. Вычислимость $f(x)$ легко следует из ее определения, поэтому согласно тезису Тьюринга–Поста существует МТ $\mathfrak{M}'$, вычисляющая ее значения.

$$\mathfrak{M}'(x) = 1, \text{ если } x \in B,$$

$$\mathfrak{M}' \text{ не применима к } x \in \bar{B}.$$

По МТ $\mathfrak{M}'$ построим полугруппу $\Pi_{\mathfrak{M}'}$. Она и будет искомой конечно определенной полугруппой с неразрешимой проблемой равенства слов. Действительно, если бы в $\Pi_{\mathfrak{M}'}$ была разрешима проблема равенства слов, то существовал бы алгоритм, позволяющий распознавать эквивалентность обобщенных машинных слов вида

$$\tilde{P}_x = hq_0\underbrace{11\ldots1}_{x}h \text{ и } \tilde{Q} = hq_m1h.$$

Однако по следствию из лемм 1.4, 1.5

$$\tilde{P}_x \sim \tilde{Q} \Leftrightarrow \mathfrak{M}'(x) = 1.$$

Следовательно, вычислимой является функция

$$q(x) = \begin{cases} 1, & \text{если } x \in B \\ 0, & \text{если } x \in \bar{B} \end{cases}.$$

Тогда по тезису Тьюринга–Поста существует вычисляющая ее МТ, и мы пришли к противоречию. □

Образующими элементами и определяющими соотношениями можно задавать не только полугруппы, но и группы. При этом в качестве слов надо будет рассматривать конечные выражения вида

$$a_{i_1}^{\varepsilon_1}a_{i_2}^{\varepsilon_2}\ldots a_{i_k}^{\varepsilon_k},$$

где a_j — буквы из системы образующих A, $\varepsilon_i = \pm 1$, а к определяющим соотношениям добавить все соотношения вида

$$(a_i\, a_i^{-1}, \Lambda),\ (a_i^{-1}a_i, \Lambda).$$

Так же, как и для полугрупп, можно сформулировать проблему равенства слов для группы, заданной конечными системами образующих элементов и определяющих соотношений. Эта проблема оказалась весьма сложной. Ее в 1952 г. отрицательно решил отечественный математик П. С. Новиков. Ему удалось свести эту проблему к проблеме равенства для полугрупп. При этом он использовал так называемый метод

вложения. А именно, он нашел эффективный способ вложения конечно определенной полугруппы Π с неразрешимой проблемой равенства слов в подходящую конечно определенную группу G. Отсюда и следовала неразрешимость проблемы равенства слов в G. Позднее метод вложения неоднократно использовался для доказательства других алгоритмических проблем.

3. Проблема разрешимости исчисления предикатов

В главе 4 части I были определены понятия формулы и тождественно истинной формулы исчисления предикатов заданной сигнатуры Σ. Естественно возникает *проблема нахождения алгоритма, позволяющего для каждой формулы выяснить, тождественно истинна она или нет*. Эта задача и называется проблемой разрешимости исчисления предикатов сигнатуры Σ. Аналогичная проблема может быть поставлена и для любой элементарной теории. В 1936 г. проблему разрешимости для формальной арифметики отрицательно решил А. Чёрч. По словам А. И. Мальцева, «это был первый крупный успех только родившейся в те годы теории алгоритмов» ([38], с. 291). Отрицательное решение проблемы разрешимости исчисления предикатов при небольшом ограничении на сигнатуру Σ можно получить, исходя из существования конечно определенной полугруппы с неразрешимой проблемой равенства слов.

Теорема 1.7. *Если сигнатура Σ содержит символ двуместной операции $\cdot$ и предикат равенства, то проблема разрешимости исчисления предикатов сигнатуры Σ неразрешима (т. е. не существует алгоритма, позволяющего для любой формулы исчисления выяснять, тождественно истинна она или нет).*

□ Пусть

$$\Pi =< A, S >$$

есть конечно определенная полугруппа с неразрешимой проблемой равенства слов,

$$A = \{x_1, \ldots, x_n\},\ S = \{(P_1, Q_1), \ldots, (P_m, Q_m)\}.$$

Если слова P_i, Q_i рассматривать как термы, построенные с использованием переменных $x_1, \ldots, x_n$ и символа операции $\cdot$, то определяющее соотношение (P_i, Q_i) можно считать формулой $P_i = Q_i$ рассматриваемого исчисления $\mathfrak{L}'$. Легко видеть, что слова P, Q эквивалентны в П тогда и только тогда, когда в $\mathfrak{L}'$ тождественно истинной является формула

$$\forall x_{n+1}, x_{n+2}, x_{n+3}\ ((x_{n+1}(x_{n+2}x_{n+3}) = (x_{n+1}x_{n+2})x_{n+3})\&$$

$$\&P_1 = Q_1\& \ldots \$P_m = Q_m) \to P = Q.$$

Следовательно, если бы существовал алгоритм распознавания истинности формул исчисления $\mathfrak{L}'$, то существовал бы и алгоритм распознавания равенства слов в полугруппе П. А так как последнего по условию нет, то нет и первого. □

Имеется много результатов отрицательного характера по проблемам разрешимости элементарных теорий различных классов алгебраических систем. Так, из существования конечно определенных полугрупп и групп с неразрешимой проблемой равенства слов сразу следует неразрешимость элементарных теорий для классов всех полугрупп и всех групп. Неразрешимыми являются также элементарные теории класса конечных групп (А. И. Мальцев, 1961), класса конечных симметрических групп (Ю. Л. Ершов, 1964). Много интересных результатов по разрешимости и неразрешимости элементарных теорий и вообще по теории моделей опубликовано в основанном в 1962 г. А. И. Мальцевым журнале «Алгебра и логика».

Глава 2

СЛОЖНОСТЬ АЛГОРИТМОВ И ВЫЧИСЛЕНИЙ

Выше отмечалось, что первоначально главная цель теории алгоритмов заключалась в разработке методов доказательства существования или отсутствия алгоритмов решения массовых задач. К настоящему времени такие методы созданы и успешно применены к решению многих практически интересных задач алгоритмического характера. При решении принципиального вопроса о существовании алгоритма авторов, как правило, мало интересовал вопрос о реальной осуществимости алгоритма, была важна лишь его принципиальная осуществимость (при наличии неограниченного времени и объема памяти). Однако для практических приложений вопрос о реализации алгоритма за приемлемое время имеет такое же принципиальное значение, как и вопрос о существовании алгоритма. В связи с этим перед теорией алгоритмов естественно возникла важнейшая для практики задача о сравнении алгоритмов, решающих одну и ту же массовую проблему, и о выборе в каком-то смысле наилучшего алгоритма. Ясно, что для сравнения алгоритмов необходимо введение для них каких-то количественных характеристик. Такие характеристики можно ввести поразному. Можно, например, интересоваться сложностью задания алгоритма — числом формул НА, числом внутренних состояний или суммарной длиной системы команд МТ и т. п. Обо всех таких параметрах принято говорить, что они характеризуют сложность алгоритма. Естественно, что на практике удобнее иметь дело с более простыми алгоритмами.

Другой подход к сравнению алгоритмов основан на оценках времени и памяти, затрачиваемых в процессе работы алго-

ритма. Эти оценки называют оценками временной и емкостной сложности вычислений.

Указанные подходы к оценкам качества алгоритмов существенно различны. Существуют простые по заданию, но очень медленно работающие алгоритмы, и, наоборот, сложные, но «быстрые» алгоритмы. Естественно, что между оценками сложности алгоритмов и вычислений имеются определенные связи, однако связи эти настолько нетривиальны, что указанные характеристики качества алгоритмов обычно изучаются порознь.

Имеется большое число работ отечественных и зарубежных авторов, посвященных оценкам сложности алгоритмов и особенно сложности вычислений. К работам по оценке сложностей алгоритмов относятся, в частности, многочисленные работы по минимизации булевых функций и нахождению асимптотической сложности контактных, функциональных и других типов схем, реализующих булевы функции. На оценках сложности алгоритмов основан развитый в работах А. Н. Колмогорова и П. Мартин-Лёфа алгоритмический подход к определению случайности и количества информации.

Теория сложности вычислений начала разрабатываться с конца 1950-х гг. и интенсивно развивается до настоящего времени. При этом наряду с задачей оценки сложности вычислений по заданному алгоритму большое значение имеет проблема оценки временно́й сложности оптимального алгоритма для решения той или иной массовой задачи. При решении такой проблемы обычно сначала выбирают подходящую модель вычислений. Выбирать в качестве такой модели нормальные алгоритмы не всегда удобно, поскольку различные элементарные шаги (преобразования) в одном и том же НА могут существенно различаться по сложности их реализации и требуемой емкости. Поэтому в качестве модели вычислений зачастую принимается машина Тьюринга или ее модификации, приближающиеся по характеру вычислений к компьютерам.

Ниже мы приведем отдельные результаты по сложности алгоритмов и вычислений. Подробно с результатами и с библиографией работ в этой области можно познакомиться по монографиям [1, 5].

2.1. СЛОЖНОСТЬ НОРМАЛЬНЫХ АЛГОРИТМОВ, ВЫЧИСЛЯЮЩИХ БУЛЕВЫ ФУНКЦИИ

Вопрос о сложности реализации булевых функций нормальными алгоритмами, машинами Тьюринга и автоматами был рассмотрен В. А. Кузьминым в работе [31]. При этом под сложностью НА понималась длина его изображения, а под сложностью МТ и автоматов — число их состояний. Для НА и МТ с k-буквенными алфавитами и для автоматов, вычисляющих булевы функции от n переменных, были введены функции Шеннона

$$L_{HA}(n,k),\ L_{MT}(n,k),\ L_A(n).$$

Величина $L_{HA}(n,k)$ определяется как наибольшая из сложностей самых простых НА в k-буквенном алфавите, вычисляющих булевы функции от n переменных; аналогичный смысл имеют и две другие функции Шеннона. В работе [31] были доказаны асимптотические (при $n \to \infty$) равенства:

$$L_{HA}(n,k) \sim \frac{2^n}{log_2(k)},$$

$$L_{MT}(n,k) \sim \frac{2^n}{n(k-1)},$$

$$L_A(n) = \frac{c(n)2^n}{n},$$

где $c(n)$ — константа из промежутка $[1,2]$.

Для иллюстрации методики получения оценок указанных функций приведем более грубый, но сравнительно просто доказываемый результат А. А. Маркова по оценкам функции $L_{HA}(n,5)$, (см. [42]). При этом, ради сохранения принятых в части II обозначений для табличных заданий булевых функций, нам пришлось внести в формулировку теоремы Маркова и в ее доказательство несущественные изменения.

Теорема 2.1. *Имеют место неравенства:*

$$\frac{1}{3} \cdot 2^n \le L_{HA}(n,5) \le 2^n + 91.$$

□ Справедливость нижней оценки следует непосредственно из количественных соображений. Действительно, изображение НА в алфавите из 5 букв будет словом в алфавите из 8 букв, поскольку к 5 буквам основного алфавита понадобится добавить еще три символа: $\rightarrow, \cdot, \Lambda$. Отсюда следует, что число различных НА сложности, не превосходящей l, будет не более 8^l. В то же время число булевых функций от n переменных равно 2^{2^n}. Значит, для реализации всех таких функций необходимо выбрать $l \geq \frac{1}{3} \cdot 2^n$.

Для получения верхней оценки укажем конкретный НА в алфавите $A = \{0, 1, a, b, c\}$, вычисляющий булеву функцию f от n переменных. При этом будем пользоваться известными, а также некоторыми новыми обозначениями:

$W(A)$ — множество всех слов в алфавите A;

$W_m(A)$ — множество всех слов длины m из $W(A)$;

$l(P)$ — длина слова P в любом алфавите;

$\mid P \mid$ — число $\sum_{i=0}^{m-1} a_i \cdot 2^{m-i-1}$ для слова $P = a_0 a_1 ... a_{m-1}$;

$\overrightarrow{f}$ — вектор-строка значений булевой функции f при расположении наборов значений переменных P по возрастанию числа $\mid P \mid$;

$\gamma_i(P)$ — $i{+}1$-я буква слова P, так что $\gamma_i(a_0 a_1 ... a_{m-1}) = a_i$;

$\Omega = \{0, 1\}$.

В указанных обозначениях имеем: для любой булевой функции от n переменных и любого слова $P \in W_n(\Omega)$ выполняется равенство

$$f(P) = \gamma_{|P|}(\overrightarrow{f}). \tag{2.1}$$

Теперь рассмотрим НА $\mathcal{A}$ в алфавите A со схемой:

$$0a \rightarrow a1,$$

$$1a \rightarrow 0c,$$

$$cx \rightarrow xc,\ x \in \Omega,$$

$$cby \rightarrow ab,\ y \in \Omega,$$

$$a1 \rightarrow a,$$

$$abxy \rightarrow abx,\ x, y \in \Omega,$$

$$ab \to \cdot\Lambda,$$
$$c \to ab\overrightarrow{f},$$
$$\Lambda \to c.$$

Лемма 2.2. *Если* $P, Q \in W(\Omega), l(P) = n, l(Q) > \mid P \mid$, *то*

$$\mathcal{A}(PabQ) = \gamma_{|P|}(Q). \tag{2.2}$$

□ Докажем лемму индукцией по $\mid P \mid$. При $\mid P \mid = 0$ равенство (2.2) проверяется непосредственно. Пусть теперь

$$P = x_1x_2...x_n, \ \mid P \mid > 0, \ Q = \eta_1\eta_2...\eta_t.$$

Рассмотрим два случая.

1) $x_{k+1} = 1, x_{k+2} = ... = x_n = 0; 1 \le k+1 < n$.

Тогда переведем сначала алгоритмом $\mathcal{A}$ слово $PabQ$ в слово

$$P'abQ' = x_1...x_k01...1ab\eta_2...\eta_t$$

и воспользуемся предположением индукции:

$$\mathcal{A}(P'abQ') = \gamma_{|P'|}(Q'). \tag{2.3}$$

Так как

$$\mid P' \mid = \mid P \mid -1, \gamma_{|P'|}(Q') = \gamma_{|P'|+1}(Q) = \gamma_{|P|}(Q)$$
$$\text{и} \ \ \mathcal{A}(PabQ) = \mathcal{A}(P'abQ'),$$

то из (2.3) следует (2.2).

2) $x_n = 1$. В этом случае слово $PabQ$ следует перевести алгоритмом $\mathcal{A}$ сначала в слово

$$P'abQ' = x_1...x_{n-1}0ab\eta_2...\eta_t,$$

а затем, как и в случае 1, воспользоваться предположением индукции. □

Вернемся к доказательству теоремы 2.1. Возьмем произвольное слово $P = x_1x_2...x_n \in W_n(\Omega)$ и применим к нему алгоритм $\mathcal{A}$. Вначале получим цепочку преобразований

$$x_1...x_n \to cx_1...x_n \to ... \to x_1...x_nc \to Pab\overrightarrow{f}.$$

Отсюда, используя лемму и равенство (2.1), будем иметь

$$\mathcal{A}(P) = \mathcal{A}(Pab\overrightarrow{f}) = \gamma_{|P|}(\overrightarrow{f}) = f(P).$$

Следовательно, алгоритм $\mathcal{A}$ вычисляет булеву функцию f. Осталось заметить, что сложность $\mathcal{A}$ (т. е. длина его изображения) равна $2^n + 91$. $\square$

Заметим, что у А. А. Маркова верхняя оценка равна $2^n + 81$.

2.2. СЛОЖНОСТИ ВЫЧИСЛЕНИЙ НА МАШИНАХ ТЬЮРИНГА

Для оценки сложности вычислений на МТ чаще всего используют две функции: сигнализирующую времени и сигнализирующую емкости.

Определение 2.1. *Пусть M есть МТ, предназначенная для переработки слов в алфавите A, и W' — множество всех слов из $W(A)$, к которым применима M.*

Сигнализирующей функцией времени (емкости) МТ M называется функция, определенная на W' со значениями в $\mathbb{N}$, равная на слове $P \in W'$ числу тактов (числу ячеек ленты), необходимых для переработки слова P.

Так, определенные функции обозначаются соответственно t_M и s_M.

