Вадим Дунаев

Занимательная математика

Множества и отношения

Санкт-Петербург

«БХВ-Петербург»

2008

УДК 681.3.06
ББК 32.973.26
Д83

Дунаев В. В.

Д83 Занимательная математика. Множества и отношения. — СПб.: БХВ-Петербург, 2008. — 336 с.: ил.

ISBN 978-5-94157-988-4

Книга в занимательной форме вводит читателя в мир математики и логики. Она адресована всем, кто любит поразмышлять и интересуется головоломками и парадоксами. Материал первой части изложен в форме диалогов Профессора, Простака и Зануды. На занимательных примерах и задачах читатель приобщается к алгебре логики и элементам теории множеств и постоянно встречается с парадоксальными ситуациями, пытаясь их разрешить. Для всех предлагаемых задач приведены развернутые решения. Во второй части рассказывается о теории отношений и ее применении к таким практическим вещам, как реляционные базы данных и классификационная деятельность.

Для широкого круга читателей

УДК 681.3.06
ББК 32.973.26

Группа подготовки издания:

Главный редактор	*Екатерина Кондукова*
Зам. главного редактора	*Игорь Шишигин*
Зав. редакцией	*Григорий Добин*
Редактор	*Татьяна Темкина*
Компьютерная верстка	*Натальи Караваевой*
Корректор	*Виктория Пиотровская*
Дизайн обложки	*Инны Тачиной*
Зав. производством	*Николай Тверских*

Лицензия ИД № 02429 от 24.07.00. Подписано в печать 26.09.07.
Формат $70\times100^1/_{16}$. Печать офсетная. Усл. печ. л. 27,09.
Тираж 2000 экз. Заказ №
"БХВ-Петербург", 194354, Санкт-Петербург, ул. Есенина, 5Б.

Санитарно-эпидемиологическое заключение на продукцию
№ 77.99.02.953.Д.006421.11.04 от 11.11.2004 г. выдано Федеральной службой
по надзору в сфере защиты прав потребителей и благополучия человека.

Отпечатано с готовых диапозитивов
в ГУП "Типография "Наука"
199034, Санкт-Петербург, 9 линия, 12

ISBN 978-5-94157-988-4

Оглавление

Светлой памяти отца, привившего мне любовь к математике,
а также сыну, для которого сделать того же я не сумел

Введение

И хотя в подобных случаях трудно дать общие предписания и каждый должен в них следовать указаниям собственного разума, я попытаюсь все же указать путь начинающим.

И. Ньютон

Математика увлекательна и занимательна совсем не потому, что некоторые ее разделы могут быть изложены как анекдоты и головоломки, доступные пониманию широкого круга читателей. Она интересна главным образом потому, что в сжатой, но емкой форме выражает способности и особенности нашего мышления и психологии познания окружающего мира. Математики сначала ставят себе простые, с обывательской точки зрения, задачи, а затем показывают, что не все так просто, как кажется. Простые, на первый взгляд, задачи могут вообще не иметь решения, а решение, которое все-таки предлагается, оказывается решением совсем другой задачи. Математика, в отличие от других наук, не устаревает ни в одном из своих многочисленных разделов. Истины, достигнутые много веков назад, остаются истинами и сегодня, не зависимо от достижений техники, естественных и философских наук. Нематематикам математика представляется эталоном стремления к истине, определенности и доказательности логических построений. Вместе с тем, математика поучительна для нас еще и тем, как она применяет аналогии и обобщение. Мы и в повседневной жизни нередко прибегаем к методу рассуждений по аналогии, но в математике он наиболее ясно выражен. Однако в математике аналогия — не аргумент, не прецедент, на который можно сослаться как на основание для последующих умозаключений, а лишь путеводная нить для доказательства истинности или ложности некоторого утверждения. Аналогия для математика — это инструмент и технология его кухни, в которую он не любит пускать посторонних. Решая какую-то задачу, в том числе поставленную насущной практикой, математик заинтересован не столько в получении правильного ответа на нее, сколько в разработке метода решения более общей задачи. Иначе говоря, он стремится к возможно большей абстракции от кучи мелких и чересчур конкретных деталей чисто практической задачи, чтобы увидеть ее настоящую сущность и выработать наиболее универсальный метод решения, который, возможно, пригодится и в многочисленных случаях аналогичных частных задач.

Занимательнее всего в математике, пожалуй, не столько интригующие и увлекательные формулировки задач и неожиданные ответы на них, сколько порой драматический поиск решений, собственно мыслительный процесс. Правильный ответ, конечно, важен, но это не главная цель данной книги.

Эта книга — о математике не для математиков, и написана она не математиком. Она адресована всем, кто любит поразмышлять или порешать задачки и, в первую очередь, тем, кто считает свою направленность к людям гуманитарной, кому чужды точные науки, — психологам, социологам, юристам, журналистам, врачам, биологам и т. п., а также школьникам старших классов, которые еще не выбрали свою будущую стезю. Для чтения этой книги не предполагается сколько-нибудь серьезная математическая подготовка, — знаний в пределах средней школы, даже изрядно подзабытых, вполне достаточно.

Добавлю еще, что эта книга, надеюсь, окажется небесполезной будущим и уже состоявшимся инженерам, а также программистам, как изучавшим прикладную математику в вузах, так и самодеятельным. Программисты, знающие только языки общения с компьютерами, но не понимающие математическую сущность алгоритмов создаваемых ими программ, являются не более чем просто кодировщиками или весьма ограниченными переводчиками, которые переводят слова (если не буквы), но не мысли.

Данная книга посвящена лишь некоторым разделам необъятной математики — множествам и отношениям. Что может быть проще понятия множества? Под множеством мы понимаем практически все, что угодно. Это — многообразие, совокупность различающихся чем-то вещей, но рассматриваемая как некая целостность. Одно дело бесконечно разнообразный в своих деталях мир, другое — его часть или весь этот мир в моем поле зрения, рассматриваемый как нечто целое или относящееся к чему-то одному, подпадающий под некоторую общую идею. Простая в своей сущности идея множества сначала была положена в основу очень специфических задач математики, а затем вдруг оказалось, что ее можно попытаться развить в качестве фундамента математики в целом. Этому способствовал общий кризис математики, долго зреющий, но явно обнаружившийся только в конце XIX века. Тогда лучшие умы мира самоотверженно бросились спасать самую лучшую, с точки зрения определенности, обоснованности и истинности, науку. Они не могли допустить краха веры в непогрешимость математики. Ведь она и создавалась веками как особая наука об истине, как демонстрация могущества человеческого разума. Только во вторую очередь ее рассматривали как собрание методов решения насущных инженерных и перспективных прикладных задач. Драма математики конца XIX — первой половины XX веков увлекательна и поучительна сама по себе. Однако это не драма одной лишь математики, это — кризис чрезмерной самоуверенности человеческого разума и порожденной им науки, это — многими ожидаемая расплата за гордыню. Однако в данной книге мы лишь коснемся данной проблемы, не останавливаясь на нюансах. Эта драма должна нас научить быть более внимательным, а также примириться с мыслью о том, что истина не абсолютна — она относительна. Она где-то рядом, но не лежит на поверхности, доступная и тривиально-очевидная. Ее, истины, в большинстве случаев надо дознаться, потратив на это свои интеллектуальные и эмоциональные ресурсы, т. е. испытав сначала недоумение, раздражение и лишь потом — надежду и восторг. Последнее и будет настоящей наградой за все перипетии пройденного пути к тому, что называют истиной, пониманием, осознанием или, короче, жизнью интеллекта и чувств.

Немного уяснив понятие множества, мы должны ощутить готовность двигаться дальше. А именно: мы должны разобраться в том, что такое отношения между объектами различной природы. При этом нам не важно, находятся ли конкретные индивидуумы в интересующем нас отношении, например, любят ли они друг друга, больше ли один другого или главнее. Для нас важно понять, чем одни отношения отличаются от других вне зависимости от того, как конкретно они определены, между какими объектами и по каким правилам или признакам. Как бы ни было задано конкретное отношение между объектами из одного, двух или более множеств, мы, при более пристальном рассмотрении, можем установить, что это, например, отношение сходства, или эквивалентности, или порядка. А установив тип отношения, будучи вооруженными знаниями о нем, мы сможем чисто логически вывести другие важные следствия о данных конкретных отношениях и увеличить свою мощь в понимании мира, где как будто все взаимосвязано, все со всем находится в отношениях или вступает в определенные отношения.

Исключительно практическая область деятельности, такая как синтез и применение реляционных баз данных, покоится на теории отношений. Можно с уверенностью утверждать, что успех реляционных баз данных объясняется счастливым симбиозом насущных практических потребностей с адекватной теорией, возникшим совсем недавно — в середине XX века. Теперь мы можем изучать реляционные базы данных как инженерно-программистские конструкции, обросшие к настоящему времени множеством деталей и нюансов. А можем взглянуть на это сложное технологическое хозяйство с позиций теории, чтобы абстрагироваться от затемняющих суть дела деталей и преходящих технических усовершенствований. В данной книге мы рассмотрим некоторые из основных проблем реляционных баз данных именно с позиций теории отношений.

В теории множеств любую (или почти любую) совокупность объектов можно рассматривать как некий новый объект, называемый множеством. Этот тезис развязывает нам руки в творении множеств. Однако жизнь показывает, что творчество в данном направлении не вполне произвольно. В зависимости от интересов исследования и практических нужд мы создаем одни множества, а другие — нет. Наиболее ярко синтез множеств проявляется в так называемой классификационной деятельности, когда явно требуется разделить объекты некоторой предметной области на классы (т. е. множества). Почему дети относят китов и дельфинов к рыбам, а мы, взрослые, — нет? Почему даже самые маленькие дети отличают жуков и бабочек от птиц, хотя и те, и другие летают? Почему дети научаются считать быстрее, чем сравнивать количества? Не потому ли, что для счета требуется лишь научиться переходить от одного элемента к другому, называя их разными именами, а для сравнения совокупностей нужна предварительная абстракция от всего, что в них есть индивидуального? И то, и другое — различные проявления единого понятия числа. Это все не просто психология, это — математика. Математика и есть концентрированная психология, только обращенная не к чувственной жизни тела, а к парению духа.