По функциям t_M и s_M обычным образом определяются арифметические функции Шеннона

$$T_M(n) = \max t_M(P) \text{ и } S_M(n) = \max s_M(P),$$

где максимум берется по всем словам P длины n из W'. Функции T_M и S_M не определены в точке x_0, если W' не содержит слов длины x_0.

В общем случае функции T_M и S_M могут быть сколь угодно сложными. Чтобы убедиться в этом достаточно рассмотреть машину M, вычисляющую рекурсивную функцию вида

$$f(x) = 2^{2^{\cdots^2}},$$

где в показателе двойка повторяется x раз.

При ее вычислении только на запись результата вычисления понадобятся время и емкость, равные по порядку значению функции $f(x)$. Следовательно, в данном случае функции T_M и S_M имеют порядок, не меньший чем f. Вместе с тем приведенный пример не показателен, поскольку в нем вывод о сложности работы алгоритма делается на основании лишь сложности записи результата. Для практических приложений, по-видимому, более распространенными являются задачи по вычислению предикатов, т. е. задачи, требующие ответов «да» или «нет». Оказывается, что и среди таких задач имеются сколь угодно сложные. Приведем доказательство этого факта для временной сложности, принадлежащее Г. С. Цейтену.

Теорема 2.3. *Для любой вычислимой арифметической функции $f(x)$ существует вычислимый предикат p на множестве всех слов в алфавите $\Omega = \{0, 1\}$, такой что для любой машины Тьюринга, вычисляющей p, выполняется неравенство*

$$t_M(x) > f(x). \tag{2.4}$$

□ Так как нас интересуют машины, вычисляющие предикаты на множестве слов в алфавите Ω, то будем рассматривать машины Тьюринга с внешним алфавитом $S = \{s_0, s_1, ..., s_m\}$, где $s_0 = 0, s_1 = 1$, а m пробегает значения 1, 2, 3,... . В качестве внутренних алфавитов машин будем брать множества q_0, q_l,..., q_k, $k = 1, 2, 3, ...$, считая q_0 начальным состоянием, а q_k — стоп-состоянием соответствующей машины. Все такие машины Тьюринга можно занумеровать натуральными числами, а записав номер машины в двоичной системе счисления, можно пытаться применять машину к своему номеру.

Назовем машину M_x f-самоприменимой, если

$$M_x(x) = 1 \text{ и } t_{M_x}(x) \leq f(x), \tag{2.5}$$

и f-несамоприменимой в противном случае.

Определим на множестве слов $W(\Omega)$ предикат p, положив для $x \in W(\Omega)$:

$$p(x) = 1 \Longleftrightarrow M_x \text{ — } f\text{-самоприменима.}$$

Из вычислимости функции f легко следует вычислимость предиката p, а потому и его отрицания $\overline{p}$. Следовательно, существуют МТ, вычисляющие значения предиката $\overline{p}$. Пусть M_y — любая из таких машин, т. е. M_y применима к любому слову $x \in W(\Omega)$ и

$$M_y(x) = \overline{p(x)}. \tag{2.6}$$

Докажем от противного, что $t_{M_y}(y) > f(y)$. Допустим, что

$$t_{M_y}(y) \leq f(y). \tag{2.7}$$

Машина M_y не может быть f-самоприменимой, поскольку в этом случае из (2.5) следует $M_y(y) = 1$, а из (2.7) следует $M_y(y) = 0$. Если же M_y f-несамоприменима, то из отрицания условия (2.5) и из неравенства (2.7) следует $M_y(y) = 0$, в то время как из (2.6) получается $M_y(y) = 1$. Таким образом, во всех случаях приходим к противоречию, и потому $t_{M_y}(y) > f(y)$.

Так как при доказательстве последнего неравенства использовалось лишь действие машины M_y на словах в алфавите Ω, то все рассуждения останутся верными при замене M_y машиной M_z, полученной из M_y расширением внешнего алфавита с произвольным доопределением программы. Но любая такая машина M_z перерабатывает Z так же, как и машина M_y. Поэтому имеем

$$t_{M_z}(z) > f(z)$$

для бесконечного множества значений z. □

Заметим, что известна более сильная теорема М. Рабина о существовании предиката p, для которого условие (2.4) выполняется для всех значений x, кроме конечного их множества.

Получение нетривиальных верхних и особенно нижних оценок сложности вычисления конкретных предикатов на МТ чаще всего сопряжено с большими трудностями. Один из применяемых для этой цели методов основан на анализе следов работы машины. Проиллюстрируем его на задаче распознавания симметрии слов в алфавите $\Omega = \{1, 0\}$. С одной стороны,

нетрудно построить МТ, распознающую симметрию слова длины n за время порядка n^2. Для этого достаточно программу машины составить так, чтобы при переработке любого слова P головка машины, двигаясь по ленте, сравнивала 1-ю букву слова P с последней, затем предпоследнюю со 2-й и т. д. В конце работы машина должна стереть слово, оставив лишь символ 1, если P симметрично, и символ 0 в противном случае. Запоминание буквы головкой можно осуществить, например, переходом к новому состоянию машины. Составьте программу такой машины в качестве упражнения.

А с другой стороны, имеет место доказанная впервые Я. М. Барздинем

Теорема 2.4. *Если M — машина Тьюринга с k состояниями, распознающая симметрию слов в алфавите $\Omega = \{1, 0\}$, то*

$$T_M(n) \geq c \cdot n^2,$$

где c — константа, зависящая от M, но не зависящая от n.

□ Пусть в начале работы машины M на ее ленте записано слово

$$P = a_1 a_2 ... a_n.$$

Занумеруем края, или границы, ячеек с буквами a_i числами $0, 1, ..., n$ (слева направо). В ходе работы машины головка будет двигаться вдоль ленты, пересекая границы ленты.

Последовательность внутренних состояний машины, в которых она пересекает границу с номером j в процессе переработки слова P в 0 или 1, называется следом машины в точке j и обозначается через (P, j).

Очевидно, что

$$t_M(P) \geq \sum_{j=0}^{n} l(P, j), \tag{2.8}$$

где $l(P, j)$ — длина следа (P, j).

Займемся оценкой величины $l(P, j)$ при

$$j = \frac{n}{4} + 1, \frac{n}{4} + 2, ..., \frac{n}{2},$$

полагая для простоты n кратным 4.

Лемма 2.5. *Для каждого достаточно большого n найдется симметричное слово P длины n, такое что*

$$l(P,j) \geq c_1 \cdot n; j = \frac{n}{4}+1, \frac{n}{4}+2, ..., \frac{n}{2}, \tag{2.9}$$

где $c_1 = \frac{1}{8} \cdot log_k 2$.

Для доказательства леммы введем обозначения: $S(n)$ — множество всех симметричных слов длины n в алфавите Ω; $N(n)$ — число тех из них, которые не удовлетворяют условию (2.9), и $N(j,n)$ — число слов P из $S(n)$, удовлетворяющих условию $l(P,j) < c_1 \cdot n$.

Докажем сначала неравенство

$$N(j,n) < 2^{3/8}. \tag{2.10}$$

Для этого представим число $N(j,n)$ в виде суммы

$$N(j,n) = L(\sigma_1) + ... + L(\sigma_m), \tag{2.11}$$

где $\sigma_1, ..., \sigma_m$ суть все различные следы машины в точке j, а $L(\sigma_i)$ — число различных слов $P \in S(n)$ со следом $(P,j) = \sigma_i$. Так как след есть слово в k-буквенном алфавите (внутренних состояний машины), то

$$m \leq 1 + k + ... + k^{c_1 n - 1} = \frac{k^{c_1 n} - 1}{k-1} \leq k^{c_1 n} = 2^{n/8}. \tag{2.12}$$

Оценим теперь число $L(\sigma_i)$. Пусть P, Q — два слова из $S(n)$ с одним и тем же следом в точке j. Представим их в виде $P = P_1 P_2, Q = Q_1 Q_2$, где P_1, Q_1 — слова длины j. Так как $(P,j) = (Q,j)$, то при переработке слова $P_1 Q_2$ слово P_1 будет перерабатываться так же, как и в слове , а Q_2, как в Q. Следовательно, $M(P_1 Q_2)$ совпадает с $M(P)$, если машина заканчивает переработку слова левее точки j, и с $M(Q)$ в противном случае. Так как P, Q — симметричны, то $M(P) = M(Q) = 1$ и $M(P_1 Q_2) = 1$, т. е. слово $P_1 Q_2$ симметрично. Отсюда следует, что $P_1 = Q_1$ при всех $j \leq n/2$ и, в частности, при интересующих нас значениях $j = \frac{n}{4}+1, \frac{n}{4}+2, ..., \frac{n}{2}$.

Значит, у слов из $S(n)$ с одним следом в точке j начальные отрезки длины j определены однозначно. Отсюда имеем

$$|L(\sigma_i)| \le 2^{n/2-j},\ j = \frac{n}{4}+1, \frac{n}{4}+2, ..., \frac{n}{2}. \qquad (2.13)$$

Теперь, учитывая (2.11), (2.12), (2.13), получим (2.10).

Из определения величин $N(n), N(j,n)$ и оценки (2.10) следует, что

$$N(n) \le \sum_{j=n/4+1}^{n/2} N(j,n) \le \frac{n}{4} \cdot 2^{3n/8} = n \cdot 2^{3n/8-2}.$$

Отсюда видно, что доля числа симметричных слов P, не удовлетворяющих условию (2.9), стремится к нулю при $n \to \infty$, что и доказывает даже более сильное утверждение, чем лемма. Поэтому требуемое леммой слово P существует, если выполняется неравенство

$$2^{n/2} - n \cdot 2^{3n/8-2} > 0.$$

Легко видеть, что последнее неравенство выполняется при любом $n > 0$, значит, лемма верна.

Теперь утверждение теоремы 2.3 следует непосредственно из леммы и неравенства (2.8). При этом в качестве константы c можно взять число $\frac{1}{16} log_k 2$.

Из теоремы 2.4 и существования МТ, распознающей симметрию слов из $W(\Omega)$ за время порядка n^2, следует, что сложность самой задачи распознавания симметрии слов из $W(\Omega)$ в классе машин Тьюринга имеет порядок n^2.

2.3. КЛАССИФИКАЦИИ ЗАДАЧ ПО СЛОЖНОСТИ ИХ РЕШЕНИЯ НА МАШИНАХ ТЬЮРИНГА

В данном параграфе под задачей всегда будет пониматься задача вычисления предиката на некотором множестве слов в каком-либо конечном алфавите. При этом и предикат, и задача по его вычислению будут обозначаться одной и той же буквой. Интересоваться будем только временно́й сложностью.

Будем говорить, что сложность задачи R ограничена арифметической функцией $T(n)$, если существует решающая задачу R МТ M, такая что

$$T_M(n) \leq T(n),$$

для всех n, в которых определена функция $T_M(n)$.

Из теоремы 2.3 следует существование задач со сколь угодно большой временной сложностью решения на МТ. Однако многие массовые задачи математической логики, комбинаторики, теории графов и т. п. имеют ограниченную сложность. Такие задачи и пытаются классифицировать на менее сложные и более сложные. К первым относят задачи так называемой полиномиальной сложности.

Определение 2.2. *Говорят, что задача R имеет полиномиальную (временную) сложность, если существуют многочлены $f \in \mathbb{Z}[x]$ и МТ M, вычисляющая предикат R, такие что*

$$T_M(n) \leq f(n)$$

для всех $n \in \mathbb{N}$, на которых определена функция T_M.

Например, рассмотренная в параграфе 2.2 задача распознавания симметрии слов в алфавите Ω имеет полиномиальную сложность.

Класс всех задач (по вычислению предикатов) полиномиальной сложности обозначим буквой P.

Для определения еще одного класса задач нам понадобится понятие недетерминированной машины Тьюринга (НМТ). Такая машина отличается от обычной МТ только тем, что в системе ее команд допускается не обязательно одна, а произвольное множество различных команд с одинаковыми левыми частями. Заметим, что это множество конечно в силу конечности внешнего алфавита и множества внутренних состояний машины Тьюринга.

Функционирование НМТ M, вычисляющей предикат R на множестве слов W, определяется так же, как и для обычной МТ, с той лишь разницей, что на каждом такте из всех команд с одинаковой левой частью разрешается выбирать любую

из них. Таким образом, в НМТ существует много различных путей переработки одного и того же входного слова A. Считается, что НМТ M применима к слову A, если в каждом из этих путей она приходит в стоп-состояние. В противном случае M не применима к A. Если M применима к A, то значение $R(A)$ считается истинным, или равным 1, в том и только том случае, когда хотя бы один из путей переработки слова A заканчивается словом 1 или каким-либо его кодом.

Нетрудно заметить, что в определении НМТ отражена широко используемая в практике ЭВМ идея распараллеливания вычислений.

Для НМТ M, как и для МТ, определяются сигнализирующая функция времени $t_M(x)$ (емкости $s_M(x)$), равная числу тактов (числу использованных ячеек ленты) в кратчайшем пути переработки слова x, если $M(x) = 1$, и в самом длинном пути, если $M(x) = 0$. Как и в случае обычной МТ, для НМТ определяются функции Шеннона $T_M(n)$ и $S_M(n)$.

Определение 2.3. *Говорят, что задача R имеет недетерминированно полиномиальную сложность, если существуют НМТ M, решающая задачу R, и многочлен $f \in \mathbb{Z}[x]$, такие что*

$$T_M(n) \le f(n)$$

для всех n, в которых определена функция $T_M(n)$.

Класс всех задач недетерминированно полиномиальной сложности обозначается через NP.

Так как МТ является частным видом НМТ, то $\mathrm{P} \in \mathrm{NP}$. С другой стороны, легко показать, что если сложность вычисления предиката R на НМТ M ограничена вычислимой функцией $T(n)$, то существует МТ M', вычисляющая предикат R с сигнализирующей функцией времени

$$T_{M'}(n) \le c^{T(n)}$$

при подходящей константе c, не зависящей от n. Идея доказательства этого утверждения интуитивно ясна. Систему команд искомой машины M' надо строить так, чтобы при переработке слова x длины n M' перебирала начальные отрезки

длин $l \leq T(n)$ всевозможных путей переработки слова x машиной M.

Из указанного утверждения следует, во-первых, что сложность любой задачи из класса NP ограничена некоторой экспоненциальной функцией c^n, а во-вторых, корректность вопроса о включении класса NP в класс P или, что то же самое, о равенстве этих классов. Вопрос этот на сегодняшний день остается открытым, хотя имеется весьма убедительная гипотеза о том, что NP $\neq$ P.

Подтверждением этой гипотезы служит наличие в классе NP многих практически интересных задач, привлекавших к себе внимание большого числа специалистов и не получивших до настоящего времени решения с полиномиальной сложностью. Примером может служить задача о распознавании разрешимости произвольной булевой системы уравнений или эквивалентная ей задача о распознавании выполнимости формулы алгебры высказываний.

Если в NP существуют задачи не из P, то к ним заведомо будут относиться наиболее сложные задачи из NP. В связи с этим естественно возникает вопрос о выделении в некотором смысле самых сложных задач из класса NP. Так появляется класс NP-полных задач.

Определение 2.4. *Говорят, что задача R_1 полиномиально сводится к задаче R, если по любой МТ M, решающей задачу R, можно построить МТ M_1, решающую задачу R_1 и такую, что*

$$T_{M_1}(n) \leq T_M(f(n))$$

для всех n, на которых определена функция T_{M_1} и для подходящего полинома $f \in \mathbb{N}[x]$.

Определение 2.5. *Задачи R_1 и R называются полиномиально эквивалентными, если каждая из них полиномиально сводится к другой.*

Легко видеть, что полиномиальная эквивалентность является отношением эквивалентности и P содержится в классе самых простых задач в этом смысле.

Определение 2.6. *Задача R называется недетерминированно полиномиально полной, или NP-полной, если она принадлежит классу NP и к ней полиномиально сводится любая задача из NP.*

Таким образом, NP-полные задачи являются с точностью до полиномиальной эквивалентности самыми сложными задачами из класса NP, и если хотя бы одна NP-полная задача имеет полиномиальную сложность, то и все задачи из класса NP имеют полиномиальную сложность, т. е. NP = P. В связи с этим одним из способов обоснования сложности задачи в настоящее время является сведение к ней какой-либо NP-полной задачи. Понятно, что для облегчения реальной осуществимости таких сведений желательно иметь по возможности широкий список NP-полных задач. Много примеров таких задач имеется в монографиях [5, 68]. В последнее время список этот непрерывно пополняется. Доказательство NP-полноты задачи обычно проводится методом сведения к ней какой-либо известной NP-полной задачи. Первым же примером задачи, NP-полнота которой была доказана, непосредственно исходя из определения, явилась задача о распознавании выполнимости формул алгебры высказываний, или булевых формул (см. [5, 68]). Ниже мы приведем схему доказательства соответствующей теоремы.

Предварительно уточним задачу о выполнимости булевых формул.

Под булевой формулой будем понимать приведенную формулу алгебры высказываний. Любую такую формулу можно записать в виде строчки символов в конечном алфавите K, если условиться вместо $\overline{B}$ писать $\neg B$, а каждое переменное высказывание x_i из счетного множества переменных высказываний $x_1, x_2, \ldots$ записывать в виде слова $xa_1 \ldots a_k$, где $a_1 \ldots a_k$ — запись числа i в двоичной системе счисления.

Задачу распознавания выполнимости булевых формул можно рассматривать как задачу вычисления предиката R, определенного на словах из $W(K)$, представляющих булевы формулы, по правилу: $R(A) = 1$, если A — выполнимая формула,

и $R(A) = 0$, если A — невыполнимая формула. Значения предиката на остальных словах из $W(K)$ нас не интересуют.

Теорема 2.6. *Задача распознавания выполнимости булевых формул NP-полна.*

□ Для краткости будем обозначать нашу задачу буквой R. Покажем сначала, что $R \in$ NP, т. е. R решается подходящей НМТ с полиномиальной сложностью. Работу искомой машины M можно представить себе следующим образом. Сначала она в записанной на ее ленте формуле A заменяет символы переменных на 0 и 1. В силу недетерминированности эти замены она может осуществлять параллельно: сначала первое переменное на 0 и на 1, затем в каждом из полученных слов второе переменное на 0 и на 1 и т. д. В итоге формуле A от m переменных сопоставится 2^m слов, полученных из A заменой переменных на 0 и 1. Если слово A имеет длину n, то $m \leq n$ и длина каждого пути переработки слова A в указанные слова не будет превосходить числа $n \cdot k$ при некотором постоянном k, не зависящем от x. Теперь в каждом из полученных 2^m слов нужно произвести все логические операции над элементами 0, 1. Эту работу также можно производить параллельно сразу для всех 2^m вариантов. А так как для одного варианта все вычисления можно произвести за полиномиальное время, то M имеет полиномиальную сложность, и $R \in$ NP.

Осталось доказать самую нетривиальную часть теоремы — полиномиальную сводимость к R любой другой задачи R_1 из NP.

Пусть R_l — предикат, определенный на некотором множестве слов W, и M_1 — НМТ, вычисляющая R_1 с полиномиальной сложностью $T(n)$. Пусть $Q = \{q_0, q_1, ..., q_\tau\}$ — множество внутренних состояний машины, q_0 – начальное состояние, q_τ — заключительное состояние и $S = \{s_0, s_1, ..., s_r\}$ — внешний алфавит НМТ M_1, s_0 — пустой символ.

Будем считать, что в начале работы исходное слово x длины n из W записывается в ячейки с номерами $1, 2, ..., n$, а остальные ячейки заполняются пустым символом s_0.

Несущественным изменением программы машины M_1 можно получить НМТ M_2, переработка слова x в которой

отличается от его переработки в M_1 лишь тем, что после остановки M_1 машина M_2 продолжает без конца работать вхолостую, оставаясь в том же состоянии q_τ, если $R_1(x) = 1$, и перейдя предварительно в некоторое новое заключительное состояние, если $R_1(x) = 0$. Таким образом, для слова $x \in W$ длины n равенство $R_1(x) = 1$ будет выполняться в том и только том случае, когда на некотором пути переработки слова x машиной M_2 машина окажется в состоянии q_τ в момент времени $T(n)$. За это время считывающая головка может отклониться от слова x влево или вправо не более, чем на $T(n)$ ячеек. Следовательно, для нахождения значения предиката $R_1(x)$ нам достаточно наблюдать лишь за ячейками ленты с номерами

$$-f(n), ..., -2, -1, 0, 1, 2, ..., g(n),$$

где $f(n) \leq T(n)$ и $g(n) \leq n + T(n)$.

Условия переработки машиной M_2 слова $x \in W$ длины n и равенство $R_1(x) = 1$ можно выразить в виде следующей последовательности утверждений:

1) в каждый из $T(n)$ тактов головка машины обозревает ровно одну ячейку ленты;

2) в каждый из $T(n)$ тактов работы каждая ячейка ленты содержит ровно один символ из S;

3) в каждом из $T(n)$ тактов машина находится во вполне определенном внутреннем состоянии из Q;

4) при переходе от t-го такта к $(t+1)$-му на ленте может измениться содержимое лишь обозреваемой ячейки;

5) при переходе от t-го такта к $(t+1)$-му положение машины может изменяться только в соответствии с системой команд машины M_2;

6) в начальный момент машина M_2 находится в состоянии q_0, исходное слово $x = s_{j_1}...s_{j_n}$ записано в n первых ячейках ее ленты и головка обозревает 1-ю ячейку;

7) в такте с номером $T(n)$ машина находится в заключительном состоянии q_τ.

Запишем всю эту информацию с помощью булевых формул. С этой целью введем элементарные высказывания:

$s_{i,j,t}$ — в момент t i-я клетка ленты содержит символ s_j;

$q_{k,t}$ — в момент t машина находится в состоянии q_k;

$h_{i,t}$ — в момент t головка обозревает i-ю ячейку ленты.

Для сокращения последующих записей условимся обозначать в виде

$$U(x_1, ..., x_r)$$

формулу

$$(\vee_{i=1}^{r} x_i)\&_{1\le i\ne j\le r}(\overline{x_i} \vee \overline{x_j}).$$

Легко видеть, что эта формула принимает значение 1 в том и только том случае, когда значение 1 имеет ровно одно из элементарных высказываний $x_1, ..., x_r$.

При указанных обозначениях утверждения 1–7 запишутся соответственно формулами:

1) $A = \&U(h_{-f(n),t}, h_{-f(n)+1,t}, ..., h_{g(n),t})$;
2) $B = \&U(s_{i,0,t}, s_{i,1,t}, ..., s_{i,r,t})$;
3) $C = \&U(q_{0,t}, q_{1,t}, ..., q_{\tau,t})$;
4) $D = \&(s_{i,j,t} = s_{i,j,t+1} \vee h_{i,t})$;
5) $E = \&(s_{i,j,t}\&h_{i,t}\&q_{k,t} \to \vee s_{i,j,t}\&h_{i,t}\&q_{k,t})$;
6) $F = q_{0,0}\&h_{1,0}\&(\bigwedge s_{i,j_i}, 0)\&(\bigwedge s_{i',0,0})$;
7) $G = q_{\tau,T(n)}$.

Здесь:

в формулах 1–5 индексы конъюнкции берутся по всем формулам, в которых соответствующие индексы i, j, t, k независимо пробегают соответственно множества

$$\{-f(n), -f(n)+1, ..., g(n)\}; \{0, 1, ..., r\};$$
$$\{0, 1, ..., T(n)\}; \{0, 1, ..., \tau\};$$

в формуле 5 все переходы от момента времени t к $t+1$ осуществляются в соответствии с системой команд машины M_2 и индекс l пробегает по всем шагам машины M_2, возможным в случае, когда M_2 находится в состоянии q_k и обозревается i-я ячейка с записанным в ней символом s_j;

в формуле 6 индекс i пробегает множество $\{1, ..., n\}$, а i' — все остальные значения из промежутка

$$\{-f(n), -f(n)+1, ..., g(n)\}.$$

Легко видеть, что $R_1(x) = 1$ тогда и только тогда, когда выполнима формула

$$\Phi = A\&B\&C\&D\&E\&F\&G.$$

Таким образом, задача вычисления предиката $R_1(x)$ свелась к задаче записи формулы Φ и распознавания ее выполнимости. А так как $g(n)$ и $f(n)$ полиномы, то длина формулы Φ также будет ограничена некоторым полиномом и ее можно записать на ленту за полиномиально ограниченное время. Отсюда и следует, что если бы задача распознавания выполнимости формул имела полиномиальную сложность, то аналогичное утверждение было бы верно и для задачи R_1. □

Заметим, что A, B, C, F, G являются конъюнктивными нормальными формами (КНФ), а КНФ для формул D, E легко получаются с помощью равносильностей алгебры высказываний с увеличением длин формул не более чем в некоторое постоянное число раз. В частности, равенство $x = y$ можно записать формулой $(\overline{x} \vee y)\&(x \vee \overline{y})$. Отсюда и из доказанной теоремы получаем

Следствие 1. *Задача распознавания выполнимости КНФ алгебры высказываний NP-полна.*

Путем сведения к задаче выполнимости КНФ сравнительно просто устанавливается NP-полнота многих комбинаторных задач. Приведем в качестве примера доказательство NP-полноты задачи о три-выполнимости булевых формул. Задача заключается в распознавании выполнимости формул алгебры высказываний представленных в виде конъюнкций элементарных дизъюнкций длины 3.

Теорема 2.7. *Задача три-выполнимости булевых формул NP-полна.*

□ Заметим, что путем введения новых переменных любой элементарной дизъюнкции A длины $m \neq 3$ можно сопоставить формулу B, представленную в виде конъюнкции элементарных дизъюнкций длины 3 и выполнимую в том и только том случае, когда выполнима A. Действительно, непосредственной

проверкой убеждаемся, что в случаях, когда A имеет длину 1, 2, $k > 3$, в качестве B можно взять соответственно формулы:

$B = (A \vee y \vee z)\&(A \vee \overline{y} \vee z)\&(A \vee y \vee \overline{z})\&(A \vee \overline{y} \vee \overline{z})$;

$B = (A \vee y)\&(A \vee \overline{y})$;

$B = (A_1 \vee A_2 \vee y_1)\&(A_3 \vee \overline{y}_1 \vee y_2)\&(A_4 \vee \overline{y}_2 \vee y_3)\&...\&(A_{k-2} \vee \vee \overline{y}_{k-4} \vee y_{k-3})\&(A_{k-1} \vee A_k \vee \overline{y}_{k-3})$;

где $y, z, y_1, ..., y_{k-3}$ — новые переменные, $A_1, ..., A_k$ — слагаемые (литералы) дизъюнкции A при $k > 3$.

Пусть теперь F — любая КНФ. Заменив в ней каждую элементарную дизъюнкцию A длины $l \neq 3$ соответствующей ей формулой B, мы получим формулу F', являющуюся конъюнкцией элементарных дизъюнкций длины 3, которая выполнима в том и только том случае, когда выполнима F. Тем самым проблема выполнимости булевых формул сведена к проблеме три-выполнимости. При этом сведение — полиномиально, поскольку переход от F к F' осуществим с полиномиальной сложностью, а длина формулы F' не более чем в 3 раза больше длины формулы F. □

Замечание. Естественно возникает вопрос об NP-полноте задачи 2-выполнимости. Известно, что этот вопрос решается отрицательно. Нетрудно заметить, что задача 2-выполнимости булевых формул решается с полиномиальной сложностью приведенным выше методом резолюций.

2.4. О СЛОЖНОСТНОЙ КЛАССИФИКАЦИИ СИСТЕМ БУЛЕВЫХ УРАВНЕНИЙ

Булевыми уравнениями называют такие уравнения, в которых левые и правые части являются булевыми функциями.

Сопоставим каждому классу F ненулевых булевых функций от переменных из множества $X = \{x_1, x_2, ...\}$ его расширение $< F >$, полученное путем его замыкания относительно всевозможных замен переменных в функциях на константы из $\Omega = \{0, 1\}$ и произвольные переменные из X, и два класса систем уравнений:

класс $[F]_{NC}$ всех систем уравнений вида

$$f_i(x_{i1}, ..., x_{ik_i}) = 1,\ i = 1, ..., m, \tag{2.14}$$

где $f_i \in F$, $x_{i1}, ..., x_{ik_i} \in X$, $m \in \mathbb{N}$, и класс $[F]_C = [< F >]_{NC}$.

Для каждой системы булевых уравнений возникают задачи по оценке сложности алгоритмов распознавания совместности, нахождения числа решений и самих решений. В данном параграфе будет рассмотрен вопрос о сложности задачи распознавания совместности систем уравнений вида (2.14). Эта задача обычно обозначается через $Sat([F]_{NC})$. Если булевы функции f_i системы (2.14) заданы формулами алгебры высказываний, то задача распознавания совместности системы (2.14) совпадает с задачей выполнимости формулы

$$f_1(x_{11}, ..., x_{1k_1}) \& ... \& f_m(x_{m1}, ..., x_{mk_m}).$$

Следовательно, в силу теоремы 2.6 задачи $Sat([F]_{NC})$ и $Sat([F]_C)$ принадлежат классу NP.

Заметим, что если в системе уравнений (2.14) используется нулевая функция, то система несовместна. Если же используется функция, тождественно равная 1, то соответствующее уравнение не влияет на решения системы и его можно удалить. В связи с этим всюду далее, без оговорок, будем считать, что F — класс булевых функций, не содержащий функций-констант.

В работе [72] выделены классы функций F, для которых проблема $Sat([F]_{NC})$ полиномиальна, и доказано, что в остальных случаях она NP-полна. В данном пособии мы приводим эти результаты в изложении С. П. Горшкова ([15]).

Определение 2.7. *Булева функция $f(x_1, ..., x_k)$ называется:*

1) 1-выполнимой, если $f(1, ...1) = 1$;

2) 0-выполнимой, если $f(0, ...0) = 1$;

3) мультиаффинной, если f представляется в виде конъюнкции аффинных функций;

4) биюнктивнной, если f представляется в виде конъюнкции дизъюнктов длины 2;

5) слабо положительной, если f представляется КНФ с дизъюнктами вида

$$x_{i1}^{r_i} \vee x_{i2} \vee ... \vee x_{ik_i}, \tag{2.15}$$

где $r_i \in \{0,1\}$;

б) слабо отрицательной, если f *представляется в виде КНФ вида*

$$\&_{i=1}^{t}(x_{i1}^{r_i} \vee \overline{x_{i2}} \vee ... \vee \overline{x_{ik_i}}),$$

где $r_i \in \{0,1\}$.

Заметим, что класс 1 совпадает с замкнутым классом T_1 всех функций, сохраняющих константу 1.

Множество всех функций классов 1–6 обозначим соответственно через

$$T_1,\ S_0,\ A,\ Bi,\ WP,\ WN.$$

Ниже функции каждого из классов A, Bi, WP, WN будут считаться заданными в форме, указанной определением 2.7, поскольку задача перехода от любого одного задания функции к любому другому является полиномиальной.

Введенные выше классы функций 1–6 можно охарактеризовать через свойства множеств истинности функций. Любая булева функция $f(x_1, ..., x_n)$ однозначно определяется своим множеством истинности из Ω^n:

$$M(f) = \{(a_1, ..., a_n) : f(a_1, ..., a_n) = 1\}.$$

Для функций классов 1, 2 такая характеризация очевидна, для остальных она дается следующей теоремой.

Теорема 2.8. *Для любой булевой функции* $f(x_1, ..., x_n) \neq 0$ *имеют место утверждения:*

а) $f \in A \iff \forall a, b, c \in \Omega^n : a, b, c \in M(f) \to a \oplus b \oplus c \in M(f)$;

б) $f \in Bi \iff \forall a, b, c \in \Omega^n : a, b, c \in M(f) \to (a \vee b) \oplus (a \vee c) \oplus (b \vee c) \in M(f)$;

в) $f \in WP \iff \forall a, b \in \Omega^n : a, b \in M(f) \to a \vee b \in M(f)$;

г) $f \in WN \iff \forall a, b, c \in \Omega^n : a, b \in M(f) \to a\&b \in M(f)$.

□ а) По определению 2.7 $f \in A$ тогда и только тогда, когда f представляется в виде

$$f = f_1\&...\&f_k, \tag{2.16}$$

где f_i — аффинные функции. Так как уравнение $f(x_1, ..., x_n) = 1$ эквивалентно системе уравнений

$$f_i(x_1, ..., x_n) = 1, i \in \{1, ..., k\}$$

над полем $GF(2)$, то получаем: $f \in A$ в том и только том случае, когда $M(f)$ есть множество решений подходящей системы линейных уравнений. Из курса алгебры известно, что последнее условие выполняется тогда и только тогда, когда $M(f) = \varnothing$ или $M(f)$ является линейным многообразием, т. е. смежным классом пространства Ω^n над полем $GF(2)$ по некоторому линейному подпространству, или, что то же самое, по некоторой подгруппе группы $(\Omega^n, \oplus)$. Так как $f \neq 0$, то $M(f) \neq \varnothing$. Следовательно, для доказательства утверждения *а)* остается заметить, что подмножество $V \subset \Omega^n$ замкнуто относительно сложения троек векторов тогда и только тогда, когда оно является смежным классом группы $(\Omega^n, \oplus)$ по подгруппе $V_0 = \{a \oplus b : a, b \in V\}$. Заметим, что используемые в доказательстве алгебраические факты являются упражнениями по алгебре (см. [14], с. 168, 302).

б) Пусть $f(x_1, ..., x_n) \in Bi$. Тогда f представляется в виде (2.1), где f_i — дизъюнкты длины 2. Так как операции $\vee$ и $\oplus$ над векторами осуществляются покоординатно, то требуемое утверждением *б)* свойство векторов из $M(f)$ достаточно проверить лишь для векторов длины 2 из $M(f_i)$, что легко осуществляется непосредственной проверкой.

Обратно, пусть функция $f(x_1, ..., x_n)$ обладает свойством:

$$\forall a, b, c \in M(f) : (a \vee b) \oplus (a \vee c) \oplus (b \vee c) \in M(f). \qquad (2.17)$$

Допустим, что f не биюнктивна. Тогда любая ее КНФ, и в частности сокращенная КНФ $A = A(x_1, ..., x_n)$, будет содержать дизъюнкты длины, большей 2. Зафиксируем некоторый из таких дизъюнктов D. Очевидно, что любая замена переменного x_i на x_j или на $\overline{x_j}$ не выводит функцию из класса Bi. Непосредственной проверкой можно убедиться, что такие замены сохраняют и свойство (2.17). Поэтому, не теряя общности, будем считать, что

$$D = x_1 \vee x_2 \vee, ..., \vee x_m \ , \ 3 \leq m \leq n.$$

Так как D — имплицента формулы A, или, что то же самое, функции f, то

$$D\&A \equiv A. \tag{2.18}$$

Докажем, что

$$A(x_1, ..., x_i, 0, ..., 0, x_{m+1}, ..., x_n) \not\equiv 0,\ i = 1, ..., m. \tag{2.19}$$

Допустим, например, что для $i = 1$ имеет место тождество

$$A(x_1, 0, ..., 0, x_{m+1}, ..., x_n) \equiv 0. \tag{2.20}$$

Покажем, что тогда при $D' = x_2 \vee, ..., \vee x_m$ выполняется равенство

$$D'\&A \equiv A. \tag{2.21}$$

Пусть $a = (a_1, ..., a_n)$ — любой набор значений переменных. Если среди $a_2, ..., a_m$ есть 1, то $D'(a) = D(a)$ и равенство (2.21) в этом случае следует из (2.19). Если же $a_2 = ... = a_m = 0$, то $D(a) = 0$ и равенство (2.21) в этом случае следует из (2.20). В итоге равенство (2.21) доказано. Однако оно означает, что D' есть имплицента функции f. Это противоречие с простотой имплиценты D и доказывает неравенство (2.19) при $i = 1$. Аналогично оно доказывается и при других значениях i.

Из неравенств (2.19) следует, что $M(f)$ содержит векторы вида

$$a = (1, 0, 0, ..., 0, a_{m+1}, ..., a_n),$$

$$b = (0, 1, 0, ..., 0, b_{m+1}, ..., b_n),$$

$$c = (0, 0, 1, ..., 0, c_{m+1}, ..., c_n).$$

Легко видеть, что при таких a, b, c указанный в (2.17) вектор α имеет вид $d = (0, 0, ..., 0, d_{m+1}, ..., d_n)$ и не содержится в $M(f)$. Следовательно, наше допущение неверно, и $f \in Bi$.

в) Пусть $f \in WP$. Тогда f имеет вид (2.16), где f_i — дизъюнкты вида (2.15). Непосредственной проверкой убеждаемся, что для любого дизъюнкта D вида (2.15) из $a, b \in M(D)$ следует $a \vee b \in M(D)$. Поэтому и из $a, b \in M(f)$ следует $a \vee b \in M(f)$.

Обратно, пусть последняя импликация выполнена для любых $a, b \in \Omega^n$, а $f \notin WP$. Тогда сокращенная КНФ $A = A(x_1, ..., x_n)$ функции f будет иметь дизъюнкт D, содержащий не менее двух переменных с отрицанием. Не теряя общности, будем считать, что

$$D = \overline{x}_1 \vee ... \vee x_t \vee x_{t+1} \vee ... \vee x_m,\ 2 \le t \le m \le n.$$

Далее по той же схеме, что и в пункте *б)*, докажем неравенства

$$A(x_1, 1, 1, ..., 1, 0, ..., 0, x_{m+1}, ..., x_n) \not\equiv 0,$$

$$A(1, x_2, 1, ..., 1, 0, ..., 0, x_{m+1}, ..., x_n) \not\equiv 0.$$

Следовательно, в $M(f)$ найдутся векторы вида

$$a = (0, 1, 1, ..., 1, 0, ..., 0, x_{m+1}, ..., x_n),$$

$$b = (1, 0, 1, ..., 1, 0, ..., 0, x_{m+1}, ..., x_n),$$

при которых $\alpha = a \vee b \notin M(f)$, поскольку $D(\alpha) = 0$. Полученное противоречие с условием и доказывает, что $f \in WP$.

г) В этом случае доказательство проводится по схеме, двойственной случаю *в)*. □

Теорема 2.9. *Если для конечного набора F ненулевых булевых функций выполняется хотя бы одно из включений*

$$F \subset T_1,\ F \subset S_0,\ F \subset A,\ F \subset Bi,\ F \subset WP,\ F \subset WN, \quad (2.22)$$

то задача $Sat([F]_{NC})$ полиномиальна.

□ Если $F \subset T_1$ или $F \subset S_0$, то утверждение очевидно, поскольку в этих случаях любые системы уравнений из $[F]_{NC}$ совместны и имеют соответственно решения $(1, ..., 1)$ или $(0, ..., 0)$.

В случае $F \subset A$ задача $Sat([F]_{NC})$ эквивалентна задаче распознавания совместности системы линейных уравнений, которая полиномиальна в силу полиномиальности алгоритма Гаусса.

В случае $F \subset Bi$ задача $Sat([F]_{NC})$ эквивалентна задаче 2-выполнимости КНФ, которая, как указано в замечании к теореме 2.7, полиномиальна.

Пусть $F \subset WP$. В этом случае любая система из класса $[F]_{NC}$ полиномиально сводится к равносильной ей системе уравнений вида

$$x_{i1}^{r_i} \vee x_{i2} \vee ... \vee x_{ik_i} = 1,\ i = \{1, ..., m\}\ ,\ r_i \in \Omega.$$

Укажем алгоритм распознавания совместности такой системы.

Если каждое уравнение системы содержит хотя бы одно неизвестное без отрицания, то система имеет решение $(1, ..., 1)$ и, значит, совместна. В противном случае, просматривая подряд все уравнения системы, находим первое уравнение вида $\overline{x}_{sj} = 1$, полагаем $x_{sj} = 0$ и удаляем из системы все уравнения, содержащие $\overline{x}_{sj}$. Если при этом будут удалены все уравнения, то система совместна и ее решением будет любой набор значений переменных, в котором $x_{sj} = 0$. Если удалятся не все уравнения и среди оставшихся уравнений найдется уравнение $x_{sj} = 1$, то система не совместна. Если же в оставшихся уравнениях уравнения $x_{sj} = 1$ не найдется, то удаляем из них все литералы вида x_{sj}, если такие существуют. В итоге получим систему уравнений от меньшего числа неизвестных, которая, очевидно, совместна в том и только том случае, когда совместна исходная система. Далее ту же процедуру применяем к полученной системе уравнений. Нетрудно заметить, что указанный алгоритм полиномиален. Аналогично (с заменой лишь x_{sj} на $\overline{x}_{sj}$) устанавливается этот факт и при $F \subset WN$. □

Для теоремы 2.8 справедлива и обратная теорема. Для ее доказательства нам понадобятся некоторые новые понятия и вспомогательные утверждения.

Определение 2.8. *Сужением функции $f(x_1, ..., x_n)$ по ее переменным $x_{i_1}, ..., x_{i_k}$, $0 \le k < n$, называется функция $\tilde{f}$ от остальных переменных $x_{j_1}, ..., x_{j_{n-k}}$, для которой $M(\tilde{f})$ получается удалением во всех векторах из $M(f)$ координат с номерами $i_1, ..., i_k$.*

Введем для функции $\tilde{f}$ обозначение:

$$\tilde{f}(x_{j_1}, ..., x_{j_{n-k}}) = [x_{i_1}...x_{i_k}]f(x_1, ..., x_n). \qquad (2.23)$$

Пример. Пусть $f(x_1, x_2, x_3) = (\overline{x}_1 \vee \overline{x}_2)\&(x_2 \oplus x_3)$.

Пользуясь табличным заданием функции f, находим

$[x_1]f(x_1, x_2, x_3) = x_2 \oplus x_3$,

$[x_2]f(x_1, x_2, x_3) = \overline{x}_1 \vee x_3$,

$[x_3]f(x_1, x_2, x_3) = \overline{x}_1 \vee \overline{x}_2$.

Заметим, что эти функции не являются подфункциями для f.

Отметим некоторые очевидные свойства сужений, доказать которые предлагается в качестве упражнений: сама функция f является своим сужением (по пустому множеству переменных); сужение (2.23) не изменится при замене переменных $x_{i_1}...x_{i_k}$ на любые переменные, отличные от $x_{j_1}, ..., x_{j_{n-k}}$, произведение сужений двух функций является сужением их произведения по подходящей системе переменных (для доказательства переобозначьте сначала переменные в исходных функциях так, чтобы наборы переменных, по которым сужаются функции не пересекались между собой и с системой остальных переменных); любое сужение функции f является выполнимой функцией в том и только том случае, когда выполнима f. Сужение функции $(x_1 \vee A)(\overline{x}_1 \vee B)$ по переменному x_1 совпадает с функцией, представленной резолюцией $A \vee B$ формулы $(x_1 \vee A)(\overline{x}_1 \vee B)$ при любых не равных нулю и не зависящих от x_1 формулах A, B.

Теперь для класса функций F от n переменных построим еще два расширения: $\{F\}$ и (F):

$\{F\} = \{\&_{i=1}^{cn} f_i(x_{i1}, ..., x_{ik_i}) : f_i \in < F >\}$, где c пока произвольная, достаточно большая константа;

(F) — множество всевозможных сужений всех функций из F.

Имеют место очевидные включения:

$$F \subset < F > \subset \{F\} \subset (F).$$

Лемма 2.10. *Если $h(x_1, ..., x_k) \in (F)$ и задача $Sat([h]_C)$ — NP-полна, то и задача $Sat([F]_C)$ — NP-полна.*

□ Достаточно показать, что задача $Sat([h]_C)$ полиномиально сводится к задаче $Sat([F]_C)$. По условию h, с точностью до перенумерации переменных, представляется в виде

$$h(x_1, ..., x_k) = [x_{k+1}...x_p]g(x_1, ..., x_p),\ g \in \{F\}.$$

Пусть

$$h(y_{i1}, ..., y_{ik}) = 1,\ i = 1, ..., m \tag{2.24}$$

— любая система уравнений из класса $([h]_C)$, где $y_{ij} \in \in \{0, 1, x_1, ..., x_n\}$. Так как среди переменных $x_1, ..., x_n$ могут встретиться и $x_{k+1}, ..., x_p$, то в указанном представлении необходимо заменить переменные $x_{k+1}, ..., x_p$ на новые переменные, не встречающиеся в системе (2.24). В связи с этим будем считать, что

$$h(x_1, ..., x_k) = \\ = [x_{mk+1}...x_{mk+p-k}]g(x_1, ..., x_k, x_{mk+1}, ..., x_{mk+p-k}).$$

Тогда система уравнений (2.24) запишется в виде

$$[x_{mk+1}, ..., x_{mk+p-k}]g(y_{i1}, ..., y_{ik}, x_{mk+1}, ..., x_{mk+p-k}) = 1,$$

$$i = 1, ..., m.$$

При этом в силу свойств сужений система эта совместна тогда и только тогда, когда совместна система

$$g(y_{i1}, ..., y_{ik}, x_{mk+1}, ..., x_{mk+p-k}), i = 1, ..., m.$$

Таким образом, вопрос о совместности любой системы из $([h]_C)$ сводится к его решению для системы из $([F]_C)$. Очевидно, что это сведение полиномиально, поскольку требуется лишь замена переменных. □

Лемма 2.11. *Если $F \not\subset WP, WN$, то $\{F\}$ содержит функцию $x_1 \oplus x_2$.*

□ Пусть $f(x_1, ..., x_k) \notin WP$. Тогда по теореме 2.7 найдутся наборы $a, b \in \Omega^k$, такие что

$$f(a) = f(b) = 1, f(a \vee b) = 0. \tag{2.25}$$

Не теряя общности, можно считать, что

$$a = (0, ..., 0, 0, ..., 0, 1, ..., 1, 1, ..., 1),$$

$$b = (0, ..., 0, 1, ..., 1, 0, ..., 0, 1, ..., 1),$$

$$a \vee b = (0, ..., 0, 1, ..., 1, 1, ..., 1, 1, ..., 1).$$

Здесь координаты векторов разбиты на четыре блока, длины которых обозначим соответственно через p, q, r, s. Из условия (2.25) следует, что $q > 0, r > 0$ (в противном случае мы бы имели $a \vee b = a$ или b). Отсюда же, заменяя в функции f переменные в указанных блоках соответственно на $0, x_1, x_2, 1$, получим функцию

$$g(x_1, x_2) = f(0, ..., 0, x_1, ..., x_1, x_2, ..., x_2, 1, ..., 1),$$

удовлетворяющую условиям:

$$g(0, 1) = 1, g(1, 0) = 1, g(1, 1) = 0.$$

Аналогично из функции не из класса WN получим функцию $h(x_1, x_2)$ такую, что

$$h(0, 1) = 1, h(1, 0) = 1, h(0, 0) = 0.$$

Теперь очевидно, что $g\&h = x_1 \oplus x_2 \in \{F\}$. □

Лемма 2.12. *Если* $f(x_1, ..., x_k) \in < F >$, *то*

$$f(x_1 \oplus a_1, ..., x_k \oplus a_k) \in (F)$$

для любых $a_1, ..., a_k \in \Omega$.

□ Утверждение очевидно при $(a_1, ..., a_k) = (0, ..., 0)$. Поэтому будем считать, что $(a_1, ..., x_k) \neq (0, ..., 0)$, и для упрощения записей положим

$$a_1 = ... = a_t = 0,\ a_{t+1} = ... = a_k = 1,\ 1 \leq t \leq k.$$

Из условия и леммы 2.11 следует, что (F) содержит функцию

$$\tilde{f}(x_1, ..., x_k) = [z_{t+1} ... z_k] f(x_1, ..., x_t, z_{t+1}, ..., z_k) \& \\ \&(x_{t+1} \oplus z_{t+1}) \& ... \& (x_k \oplus z_k),$$

где $z_{t+1}, ..., z_k$ — переменные из X, отличные от $x_1, ..., x_k$.

Замечаем, что для любых $b_1, ..., b_k \in \Omega$:

$$\tilde{f}(b_1, ..., b_k) = 1 \Longleftrightarrow f(b_1, ..., b_t, \overline{b}_{t+1}, ..., \overline{b}_k) = 1.$$

Отсюда видно, что

$$\tilde{f}(x_1, ..., x_k) = f(x_1, ..., x_t, \overline{x_{t+1}}, ..., \overline{x_k}) = \\ = f(x_1 \oplus a_1, ..., x_k \oplus a_k).$$

□

Лемма 2.13. *Если $F \not\subset WP, WN, A$, то F содержит функции $x_1^{u_1} \vee x_2^{u_2}$ при любых $u_1, u_2 \in \Omega$.*

□ Пусть $f(x_1, ..., x_k) \notin A$. Тогда по теореме 2.8 найдутся $a, b, c \in \Omega^k$, такие что

$$f(a) = f(b) = f(c) = 1, f(a \oplus b \oplus c) = 0.$$

Обозначим $x = (x_1, ..., x_k), \theta = (0, ..., 0)$ и рассмотрим функцию $g(x) = f(x \oplus a)$. Заметим, что

$$\begin{array}{l} g(\theta) = f(a) = 1, \;\; g(a \oplus b) = f(b) = 1, \\ g(a \oplus c) = f(c) = 1, \;\; g(b \oplus c) = f(a \oplus b \oplus c = 0. \end{array} \tag{2.26}$$

Не теряя общности, можно считать, что

$$a \oplus b = (0, ..., 0, 0, ..., 0, 1, ..., 1, 1, ..., 1),$$

$$a \oplus c = (0, ..., 0, 1, ..., 1, 0, ..., 0, 1, ..., 1),$$

и тогда

$$b \oplus c = (0, ..., 0, 1, ..., 1, 1, ..., 1, 0, ..., 0).$$

Здесь, как и в лемме 2.11, координаты векторов разбиты на четыре блока соответственно длин p, q, r, s. При этом возможны различные варианты.

Пусть $q > 0$, $r > 0$, $s > 0$. Тогда, заменяя в функции g переменные в указанных блоках соответственно на $0, x_1, x_2, x_3$, получим функцию $u(x_1, x_2, x_3)$, такую что

$$u(0, 0, 0) = u(0, 1, 1) = u(1, 0, 1) = 1 \;,\; u(1, 1, 0) = 0.$$

Далее непосредственной проверкой убеждаемся, что:

$[x_3]u(x_1, x_2, x_3) = \overline{x}_1 \vee \overline{x}_2$; если $u(1,1,1) = 0$,

$[x_1]u(x_1, x_2, x_3) = x_1 \vee \overline{x}_2$, если $u(1,1,1) = 1$ и $u(0,1,0) = 0$;

$u(x_1, 1, x_3) = \overline{x}_1 \vee x_2$, если $u(1,1,1) = 1, u(0,1,0) = 1$.

Таким образом, $\{F\}$ содержит функцию вида $x_1^{u_1} \vee x_2^{u_2}$ при некоторых $u_1, u_2 \in \Omega$.

К такому же результату непосредственно из условий (2.26) приходим и в том случае, если одно из чисел q, r, s равно нулю. Случаи, когда два из чисел q, r, s равны нулю, приводят к противоречию с условиями (2.26).

Теперь, используя лемму 2.12, мы из любой одной функции вида $x_1^{u_1} \vee x_2^{u_2}$ получим все функции такого вида. □

Лемма 2.14. *Если* $F \not\subset WP, WN, A, Bi$, *то* (F) *содержит функцию*

$$w(x_1, x_2, x_3) = x_1 \& x_2 \& x_3 \oplus x_1 \oplus x_2 \oplus x_3.$$

□ Пусть $f(x_1, ..., x_k) \notin Bi$. Тогда по теореме 2.7 найдутся $a, b, c \in \Omega^k$, такие что

$$f(a) = f(b) = f(c) = 1\ ,\ f(a\&b \oplus a\&c \oplus b\&c) = 0.$$

По функции f, как и при доказательстве леммы 2.13, определим функцию $g(x)$ и найдем ее значения при $x = a \oplus b$, $x = a \oplus c$, $x = \theta$. Кроме того, учитывая равенство

$$(a \oplus b)\&(a \oplus c) \equiv a \oplus a\&b \oplus a\&c \oplus b\&c,$$

найдем $g((a \oplus b)\&(a \oplus c))$. Далее снова, как и в лемме 2.13, определим функцию $u(x_1, x_2, x_3)$. Для нее будем иметь

$$u(0,0,0) = u(0,1,1) = u(1,0,1) = 1\ ,\ u(0,0,1) = 0.$$

Теперь непосредственной проверкой легко убедиться, что функция

$$u(x_1, x_2, \overline{x}_3)\&(\overline{x}_1 \vee \overline{x}_2)\&(\overline{x}_1 \vee \overline{x}_3)\&(\overline{x}_2 \vee \overline{x}_3)$$

совпадает с искомой функций $w(x_1, x_2, x_3)$. Из утверждений лемм 2.12, 2.13 следует, что эта функция содержится в (F). □

Лемма 2.15. *Если* $F \not\subset A, Bi, WP, WN$, *то задача* $Sat([F]_C)$ *является NP-полной.*

□ Учитывая теорему 2.8 и лемму 2.10, достаточно доказать, что задача три-выполнимости полиномиально сводится к задаче $Sat([w]_C)$. Для этого выразим любую функцию вида

$$x_1^{u_1} \vee x_2^{u_2} \vee x_3^{u_3} \tag{2.27}$$

через функцию w. Непосредственной проверкой убеждаемся, что

$$(\overline{x}_1 \vee \overline{x}_2 \vee \overline{x}_3) \equiv$$

$$\equiv [y_1 ... y_6] w(x_1, y_1, y_2) \& w(x_2, y_3, y_4) \& w(x_3, y_5, y_6) \& w(y_1, y_3, y_5).$$

Из леммы 2.11 следует, что для получения других функций вида (2.27) достаточно иметь функцию $x_1 \oplus x_2$. Однако эта функция также выражается через w:

$$x_1 \oplus x_2 \equiv w(x_1, x_2, 0).$$

Таким образом, задача три-выполнимости сведена к задаче $Sat([w]_C)$. Нетрудно видеть, что такое сведение полиномиально. □

Лемма 2.16. *Если для конечной системы* F *ненулевых булевых функций не выполнено ни одно из включений (2.22), то задача* $Sat([F]_C)$ *полиномиально сводится к задаче* $Sat([F]_{NC})$.

□ Для требуемого сведения задач достаточно научиться заменять константы 0 и 1 в системах уравнений из $[< F >]$ на переменные из X, переводя таким образом системы уравнений из класса $[F]_C$ в класс $[F]_{NC}$. Идея такой замены проста. Константу a в уравнениях нужно заменить какой-либо переменной x, не содержащейся в исходной системе, а систему уравнений дополнить новой системой уравнений из класса $[F]_{NC}$ с неизвестными, также не входящими в исходную систему, и имеющую решения только при $x = a$. Рассмотрим случаи, когда эта идея реализуется.

Так как $F \not\subset T_1$, то в F есть функция f, такая что $f(1,...,1) = 0$. Кроме того, по условию $f \not\equiv 0$, и значит, существует набор a, при котором $f(a) = 1$. Заменяя в a нули на x_1, а единицы на x_2, получим из f функцию $\tilde{f}(x_1, x_2)$, которая совпадает с $\overline{x}_1$ при $a = (0,...,0)$ и обладает свойствами: $\tilde{f}(0,1) = 1$, $\tilde{f}(1,1) = 0$ в противном случае. Аналогично из функции $g \in F \setminus S_0$ получим функцию $\tilde{g}(x_1, x_2)$, которая совпадает с x_1 или обладает свойствами: $\tilde{g}(0,1) = 1$, $\tilde{g}(0,0) = 0$.

Рассмотрим систему уравнений

$$\tilde{f}(x_1, x_2) = 1 \ , \ \tilde{g}(x_1, x_2) = 1. \tag{2.28}$$

Если $\tilde{f}(x_1, x_2) = \overline{x}_1$ или $\tilde{g}(x_1, x_2) = x_1$, то система 2.28 будет иметь единственное решение $x_1 = 0, x_2 = 1$ и константы получены. К тому же приходим и в случае, когда $\tilde{f} \not\equiv \overline{x_1}, \tilde{g} \not\equiv x_1$, и при этом $\tilde{f}(1,0) = 0$ или $\tilde{g}(1,0) = 0$.

Остается рассмотреть случай, когда

$$\tilde{f}(0,1) = 1, \tilde{f}(1,1) = 0, \tilde{f}(1,0) = 1,$$

$$\tilde{g}(0,1) = 1, \tilde{g}(0,0) = 0, \tilde{g}(1,0) = 1.$$

Заметим, что в этом случае $\tilde{f}\&\tilde{g} = x_1 \oplus x_2$.

Далее различаются два подслучая.

1. Класс F содержит несамодвойственную функцию h. Тогда найдется набор a, такой что $h(a) = h(\overline{a}) = 1$. Заменяя в a единицы на x_1, а нули на x_2, мы получим функцию $\tilde{h}$, такую что добавление к системе (2.28) уравнения $h(x_1, x_2) = 1$ приводит к системе, имеющей единственное решение $(0,1)$ или $(1,0)$. Значит, константы получены.

2. Все функции из F самодвойственны. В этом случае указанная выше идея не проходит и предлагается другое решение.

В исходной системе уравнений

$$f_i(x_{i1}, ..., x_{ik_i}) = 1 \ , \ i = 1, ..., m, \tag{2.29}$$

где $x_{ij} \in X \cup \Omega$, заменим нули на x_1, а единицы на x_2, предполагая, что переменные x_1, x_2 не входят в систему (2.29) (этого

можно добиться перенумерацией переменных). Объединив полученную систему с системой (2.28), получим систему, которую обозначим через (S). Очевидно, что совместность системы (2.29) влечет совместность системы (S). Обратно, пусть совместна система (S) и a — ее решение. Так как (S) содержит (2.28), то значения переменных x_1, x_2 в a не могут совпадать. Если они равны соответственно 0, 1, то a удовлетворяет и системе (2.28). В противном случае решением системы (2.29) будет вектор $\overline{a}$, противоположный a, поскольку все функции системы F самодвойственны. Таким образом, во всех случаях распознавание совместности системы уравнений из $[F]_C$ сводится к распознаванию совместности системы уравнений из $[F]_{NC}$. Легко видеть, что такое сведение полиномиально. □

Из доказанных лемм легко следует основной результат данного параграфа, т. е. теорема, обратная к теореме 2.8.

Теорема 2.17. *Если для конечной системы F ненулевых булевых функций не выполнено ни одно из включений (2.22), то задача $Sat([F]_{NC})$ NP-полна.*

□ Обозначим в виде $A \prec B$ тот факт, что задача A полиномиально сводится к задаче B. Тогда результаты лемм 2.15, 2.10, 2.16 при выполнении их условий запишутся в виде

$$3\text{-вып} \prec Sat([w]_C) \prec Sat([F]_{NC}). \tag{2.30}$$

При доказательстве лемм использовались некоторые функции класса $\{F\}$, который зависел от константы c. Теперь нетрудно оценить сверху ее значение, при котором будут корректными все утверждения, используемые в доказательстве лемм. При сводимости 3-вып $\prec Sat([w]_C)$ использовались все дизъюнкты длины 3 и функция w. Нетрудно посчитать, что для w понадобится перемножать не более $8n$ функций из $< F >$, а для нахождения дизъюнкты длины 3 понадобится перемножать не более $4 \cdot 8n$ функций. Следовательно, в качестве константы c можно взять число 32. При этом будут справедливы все соотношения из (2.30), и в силу теоремы 2.7 задача $Sat([F]_{NC})$ — NP-полна. □.

Замечание. Следует иметь в виду, что функция Шеннона для временной (емкостной) сложности массовой задачи определяется по максимально сложной ее индивидуальной задаче из всех индивидуальных задач с фиксированной длиной входа. Поэтому NP-полнота той или иной массовой задачи даже при гипотезе $P \notin NP$ не означает, что в ней нет подклассов задач полиномиальной сложности. Так, например, задача тривыполнимости формул алгебры высказываний NP-полна, однако ее подзадача 2-выполнимости — полиномиальна. Следовательно, доказательство NP-полноты задачи $Sat([F]_{NC})$ (как и любой другой массовой задачи) не исключает поиска более простых решений в ее частных случаях, для различных частных видов систем уравнений.

2.5. АСИМПТОТИЧЕСКИЕ ОЦЕНКИ СЛОЖНОСТИ АЛГОРИТМОВ

Машина Тьюринга была изобретена в свое время с целью построения универсального инструмента, позволяющего моделировать любые вычислительные процессы путем раздробления их на самые простые операции. Ее изобретение сыграло огромную роль в появлении и развитии теории алгоритмов. МТ положена в основу одного из общепринятых в настоящее время строго математических определений понятия алгоритма. Наличие таких определений позволило доказать отсутствие алгоритмов для решения многих массовых задач. МТ широко использовались также и для сложностной классификации задач.

Вместе с тем процесс функционирования МТ не вполне соответствует процедурам, осуществляемым в ходе работы современных вычислительных машин. Элементная база современных компьютеров, позволяет однотактно осуществлять более крупные операции, чем в МТ. Кроме того, в них и само управление процессом работы является более сложным. Вспомните, как МТ распознавала симметрию слова. Для сравнения первой буквы с последней она вынуждена была пробегать все промежуточные ячейки ленты. В современных компьютерах допускаются и параллельные вычисления, и прямой

доступ к памяти. В связи с этим на практике сложность алгоритма чаще всего оценивается не временем работы соответствующей МТ, а минимальным числом некоторых элементарных операций, необходимых для переработки входных данных в выходные. При этом в разных алгоритмах за элементарную операцию могут приниматься различные конкретные процедуры: сравнения элементов, перестановка элементов, алгебраические операции над элементами и т. д. При работе с числами зачастую, учитывая двоичный алфавит компьютеров, стремятся записывать данные в двоичном алфавите Ω, а за элементарную операцию брать вычисление значения любой булевой функции от двух существенных переменных.

Как и в машинах Тьюринга, здесь следует различать сложность задачи и сложность алгоритма, решающего эту задачу. Грубо говоря, сложность задачи — это минимум сложностей всех алгоритмов, решающих эту задачу. Для точного определения этих понятий необходимо предварительно условиться, что понимать под элементарной операцией и как измерять в числах сложность (размер) входных данных. При этих условиях фиксированный алгоритм F, будучи детерминированной процедурой, затрачивает вполне определенное число операций при переработке каждого допустимого входа, и функция сложности $T_F(n)$ алгоритма F определяется как максимум указанного числа операций по всем входам размера n, а функция сложности $T_F(n)$ задачи — как минимум функций $T_F(n)$ по всем алгоритмам F, решающим эту задачу.

Так как для решения каждой разрешимой задачи может существовать много различных алгоритмов, то найти точное значение сложности задачи, как правило, не удается. Поэтому на практике обычно оценивают не сложность задачи, а сложности конкретных алгоритмов ее решения. Да и в этом случае точное значение функции $T_A(n)$ при каждом значении n найти также непросто. Поэтому сложность алгоритмов чаще всего оценивают лишь с точностью до порядка, т. е. с точностью до мультипликативной константы, не зависящей от размера входа. Кроме того, для большей выразительности скорости роста функции сложности алгоритма при росте длины входа, указывают обычно асимптотическую оценку сложности при $n \to \infty$.

В этом случае пишут

$$T_F(n) = O(f(n)),$$

понимая под этим существование константы $c \in \mathbb{R}$ и значения n_0, таких что

$$\forall n \geq n_0 : T_F(n) \leq c \cdot f(n).$$

Заметим, что при использовании асимптотической оценки $O(f(n))$ сложности алгоритма F на практике важную роль играет не только функция f, но и константа c. Поэтому, находя асимптотическую оценку алгоритма, обычно стремятся оценивать и по возможности уменьшать входящую в нее константу c.

Один из распространенных приемов построения сравнительно быстрых алгоритмов основан на сведении решения задачи к решению ее подзадач с входами меньших размеров. Получающиеся таким образом алгоритмы называются рекуррентными, или рекурсивными, алгоритмами. Они также носят еще название «разделяй и властвуй». Для оценки порядков сложности таких алгоритмов обычно используется теорема о рекуррентном неравенстве.

Теорема 2.18. *Если арифметическая монотонно неубывающая функция f при всех $n \in \mathbb{N}$ и некоторых константах $a > 0, c > 0, b$ из $\mathbb{R}$ и $k > 1$ из $\mathbb{N}$ удовлетворяет условию*

$$f(n) \leq af(n/2) + bn^c, \tag{2.31}$$

то:

1) $f(n) = O(n^c)$, *если* $c > \log a$,
2) $f(n) = O(n^{\log a})$, *если* $c < \log a$,
3) $f(n) = O(n^c \log n)$, *если* $c = \log a$,

где все логарифмы берутся при основании k.

□ Рассмотрим сначала случай, когда $n = 2^t$. Индукцией по k легко доказывается, что

$$f(n) \leq bn^c(1 + (\frac{a}{2^c}) + (\frac{a}{2^c})^2 + ... + (\frac{a}{2^c})^{t-1}) + a^t f(1).$$

Отсюда при $d = \max(b, f(1))$ имеем

$$f(n) \le dn^c \sum_{r=0}^{t} (\frac{a}{2^c})^r.$$

Подставляя в найденное соотношение значения c из пунктов 1–3, мы и получим требуемые результаты. При этом в первом случае следует учесть, что $t \to \infty$, а во втором — равенство $a^{\log n} = n^{\log a}$.

В том случае, когда n не является степенью двойки, найдем такое t, что $2^t < n \le 2^{t+1}$, и воспользуемся монотонностью функции f и соотношениями, доказанными для $n = 2^{t+1}$. □

Приведем примеры на использование этой теоремы.

Пример 1. Умножение целых чисел (методом Карацубы–Офмана). Входами этой задачи являются пары n-разрядных двоичных чисел, т. е. чисел, записанных в двоичной системе счисления, а выходами — произведения входных чисел. Размером задачи считается число n, а элементарной операцией — сложение и умножение чисел из Ω.

Обозначим через $S(n), M(n)$ сложности соответственно задач сложения и умножения n-разрядных двоичных чисел. Заметим сначала, что школьный метод сложения двух n-разрядных чисел столбиком реализуется со сложностью, не большей $2n$ (сложение значений разрядов и соответствующих переносов), и потому $S(n) \le 2n$. С другой стороны, операция сложения является функцией, существенно зависящей от $2n$ переменных (разрядов исходных слагаемых). Отсюда видно, что для вычисления суммы в общем случае потребуется не менее n элементарных двоичных операций. В связи с этим можно считать, что $S(n) = O(n)$.

Умножение n-разрядных двоичных чисел столбиком требует, очевидно, порядка n^2 операций сложения и умножения чисел из Ω, и потому $M(n) \le O(n^2)$. Однако для умножения чисел существуют более экономные алгоритмы. Примером такого алгоритма может служить известный рекуррентный алгоритм Карацубы–Офмана.

Пусть требуется перемножить два n-разрядных двоичных числа A, B.

Не теряя общности, можно считать, что $n = 2^t$, так как этого можно добиться приписыванием к n разрядам слева нулей. Представим числа A, B в виде

$$A = 2^{t-1}A_1 + A_2,\ B = 2^{t-1}B_1 + B_2,$$

где A_i, B_i — 2^{t-1}-разрядные числа. Непосредственной проверкой убеждаемся в справедливости равенства

$$AB = ((2^{2t-2} + 2^{t-1})A_1B_1 + \\ + 2^{t-1}(A_1 - A_2)(B_2 - B_1) + (2^{t-1} + 1)B_1)A_2B_2.$$

Из него видно, что для умножения чисел A, B достаточно произвести три умножения $n/2$-разрядных чисел, несколько сложений не более чем (2^{t+2})-разрядных чисел и умножений на степени двойки, которые осуществляются путем приписывания нулей справа. Так как операции сложения и сдвига имеют сложность $O(n)$, то получаем соотношение

$$M(n) \leq 3M(\frac{n}{2}) + bn, b = const > 0,$$

из которого следует, что условие (2.31) теоремы 2.18 выполняется при $k = 2, a = 3, c = 1$. Согласно этой теореме имеем

$$M(n) \leq O(n^{\log_2 3}).$$

Так как $\log_2 3 = 1,585... \leq 2$, то полученная оценка при достаточно больших значениях n существенно лучше указанной ранее оценки $O(n^2)$. Заметим, что это улучшение достигнуто не только засчет рекуррентности алгоритма, но и засчет специфического представления произведения чисел A, B.

Из описания приведенного алгоритма видно, что он может быть использован при умножении элементов в любом кольце с единицей.

Отметим еще, что алгоритм Карацубы–Офмана не является оптимальным. Наилучшим по асимптотической оценке

в настоящее время является алгоритм Шенхаге–Штрассена, имеющий оценку $M(n) \le O(n \log_n \cdot \log\log n)$.

Асимптотической оценкой сложности умножения чисел можно воспользоваться для нахождения сложности операции возведения в степень.

Пусть a — n-разрядное и k — любое фиксированное двоичные числа. Оценим функцию сложности $Q(n)$ вычисления степени a^k при любых натуральных n. Самый распространенный ныне алгоритм нахождения числа a^k сводится к операциям возведения в квадрат и умножения на a следующим образом.

Число k можно представить в виде

$$k = 2^{i_1} + 2^{i_2} + ... + 2^{i_t},\ i_1 < i_2 < ... < i_t \le \log_2 k, t \le \log_2 k.$$

Отсюда получаем следующие представления для k и a^k:

$$k = 2^{i_1}(1 + 2^{i_2 - i_1}(1 + ... + 2^{i_{k-1} - i_{t-2}}(1 + 2^{i_k - i_{k-1}})...)),$$

$$a^k = (...((a^{2^{i_t - i_{t-1}}}a)^{2^{i_{t-1} - i_{t-2}}}a)^{2^{i_{t-2} - i_{t-3}}}...)^{2^{i_2 - i_1}}a)^{2^{i_1}}.$$

Из последнего представления видно, что вычисление числа a^k сводится к i_k возведениям в квадрат и $k-1$ умножениям на a. Следует иметь в виду, что указанные операции будут производиться уже с числами разрядности, большей чем n. Однако их разрядность будет превосходить n не более, чем в q раз, где q — константа, определяемая числом k и не зависящая от n. Из полученной выше оценки для сложности операции умножения видно, что это может отразиться лишь на изменении константы, используемой при определении функции $O(M(n))$. Следовательно, $Q(n) = O(M(n)\log_2 k)$.

Пример 2. Умножение матриц (методом Штрассена). Входом данной задачи является пара матриц A, B размером $n \times n$ над кольцом R, выходом — матрица $C = AB$. Размером такого входа будет считаться число n, а элементарными операциями — операции сложения, вычитания и умножения элементов из R. Функцию сложности задачи обозначим через $T(n)$.

Будем считать, что $n = 2^t$. Этого можно добиться добавлением к матрицам нулевых строк и столбцов. Алгоритм, основанный на определении операции умножения матриц,

потребует n^3 умножений и $n^2(n-1)$ сложений элементов из R. Значит, $T(n) \le O(n^3)$.

В алгоритме Штрассена каждая из матриц A, B, а также их произведение C разбиваются на блоки размером $2^{t-1} \times 2^{t-1}$:

$$A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}, \quad B = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}, \quad C = \begin{pmatrix} C_1 & C_2 \\ C_3 & C_4 \end{pmatrix}$$

и далее при нахождении C производятся операции над блоками матриц A, B. Обычное блочное умножение матриц потребует 8 умножений и 4 сложения матриц размером $2^{t-1} \times 2^{t-1}$. Однако Штрассеном было замечено, что можно обойтись семью умножениями за счет увеличения числа более просто реализуемых операций сложения. Действительно, непосредственной проверкой легко убедиться, что имеют место равенства:

$$C_1 = P_1 + P_4 - P_5 + P_7,\ C_2 = P_3 + P_5,$$

$$C_3 = P_2 + P_4,\ C_4 = P_1 + P_3 - P_2 + P_6,$$

где

$$P_1 = (A_1 + A_4)(B_1 + B_4),$$

$$P_2 = (A_3 + A_4)B_1,$$

$$P_3 = A_1(B_2 - B_4),$$

$$P_4 = A_4(-B_1 + B_3),$$

$$P_5 = (A_1 + A_2)B_4,$$

$$P_6 = (-A_1 + A_3)(B_1 + B_2),$$

$$P_7 = (A_2 - A_4)(B_3 + B_4).$$

Отсюда видно, что для получения матрицы C достаточно произвести 7 умножений и 18 сложений/вычитаний матриц порядка 2^{t-1}. Заметим, что для нахождения суммы или разности двух матриц порядка 2^{t-1} достаточно произвести $(2^{t-1})^2$ соответствующих операций над элементами из R. Следовательно, имеем соотношение

$$T(n) \le 7 \cdot T(n/2) + 18n^2,$$

из которого следует, что условие теоремы 2.18 выполняется при $k = 2$, $a = 7$, $b = 18$, $c = 2$. Согласно этой теореме имеем

$$T(n) \leq O(7^{\log_2 n}) = O(n^{\log_2 7}).$$

Заметим, что в настоящее время алгоритм Штрассена асимптотически не оптимален, известен алгоритм умножения матриц с оценкой $O(n^{2,376})$, тогда как $\log_2 7 = 2,807...$.

Обратим еще внимание на то, что в условии теоремы 2.18 требуется, чтобы соотношение (2.31) выполнялось для всех n. Поэтому следует иметь в виду, что в алгоритме Штрассена на блоки нужно разбивать не только исходные матрицы A, B, но затем и получившиеся при этом блоки и т. д. Это потребует дополнительный расход и времени, и памяти. В связи с этим алгоритм Штрассена имеет в основном теоретическое значение и вряд ли где применяется на практике. Кстати, аналогичное замечание можно сделать и по алгоритму Карацубы–Офмана для умножения чисел.

2.6. ДИСКРЕТНЫЕ ПРЕОБРАЗОВАНИЯ ФУРЬЕ

В данном параграфе мы познакомимся с дискретными и быстрыми преобразованиями Фурье, которые широко используются в различных областях математики и в ее приложениях. В частности, при цифровой обработке сигналов, при вычислениях различных сверток, произведений многочленов и т. д. Так, например, упомянутый выше алгоритм Шенхаге–Штрассена для умножения чисел получен с использованием ДПФ.

Зафиксируем кольцо R, с единицей e, обладающее примитивным (первообразным) корнем n-й степени из единицы, т. е. элементом w, удовлетворяющим условиям:

$w \neq e$, $w^n = e$,

$\sum_{j=0}^{n-1} w^{i \cdot j} = 0$ при всех $i = 1, ..., n-1$.

Заметим, что если R — поле, то примитивный корень степени n из единицы — суть элемент порядка n мультипликативной группы поля.

Определение 2.9. Дискретным преобразованием Фурье (ДПФ) над кольцом R называется переход от заданий многочленов над R наборами коэффициентов к заданиям наборами значений соответствующих многочленов в некоторых точках из R. При этом обратный переход называется обратным ДПФ.

Таким образом, применение ДПФ и обратного ДПФ к многочлену f сводится соответственно к вычислению значений многочлена в выбранных точках и в интерполяции многочлена по его значениям в этих точках. Естественно, что сложность ДПФ и обратного ДПФ зависит от выбора точек интерполяции. Оказалось, что в качестве таких точек удобно брать степени примитивного элемента w кольца R.

Если условиться записывать многочлен $a(x) = a_0 + a_1x^1 + \\ + ... + a_{n-1}x^{n-1}$ в виде вектора

$$a = (a_0, a_1, ..., a_{n-1}), \tag{2.32}$$

а набор его значений в точках w^i в виде вектора

$$b = (a(w^0), a(w^1), ..., a(w^{n-1}),$$

то при фиксированном n ДПФ можно будет рассматривать как преобразование множества векторов R^n. Обозначим его через F_n, а обратное ему преобразование — через F_n^{-1}. Из определения F_n следует, что

$$F_n(a) = b = (b_0, b_1, ..., b_{n-1}),$$

где

$$b_i = \sum_{j=0}^{n-1} a_j w^{i \cdot j}, i = 1, 2, ..., n-1.$$

Отсюда нетрудно заметить, что вектор $F_n(a)$ можно получить, умножив вектор a на матрицу $W_n = (w^{ij})_{n \times n}$. Следующая теорема указывает пути реализации обратного ДПФ в случае его существования.

Теорема 2.19. *Если элемент nе обратим в R, то обратное БПФ F_n^{-1} существует, причем для вектора (2.32)*

$$F_n^{-1}(a) = (c_0, c_1, ..., c_{n-1}),$$

где

$$c_i = (ne)^{-1} \sum_{j=0}^{n-1} a_j w^{-i \cdot j} \ , \ i = 1, 2, ..., n-1.$$

□ Достаточно доказать, что для всех $a \in R^n$ справедливы равенства $F^{-1}(F(a)) = F(F^{-1}(a)) = a$.

Согласно определениям преобразований F_n и F_n^{-1} имеем $F_n^{-1}(F(a)) = (c_0, ..., c_{n-1})$, где

$$c_i = n^{-1} \sum_{j=0}^{n-1} \left(\sum_{k=0}^{n-1} a_k w^{jk} \right) w^{-ij} =$$

$$= (ne)^{-1} \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} a_k w^{j(k-i)} =$$

$$= n^{-1} \sum_{k=0}^{n-1} a_j \sum_{k=0}^{n-1} w^{j(k-i)} = n^{-1} a_i n = a_i$$

для всех $i = 1, ..., n-1$. Следовательно, $F_n^{-1}(F(a)) = a$. Аналогично доказывается, что $F(F^{-1}(a)) = a$. □

Из теоремы 2.19 видно, что вектор $F_n^{-1}(a)$ равен произведению вектора a на матрицу. Заметим, что имеем несколько способов вычисления прямого и обратного ДПФ вектора, а именно: прямое ДПФ можно вычислить, находя значения многочлена в соответствующих точках или умножая вектор на матрицу $W_n = (w^{ij})_{n \times n}$, а обратное — используя интерполяционные формулы или умножая на матрицу $n^{-1} \cdot (w^{-ij})_{n \times n}$.

Оценим функцию сложности задачи реализации ДПФ, считая размером ее входа число n и используя в качестве элементарных операций сложение, вычитание и умножение в кольце R.

Алгоритм реализации ДПФ F_n путем умножения векторов на матрицу требует, очевидно, n^2 умножений и $n(n-1)$ сложений и, значит, имеет сложность $O(n^2)$. Однако существует и более эффективный алгоритм реализации ДПФ F_n, называемый быстрым преобразованием Фурье (БПФ). В нем, как и в задачах предыдущего параграфа, используется рекуррентный метод. Для его описания будем предполагать, что $n = 2^t$.

Представим многочлен $a(x)$ в виде

$$a(x) = (a_0 + a_2x^2 + ... + a_{n-2}x^{n-2}) + \\ + x_1(a_1 + a_3x^2 + ... + a_{n-1}x^{n-2}).$$

Произведя замену $x^2 = y$, получим $a(x) = p(y) + xq(y)$, где $p(y), q(y)$ — многочлены, степени которых не превосходят $n/2 - 1$. Таким образом, для вычисления значения многочлена в точке w^i достаточно вычислить значения многочленов $p(y), q(y)$ в точке w^{2i}, а затем умножить еще w^i на $q(w^{2i})$.

Заметим, что w^2 является примитивным корнем степени $n/2$ из единицы и значения многочленов $p(y), q(y)$ нужно вычислять в точках $w_0, w^2, w^4, ..., w^{n-2}$, которые образуют множество всех различных степеней элемента w^2. В связи с этим для функции $T(n)$ сложности описанного алгоритма имеем соотношение

$$T(n) \leq 2T(n/2) + 2n.$$

Применяя к нему теорему 2.18, получаем $T(n) = O(n \log n)$.

Очевидно, что если w — примитивный корень n-й степени из единицы в кольце R, то то же самое верно и для w^{-1}. Отсюда и из теоремы 2.19 следует, что обратное БПФ имеет ту же асимптотическую сложность, что и прямое БПФ.

Использование алгоритма БПФ позволяет вычислять произведение двух многочленов из $R[x]$, сумма степеней которых меньше $n = 2^k$, со сложностью $O(n \log n)$ при условии существования в кольце примитивного корня w степени n из единицы и обратимости элемента ne. Действительно, для нахождения коэффициентов многочлена-произведения достаточно найти с помощью прямого БПФ векторы значений каждого из многочленов в точках w^i, $i = 0, ..., n - 1$, перемножить их покоординатно и, используя обратное БПФ, по значениям многочлена-произведения в n точках найти вектор его коэффициентов (по китайской теореме об остатках такой вектор найдется однозначно).

2.7. ПОНЯТИЕ О ВЕРОЯТНОСТНЫХ АЛГОРИТМАХ

Снижения сложности в решении некоторых массовых задач удается достигать путем применения так называемых вероятностных алгоритмов.

Отличительными особенностями вероятностного алгоритма являются:

1) использование в процессе работы алгоритма некоторого датчика случайных элементов некоторого конечного множества, влияющих на работу алгоритма и нарушающих тем самым его детерминированность;

2) получение ответов на некоторые индивидуальные задачи лишь с некоторыми положительными вероятностями, меньшими 1.

Таким образом, в вероятностных алгоритмах ускорение решения задачи достигается за счет частичной утраты достоверности ответов. В связи с этим вероятностные алгоритмы представляют особый интерес для решения таких задач, где можно довольствоваться правильным ответом лишь с некоторой, отличной от единицы вероятностью или где детерминированные алгоритмы не осуществимы в реальное время ввиду их большой сложности. Примером задачи первого типа может служить задача декодирования корректирующего кода, когда даже неполное исправление ошибок может оказаться достаточным для правильного восстановления информации за счет ее естественной избыточности. Задачи второго типа являются более распространенными и могут встречаться в самых разных приложениях математики. Примеры приводятся ниже.

Заметим, что вероятностный алгоритм не является алгоритмом в смысле введенного в предыдущей главе строгого определения понятия алгоритма, в котором каждый шаг работы алгоритма однозначно определен программой или схемой алгоритма. В связи с этим ниже, во избежание путаницы, алгоритмы в смысле такого строгого определения будут называться детерминированными.

Следует иметь в виду, что в вероятностных алгоритмах, как правило, заложена возможность повышения вероятности

правильных ответов за счет многократного повторения алгоритма при различных значениях, поступающих с датчика случайных элементов. Более того, во многих случаях таким путем, за счет перебора достаточно большого числа различных случайных элементов, вероятность правильности ответов можно довести до 1 и получить тем самым детерминированный алгоритм. Однако такая возможность заложена не в любом вероятностном алгоритме.

Приведем пример вероятностного алгоритма распознавания простоты нечетных натуральных чисел, называемого тестом Соловея–Штрассена [8].

Алгоритм основан на следующем известном из теории чисел критерии простоты целого числа: нечетное число n является простым тогда и только тогда, когда для любого числа $a \in (\mathbb{Z}_p)^*$ выполняется равенство

$$a^{\frac{n-1}{2}} = (\frac{a}{p}), \tag{2.33}$$

где $(\frac{a}{p})$ — символ Якоби.

Входом алгоритма является нечетное натуральное число n, выходом — число 1, если n — составное, и 0, если ответ неизвестен.

Работа алгоритма состоит в общем случае из трех шагов:

1) с помощью датчика случайных чисел выбираем нечетное число $a < n$;

2) находим наибольший общий делитель $(a, n) = d$.

Если $d > 1$, то n составное число, алгоритм закончен с выходом 1. Если же $d = 1$, то делаем шаг

3) проверяем равенство (2.33).

Если оно не выполнено, то n — составное и выход 1. В противном случае ответ неизвестен и выход 0.

Таким образом, при выходе 1 мы можем точно сказать, что n — составное, при выходе 0 точного ответа нет. Однако выполнение условия (2.33) делает естественным предположение, что n простое. Выясним, какова вероятность правильности такого предположения. Оценку этой вероятности можно получить из следующего утверждения.

Теорема 2.20. *Если n — нечетное составное число, то равенство (2.33) выполняется не более чем для 1/2 части всех значений a из $(\mathbb{Z}_n)^*$.*

□ Легко видеть, что множество G элементов группы $(\mathbb{Z}_n)^*$, удовлетворяющих условию (2.33), образует подгруппу. А так как порядок подгруппы делит порядок группы (по теореме Лагранжа), то остается лишь доказать, что эта подгруппа — собственная, т. е. найти хотя бы один элемент из $(\mathbb{Z}_n)^* \setminus G$. Рассмотрим любой элемент b наибольшего порядка группы $(\mathbb{Z}_n)^*$.

Пусть n имеет следующее каноническое разложение

$$n = p_1^{k_1}...p_s^{k_s}.$$

Из курса теории чисел известно, что мультипликативная группа $(\mathbb{Z}_n)^*$ изоморфна прямому произведению циклических групп

$$(\mathbb{Z}_{p_i^{k_i}})^*,\ i = 1, ..., s,$$

порядков соответственно

$$\varphi(p_i^{k_i}) = p_i^{k_i - 1}(p_i - 1),$$

где φ — функция Эйлера. Следовательно, элемент b имеет порядок $m = \Pi_{i=1}^{s}(p_i^{k_i-1}(p_i - 1))$.

Если $b \notin G$, то нужный элемент найден. Пусть $b \in G$. Тогда $b^{n-1} = 1$, и по свойству порядков элементов группы $m|(n-1)$. Отсюда следует, в частности, что $k_i = 1$, т. е. $n = p_1 p_2 ... p_s$, $s > > 1$. В этом случае искомым элементом может служить решение c системы сравнений:

$$x \equiv a \pmod{p_1}, x \equiv 1 \pmod{p_i},\ i = 1, 2, ..., s,$$

где a — первообразный корень по модулю p_1 (или примитивный элемент группы $\mathbb{Z}_{p_i}$). Из определения элемента c находим символ Якоби:

$$(\frac{c}{n}) = (\frac{c}{p_1})(\frac{c}{p_2})...(\frac{c}{p_s}) = -1.$$

Если $c \in G$, то $c^{\frac{n-1}{2}} \equiv -1 \pmod{n}$, а потому и

$$c^{\frac{n-1}{2}} \equiv -1 \pmod{p_i},\ i = 1, 2, ..., s,$$

что противоречит условию $c \equiv 1 \pmod{p_2}$. Значит, $c \notin G$, и теорема доказана. $\square$

Из теоремы 2.20 следует, что при выходе 0 наше предположение о простоте числа n выполняется с вероятностью $P \geq 1 - 1/2$. Следовательно, при k-кратном повторении алгоритма с попарно различными значениями случайно выбираемых элементов мы сможет добиться правильного ответа с вероятностью $P \geq 1 - 1/2^k$.

Заметим, что, применяя указанный алгоритм при всевозможных элементах $a \in (\mathbb{Z}_n)^*$, мы получим точный ответ на вопрос о простоте числа n, т. е. получим детерминированный алгоритм. Однако этот алгоритм будет существенно более сложным, чем k-кратно применяемый алгоритм A_2 при сравнительно небольших значениях k.

Приложение

МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ОРГАНИЗАЦИИ ИЗУЧЕНИЯ МАТЕМАТИЧЕСКОЙ ЛОГИКИ, ТЕОРИИ АЛГОРИТМОВ, ТЕОРИИ ДИСКРЕТНЫХ ФУНКЦИЙ

Анализ требований Федерального государственного образовательного стандарта высшего профессионального образования (ФГОС ВПО)

Рассмотрим требования ФГОС ВПО третьего поколения в области информационной безопасности, на выполнение которых ориентировано изучение математической логики, теории дискретных функций и теории алгоритмов. В наибольшей степени глубокие знания в этой области математики требуются при обучении по специальностям направления подготовки 090300 «Информационная безопасность вычислительных, автоматизированных и телекоммуникационных систем».

Закончив обучение по специальности 090301 «Компьютерная безопасность», выпускник должен обладать, среди прочих, следующими компетенциями:

— способностью применять математический аппарат, в том числе с использованием вычислительной техники, для решения профессиональных задач (ПК-2);

— способностью к самостоятельному построению алгоритма, проведению его анализа и реализации в современных программных комплексах (ПК-12).

Для реализации перечисленных компетенций в результате изучения дисциплины «Математическая логика и теория алгоритмов» и «Дискретная математика» студент должен:

Знать:

основные понятия математической логики и теории алгоритмов, язык и средства современной математической логики, представления булевых функций и способы минимизации

формул, типовые свойства и способы задания функций многозначной логики, различные подходы к определению алгоритма и доказательства алгоритмической неразрешимости отдельных массовых задач, подходы к оценкам сложности алгоритмов, методы построения эффективных алгоритмов, возможности применения общих логических принципов в математике и профессиональной деятельности.

Уметь:

находить и исследовать свойства представлений булевых и многозначных функций формулами в различных базисах, оценивать сложность алгоритмов и вычислений, классифицировать алгоритмы по классам сложности, применять методы математической логики и теории алгоритмов к решению задач математической кибернетики.

Владеть:

навыками использования языка современной символической логики, навыками применения методов и фактов теории алгоритмов, относящимися к решению переборных задач, навыками упрощения формул алгебры высказываний и алгебры предикатов.

Закончив обучение по специальности 090303 «Информационная безопасность автоматизированных систем», выпускник должен обладать, среди прочих, следующими компетенциями:

— способностью выявлять естественно-научную сущность проблем, возникающих в ходе профессиональной деятельности, и применять соответствующий физико-математический аппарат для их формализации, анализа и выработки решения (ПК-1);

— способностью применять математический аппарат, в том числе с использованием вычислительной техники, для решения профессиональных задач (ПК-2).

Для реализации перечисленных компетенций в результате изучения дисциплины «Математическая логика и теория алгоритмов» и «Дискретная математика» студент должен:

Знать:

основные принципы математической логики, формализации понятия алгоритма: машины Тьюринга, рекурсивные функции, основные понятия теории сложности алгоритмов,

представления булевых функций и способы минимизации формул, типовые свойства и способы задания функций многозначной логики, возможности применения общих логических принципов в математике и профессиональной деятельности.

Уметь:

находить и исследовать свойства представлений булевых и многозначных функций формулами в различных базисах, находить и упрощать формулы алгебры предикатов, оценивать сложность алгоритмов и вычислений.

Владеть:

способами оценки сложности работы алгоритмов, навыками применения методов и фактов теории алгоритмов, относящимися к решению переборных задач, навыками упрощения формул алгебры высказываний и алгебры предикатов, навыками составления программ на машинах Тьюринга.

Закончив обучение по специальности 090305 «Информационно-аналитические системы безопасности», выпускник должен обладать, среди прочих, следующими компетенциями:

— способностью применять математический аппарат и естественнонаучный аппарат для решения профессиональных задач, интерпретировать профессиональный смысл полученного математического результата (ПК-17);

— способностью самостоятельно строить алгоритм решения задачи, проводить его анализ и реализовывать в современных программных комплексах (ПК-25).

Для ее реализации в результате изучения дисциплин «Математическая логика и теория алгоритмов» и «Дискретная математика» студент должен:

Знать:

определение исчисления высказываний и основных понятий этого исчисления, методы преобразования произвольных формул исчисления высказываний в дизъюнктивные и конъюнктивные нормальные формы, проблему дедукции, ее решение методами прямой и обратной дедукции, метод резолюций для исчисления высказываний и его роль в решении проблемы дедукции, определение исчисления предикатов и основных понятий этого исчисления, метод резолюций для исчисления предикатов, определение, свойства аксиоматических систем и

приемы работы с ними, определение и классы машин Тьюринга и их роль в теории алгоритмов.

Уметь:

формулировать задачи логического характера в рамках исчисления высказываний и исчисления предикатов, проводить исследование логических формул для доказательства их свойств, применять метод резолюций для решения проблемы дедукции в исчислении высказываний и исчислении предикатов, проводить доказательства в рамках аксиоматических систем, формулировать и решать задачи, пользуясь соответствующими классами машин Тьюринга.

Владеть:

навыками выполнения преобразования логических формул с использованием схем тождественных преобразований.

Изучение математической логики, теории дискретных функций и теории алгоритмов может также потребоваться обучающимся по другим специальностям для реализации компетенций, предусмотренных требованиями ФГОС ВПО третьего поколения в области информационной безопасности.

Цели и задачи дисциплины, замечания по методике преподавания

Целью изучения дисциплины является теоретическая и практическая подготовка специалистов по основам математической логики, методам оценки сложности алгоритмов и построения эффективных алгоритмов, а также обеспечение фундаментальной подготовки в одной из важнейших областей современной математики.

Задачами дисциплины являются: формирование научного мировоззрения, понимания широты и универсальности методов математической логики, умения применять эти методы в решении прикладных задач; развитие творческого мышления, математической грамотности, способности критически анализировать собственные рассуждения и самостоятельно их корректировать; воспитание математической культуры, которая предполагает четкое осознание необходимости и важности математической подготовки для специалиста в области информационной безопасности; ознакомление с основными объектами математической логики, а также их приложениями для реше-

ния различных задач, требующих применения вычислительных средств; ознакомление слушателей с принципами логического программирования; основными задачами теории алгоритмов и их приложениями к задачам математической кибернетики; выработка навыков обращения с дискретными конструкциями и умения строить математические модели объектов и процессов, с которыми имеет дело специалист в ходе своей профессиональной деятельности. Таким образом, дисциплина «Математическая логика и теория алгоритмов» является неотъемлемой частью профессиональной подготовки по направлению «Информационная безопасность».

Дисциплина «Математическая логика и теория алгоритмов» принадлежит базовой части математического и естественнонаучного цикла. Для успешного усвоения данной дисциплины необходимо, чтобы обучаемый владел знаниями, умениями и навыками, сформированными в процессе изучения дисциплин:

Математический анализ — основы теории пределов и действительных функций одного переменного;

Алгебра и геометрия — основы общей и линейной алгебры;

Дискретная математика — основы комбинаторики и теории графов.

Знания, полученные при изучении дисциплины «Математическая логика и теория алгоритмов», используются при изучении дисциплин «Технологии и методы программирования», «Безопасность систем баз данных», «Криптографические методы защиты информации».

Рекомендации по организации учебного процесса

Особенности данного учебного пособия связаны со спецификой основных целей изучения, указанных в его названии разделов математики для будущих специалистов в области защиты информации. Эти цели преследуют, с одной стороны, расширение фундаментального математического образования, ознакомление с проблемами оснований математики и общее развитие логического мышления, а с другой стороны, систематическое изучение основных объектов, используемых в задачах синтеза и анализа криптографических алгоритмов.

Материал пособия разбит на три взаимосвязанные части: математическая логика, дискретные функции и теория алгоритмов.

Первая глава пособия носит вспомогательный характер. Она посвящена основным понятиям всей математики — понятиям множества, операций над множествами, отображений множеств, отношений и операций на множествах. К моменту начала изучения математики в рамках данного пособия указанный материал мог быть изучен в других математических дисциплинах. В этом случае на лекциях можно ограничиться лишь беглым обзором и введением необходимых обозначений, а на практических занятиях решить ряд задач на закрепление знаний по этому материалу.

Все основные узлы современных шифрсистем конструируются с использованием дискретных функций, к которым предъявляется ряд специальных требований. При этом чаще всего используются булевы функции. В связи с этим при изучении алгебры высказываний (называемой также алгеброй логики) необходимо уделить особое внимание различным видам формул алгебры логики, их преобразованиям к эквивалентным формулам с использованием основных законов логики, их нормальным формам и минимизации формул.

Результаты о формулах алгебры логики существенно используются при изучении булевых функций во второй части пособия. А именно, формулы двоичной и k-значной алгебры логики в различных базисах служат основным средством задания и изучения булевых функций и функций на конечных полях и кольцах вычетов. При изучении материала о дискретных функциях следует больше внимания уделить способам задания функций многих переменных, вопросу о полноте различных систем функций, классификации функций относительно некоторых групп преобразований, группам инерции функций в этих группах и другим криптографическим параметрам функций, декомпозиции функций.

Раздел «Алгебра предикатов» имеет своей целью формализацию содержательных математических утверждений, т. е. их краткую и наглядную запись формулами с использованием некоторого набора основных логических операций над преди-

катами. Следует иметь в виду, что в книгах по математической логике используются различные наборы основных операций, причем зачастую стремятся к минимизации числа операций. Однако уменьшение числа основных операций ведет к усложнению формул и к потере содержательной наглядности формул. В данном пособии выбран избыточный, но наиболее естественный набор, содержащий основные операции алгебры логики и навешивания кванторов всеобщности и существования. Кроме того, при определении формулы мы запрещаем навешивать кванторы по связанным переменным. На эти моменты целесообразно обратить внимание на лекциях, а на практических занятиях решить задачи по выделению из выбранного набора минимальных подсистем, через которые выражаются формулами остальные операции. Важным также является понятие эквивалентности формул, позволяющее упрощать формулы и, в частности, приводить их к предваренным нормальным формам. И на лекции, и на практическом занятии целесообразно также привести примеры не эквивалентных формул, но эквивалентных на некоторых алгебраических системах.

Заметим, что алгебра высказываний является частью алгебры предикатов, поскольку каждое высказывание можно рассматривать как нуль-местный предикат. Однако с методической точки зрения и, учитывая важную самостоятельную роль алгебры высказываний в приложениях, более правильным представляется отдельное изучение алгебры высказываний до алгебры предикатов.

Исчисления высказываний и предикатов являются важнейшими логическими исчислениями. Они доставляют основное средство формализации доказательств математических утверждений. Их основные и вспомогательные правила вывода и доказательства формул повседневно (иногда не вполне осознанно) используются нами в доказательствах теорем. Поэтому знание этих правил приводит в порядок наше мышление и делает обоснованными наши рассуждения. На практических занятиях следует потренироваться выводить и доказывать формулы, а также формализовать доказательства некоторых известных теорем.

Вместе с тем естественно возникают вопросы о самих логических исчислениях. А именно, можно ли доказать тождественно истинную формулу средствами данного логического исчисления (проблема полноты) и не может ли в нем быть доказанной некоторая формула и ее отрицание (проблема непротиворечивости), существует ли алгоритм, позволяющий для любой формулы выяснять, доказуема эта формула или нет (проблема разрешимости). Несложно доказывается, что любая доказуемая формула исчисления предикатов является тождественно истинной в алгебре предикатов, откуда следует непротиворечивость самого исчисления предикатов. Обратное утверждение, составляющее известную теорему Гёделя о полноте, доказывается довольно сложно, и при ограниченном числе лекционных часов его достаточно пояснить, не доказывая.

Проблема разрешимости исчисления (и логики) предикатов решена А. Черчем отрицательно. Однако Дж. Эрбраном был указан алгоритм распознавания доказуемости тождественно истинных формул, который после ряда усовершенствований получил название «метода резолюций». В пособии приведена схема этого алгоритма, и на практических занятиях целесообразно решить несколько задач на применение метода резолюций, поскольку он находит применение в анализе криптографических протоколов.

Последняя часть пособия посвящена введению в теорию алгоритмов и теорию сложности алгоритмов. Здесь приводятся три подхода к определению понятия алгоритма — через машины Тьюринга, нормальные алгоритмы Маркова и рекурсивные функции. Необходимо подчеркнуть, что точное определение алгоритма понадобилось для доказательства отсутствия алгоритмов решения некоторых массовых задач.

Для специалистов в области защиты информации, имеющих дело в основном с конечными объектами, важнейшую роль играют вопросы сложности алгоритмов и сложности переборных задач. При этом точные определения алгоритмов используются как модели для определения сложности вычислений. Поэтому на практических занятиях следует потренироваться строить машины Тьюринга и схемы нормальных алгоритмов для решения некоторых конкретных задач в коль-

це целых чисел, в кольцах вычетов и в кольцах полиномов над конечными полями, а также оценивать асимптотическую сложность таких задач.

Все задачи переборного характера на конечных множествах принципиально разрешимы, однако существует определенная классификация таких задач по временно́й сложности их решения. Наиболее важными классами являются класс P задач полиномиальной сложности и класс NP недетерминированно полиномиальной сложности. Среди последних особую роль играют NP-полные задачи. Хотя строго и не доказано, что NP-полные задачи не содержатся в классе P, тем не менее доказательство NP-полноты задачи является дополнительным аргументом для отнесения ее к сложным задачам. Так как в системах по защите информации вопросы сложности переборных задач носят принципиальный характер, то желательно, чтобы будущий специалист в этой области имел определенный запас NP-полных задач и знал общий подход к доказательству NP-полноты некоторых задач.

Многие задачи анализа современных шифрсистем можно свести к решению уравнений в конечных полях и, в частности, булевых уравнений. В связи с этим представляют интерес такие системы булевых уравнений, для которых существуют полиномиальные алгоритмы распознавания совместности и нахождения решений. В связи с этим в пособии рассмотрены классы булевых уравнений, составленных из так называемых шеферовых функций. Решению таких систем уравнений необходимо уделить внимание на практических занятиях.

В заключение отметим, что необходимый перечень задач и упражнений по всем разделам пособия имеется в [75]. Кроме того, ряд упражнений для самостоятельной работы приведен в тексте пособия. Эти упражнения необходимо выполнять в ходе проработки соответствующего материала.

ЛИТЕРАТУРА

1. *Агафонов В. Н.* Сложность алгоритмов и вычислений. — Новосибирск : НГУ, 1975.
2. *Адян С. И.* Неразрешимость некоторых алгоритмических проблем теории групп // Тр. Моск. матем. о-ва. — 1957. — № 6. — С. 231–298.
3. *Алексеев В. Б.* Лекции по дискретной математике. — М. : МГУ им. М. В. Ломоносова, 2004.
4. *Амбросимов А. С.* Некоторые асимптотические разложения для числа функций с нетривиальной группой инерции / А. С. Амбросимов, Н. Н. Шаров // Проблемы кибернетики. — 1979. — № 36. — С. 65–84.
5. *Ахо А.* Построение и анализ вычислительных алгоритмов / А. Ахо, Дж. Хопкрофт, Дж. Ульман — М. : Мир, 1979.
6. *Блейхут Р.* Быстрые алгоритмы цифровой обработки сигналов. — М. : Мир, 1989.
7. *Вариченко Л. В.* Абстрактные алгебраические системы и цифровая обработка сигналов / Л. В. Вариченко, В. Г. Лабунец, М. А. Раков. — Киев : Наук. думка, 1988.
8. *Василенко О. Н.* Теоретико-числовые алгоритмы в криптографии. — М. : МЦМНО, 2003.
9. *Ван дер Варден Б. Л.* Алгебра. — СПб. : Лань, 2004.
10. *Гаврилов М. А.* Теория релейно-контактных схем. — М. ; Л. : Изд-во АН СССР, 1950.
11. *Гейтинг А.* Интуиционизм. — М. : Мир, 1965.
12. *Глухов М. М.* Алгебра высказываний и двоичные функции. — М., 1971.
13. *Глухов М. М.* Математическая логика. — М., 1981.
14. *Глухов М. М.* Алгебра : в 2 т. / М. М. Глухов, В. П. Елизаров, А. А. Нечаев. — М. : Гелиос АРВ, 2003.
15. *Горшков С. П.* Применение теории NP-полных задач для оценки сложности решения систем булевых уравнений // Обозрение прикл. и пром. мат. — 1995. — Том 2. — Вып. 3. — С. 325–467.
16. *Гэрри М.* Вычислительные машины и труднорешаемые задачи / М. Гэрри, Д. Джонсон. — М. : Мир, 1982.

17. *Драгалин А. Г.* Математический интуиционизм. Введение в теорию доказательств. — М.: Наука, 1979.
18. *Ершов Ю. Л.* Математическая логика / Ю. Л. Ершов, Е. А. Палютин. — СПб.: Лань, 2005.
19. Элементарные теории / Ю. Л. Ершов [и др.] // УМН, 20. — 1965. — № 4. — С. 37–108.
20. *Карри Х.* Основания математической логики. — М.: Мир, 1966.
21. *Касим-Заде О. М.* О влиянии базиса на мощность схем из функциональных элементов. Препринт ин-та прикладной математики АН СССР. — 1979. — № 122.
22. *Кейслер Г.* Теория моделей / Г. Кейслер, Ч. Ч. Чэн. — М.: Мир, 1977.
23. *Келли Дж. Л.* Общая топология. — М.: Наука, 1968.
24. *Клини С.* Введение в метаматематику. — М.: ИЛ, 1957.
25. *Клосс Б. М.* О классификации функций многозначной логики / Б. М. Клосс, Э. И. Нечипорук // Проблемы кибернетики. — 1963. — № 9. — С. 27–36.
26. *Колдуэлл С.* Логический синтез релейных устройств. — М.: ИЛ, 1962.
27. *Колмогоров А. Н.* Три подхода к определению понятия информации // Проблемы передачи информации. — 1965. — № I. С. 3–11.
28. *Колмогоров А. Н.* К определению алгоритма / А. Н. Колмогоров, В. А. Успенский // УМН, 13. —1958. — № 4. — С. 3–28.
29. *Кормен Т.* Алгоритмы. Построение и анализ / Т. Кормен, Ч. Лейзерсон, Р. Ривест. — М.: МЦНМО, 1999.
30. *Кузнецов А. В.* О бесповторных контактных схемах и бесповторных суперпозициях функций алгебры логики // Тр. МИАН СССР. — 1958. — № 51. — С. 186–225.
31. *Кузьмин В. А.* Реализация функций алгебры логики автоматами, нормальными алгоритмами и машинами Тьюринга // Проблемы кибернетики. — 1965. — № 13. — С. 75–96.
32. *Курош А. Г.* Лекции по общей алгебре. — СПб.: Лань, 2002.
33. *Лабунец В. Г.* Гармонический анализ булевых функций и функций k-значной логики над конечным полем / В. Г. Лабунец, О. П. Ситников // Техническая кибернетика. — 1975. — № 1.
34. *Лавров И. А.* Логика и алгоритмы. — Новосибирск: НГУ, 1970.
35. *Логачев О. А.* Булевы функции в теории кодирования и криптографии / О. А. Логачев, А. А. Сальников, В. В. Ященко. — М.: МЦНМО, 2004.
36. *Лупанов О. Б.* О синтезе некоторых классов управляющих систем // Проблемы кибернетики. — 1963. — № 10. — С. 63–98.
37. *Мальцев А. И.* Алгебраические системы. — М.: Наука, 1970.

38. *Мальцев А. И.* Алгоритмы и рекурсивные функции. — М.: Наука, 1965.
39. *Мальцев А. И.* Избранные труды. — М.: Наука, 1976. — Т. I–II.
40. *Марков А. А.* Теория алгорифмов // Тр. Матем. ин-та им. В. А. Стеклова, 1954. — Вып. 42.
41. *Марков А. А.* Избранные труды. — М.: МЦНМО, 2002–2003. — Т. I–II.
42. *Марков А. А.* О нормальных алгорифмах, связанных с вычислением булевых функций // Изв. АН СССР, 31. — 1967. — № 1. — С. 161–208.
43. *Матиясевич Ю. В.* Диофантовость перечислимых множеств // ДАН СССР, 191. — 1970. — С. 279–282.
44. *Мендельсон Э.* Введение в математическую логику. — М.: Наука, 1971.
45. *Мур Э.* Надежные схемы из ненадежных реле. Работы по теории информации и кибернетике / Э. Мур, К. Шеннон. — М., 1963. С. 114–153.
46. *Новиков П. С.* Элементы математической логики. — М.: Наука, 1973.
47. *Новиков П. С.* Избранные труды. — М.: Наука, 1979.
48. *Ноден П.* Алгебраическая алгоритмика / П. Ноден, К. Китте. — М.: Мир, 1999.
49. *Носов В. А.* Основы теории алгоритмов и анализа их сложности. — М., 1992.
50. *Поваров Г. Н.* Математическая теория синтеза контактных $(1, k)$-полюсников. // ДАН СССР, 100. — 1955. — № 5.
51. Проблемы Гильберта. — М.: Наука, 1969.
52. *Перязев Н. А.* Реализация булевых функций бесповторными формулами // Дискретная математика, 7:3. — 1995. С. 61—68.
53. *Проскуряков И. В.* Числа и многочлены. — М.: Просвещение, 1965.
54. *Робинсон А.* Введение в теорию моделей и метаматематику алгебры. — М.: Наука, 1967.
55. *Робинсон Дж.* Машинно-ориентированная логика основанная на принципе резолюции // Киб. сборник. Новая серия. М.: Мир, 1970. — Вып. 7. — С. 194–218.
56. *Рябинин И. А.* Логико-вероятностное исчисление как аппарат исследования надежности и безопасности структурно-сложных систем. — М.: АиТ: Наука, 2003. № 7. С. 178–186.
57. *Сачков В. Н.* Введение в комбинаторные методы дискретной математики. — 2-е изд. — М.: МЦНМО, 2004.
58. *Стяжкин Н. И.* Формирование математической логики. — М.: Наука, 1967.

59. *Тарский А.* Введение в логику и методологию дедуктивных наук. — М.: ИЛ, 1948.
60. *Трахтенброт Б. А.* Алгоритмы и вычислительные автоматы. — М.: Сов. радио, 1974.
61. *Успенский В. А.* Лекции о вычислимых функциях. — М.: Физматлит, 1960.
62. *Фомичев В. М.* Дискретная математика и криптология. Курс лекций. — М.: Диалог-МИФИ, 2003.
63. *Френкель А.* Основания теории множеств / А. Френкель, И. Бар-Хиллел. — М.: Мир, 1966.
64. *Чень Ч.* Математическая логика и автоматическое доказательство / Ч. Чень, Р. Ли. — М.: Наука, 1983.
65. *Черемушкин А. В.* Бесповторная декомпозиция сильно зависимых функций // Дискретная математика, 16:3. — 2004. — С. 3—42.
66. *Шеннон К.* Число двухполюсных параллельно-последовательных сетей. Работы по теории информации и кибернетике. — М.: ИЛ, 1963. С. 46–58.
67. *Шеннон К.* Синтез двухполюсных переключательных схем. Работы по теории информации и кибернетике. — М.: 1963. — С. 59–106.
68. *Шоломов Л. А.* Основы теории дискретных логических и вычислительных устройств. — М.: Наука, 1980.
69. *Яблонский С. В.* Введение в дискретную математику. — М.: Наука, 1979.
70. *Яблонский С. В.* Функции алгебры логики и классы Поста / С. В. Яблонский, Г. П. Гаврилов, В. Б. Кудрявцев. — М.: Наука, 1966.
71. *Ashenhurst R. L.* The decomposition of switching functions // Ann. Comput. Labor. Harvard Univ. — 1959. — № 29. — P. 74–116.
72. *Schaefer T.* Complexity of satisfiability probles // Proceedings of the 10 Annual ACM Symposium on Theory of Computing. — 1978. — P. 216–226.

Сборники задач

73. *Гаврилов Г. П.* Сборник задач по дискретной математике / Г. П. Гаврилов, А. А. Сапоженко. — М.: Физматлит, 2005.
74. *Гиндикин С. Г.* Алгебра логики в задачах. — М.: Наука, 1972.
75. Задачи и упражнения по математической логике, дискретным функциям и теории алгоритмов / М. М. Глухов [и др.] — СПб.: Лань, 2008.
76. *Лавров И. А.* Задачи по теории множеств, математической логике и теории алгоритмов / И. А. Лавров, Л. Л. Максимова. — М.: Наука, 1995.

ОГЛАВЛЕНИЕ

Михаил Михайлович ГЛУХОВ
Алексей Борисович ШИШКОВ

МАТЕМАТИЧЕСКАЯ ЛОГИКА ДИСКРЕТНЫЕ ФУНКЦИИ ТЕОРИЯ АЛГОРИТМОВ

Учебное пособие

Зав. редакцией физико-математической литературы
О. А. Митрофанова
Корректор *Т. А. Кошелева*
Подготовка иллюстраций *Е. В. Ляпусова*
Верстка *А. Г. Сандомирская*
Выпускающие *Н. В. Черезова, Е. П. Королькова*

ЛР № 065466 от 21.10.97
Гигиенический сертификат 78.01.07.953.П.007216.04.10
от 21.04.2010 г., выдан ЦГСЭН в СПб

Издательство «ЛАНЬ»
lan@lanbook.ru; www.lanbook.com
192029, Санкт-Петербург, Общественный пер., 5.
Тел./факс: (812) 412-29-35, 412-05-97, 412-92-72.
Бесплатный звонок по России: 8-800-700-40-71

Подписано в печать 13.08.12.
Бумага офсетная. Гарнитура Школьная. Формат 84×108 $^{1}/_{32}$.
Печать офсетная. Усл. п. л. 21,84. Тираж 1500 экз.

Заказ № .

Отпечатано в полном соответствии
с качеством предоставленных диапозитивов
в ОАО «Издательско-полиграфическое предприятие «Правда Севера».
163002, г. Архангельск, пр. Новгородский, д. 32.
Тел./факс (8182) 64-14-54; www.ippps.ru