Книга состоит из двух частей. Материал *первой части* (*главы 1*, *2* и *3*) изложен в форме диалогов, которая позволила мне задавать любые, в том числе и глупые вопросы. Глупые вопросы обычно стесняются задавать, но именно в них содержится нечто ("сермяжная правда"), благодаря чему учитель может распознать суть проблемы: то ли дело в недостаточном педагогическом искусстве изложении, то ли в дефектах собственно постановки и методов решения задачи. Для всех предлагаемых задач приведены развернутые решения, но я советую не искать в тексте правильные ответы сразу же, а прежде попробовать "вычислить" их самостоятельно. В *первой части* книги мы начинаем с задач, похожих на головоломки, а затем приобщаемся к алгебре логики и элементам теории множеств. При этом мы постоянно встречаемся с парадоксальными ситуациями, пытаясь их разрешить. Это и есть, на мой взгляд, математическое творчество, разговору о котором посвящена книга. Замышляя персонажи диалога как символы со вполне реальными прототипами, позже я понял, что они — мои собственные ипостаси. Так что это я сам рассказываю своим alter ego то, что им по силам со мной обсуждать. Один из них смел и креативен, другой осмотрителен и аккуратен.

Во *второй части* (*главы 4—7*) рассказывается о теории отношений и ее применении к таким практическим вещам, как реляционные базы данных и классификационная деятельность. Здесь изложение ведется не в форме диалога. Можно сказать, что это лекции. Пусть вас не пугает обилие символики. Она просто позволяет сделать изложение более кратким, а взаимосвязь отдельных положений — более отчетливой. В отличие от первой части, здесь вы найдете меньше увлекательных историй, но больше сведений, хотя для их усвоения потребуется, возможно, больше усилий. Доказательства теорем приведены не столько для вящей убедительности их формулировок, сколько для того, чтобы предоставить возможность поупражняться в обращении с рассматриваемыми объектами. Хотя, должен подчеркнуть, именно изучение доказательств теорем и есть настоящее изучение математики. Результат не так важен, как способ его получения. Впрочем, при первом чтении доказательства можно и пропустить без особого ущерба для понимания последующего материала. Попутно я хотел бы обратить внимание на то, что чем более интуитивно понятен тезис, тем более громоздким выглядит его доказательство. Интуитивно понятные вещи укоренились глубоко в нашем подсознании, а доказательство — попытка вывести их на уровень сознания, где правят логика и речь. Вытащить то, что и так понятно, на уровень сознания — нелегкая задача. По мере изложения материала в этой части темп и видимая сложность нарастают, но если вы неспешно начинали с первых глав, то данное обстоятельство вряд ли будет вам серьезной помехой. Дорогу осилит идущий.

В списке литературы приведены источники, которые я так или иначе использовал. Первые по списку книги я рекомендую тем, кто интересуется математикой и наукой вообще, а последующие упоминаю просто как источники фактов и доказательств, которым в данной, небольшой по объему, книге не нашлось места, либо они слишком сложны для неподготовленного читателя. В этом списке я не указал Мартина Гарднера, замечательного американского популяризатора науки XX века. Ссылки на его великолепные статьи и книги, а также выдержки из них легко найти в Интернете. Я и сам использовал в данной книге некоторые из его замечательных идей.

Между тем, чего хотелось, и тем, что действительно получилось, всегда есть разница. Я буду благодарен всем читателям, оставившим свой отзыв в моей гостевой книге по адресу **http://dunaevv1.narod.ru**.

Благодарности

Эта книга явилась результатом воспоминаний о плодотворной и интересной совместной научно-исследовательской и преподавательской работе с Олегом Маратовичем Поляковым, Владиславом Станиславовичем Блюмом и Владимиром Владимировичем Фроловым на кафедре, возглавляемой тогда Виктором Ивановичем Белицким, которым я искренне благодарен.

Я очень признателен Татьяне Темкиной за плодотворное обсуждение рукописи и придуманные ею иллюстрации, которые превосходно, на мой взгляд, воплотили в графике Инна Тачина и Наталья Караваева.

Особую благодарность я выражаю своей жене Валентине за поддержку во время работы над этой книгой.

ЧАСТЬ I

И В ШУТКУ, И ВСЕРЬЕЗ

Уважаемые Простак и Зануда, идя навстречу вашим неоднократным просьбам, я наконец-то нашел время и силы, чтобы встретиться с вами для беседы о математике. Прежде чем начать говорить по существу, я предлагаю рассмотреть несколько задач, которые следует отнести скорее к головоломкам, чем к классу математических задач, хотя их связь с математикой лежит на поверхности рассуждений и не требует особой проницательности, чтобы увериться в этом. Моя задача сейчас состоит в том, чтобы настроить вас на некоторую информационную волну, необходимую вам для облегчения последующего активного восприятия предлагаемого мной материала. Мой монолог вы можете в любое время прервать, чтобы задать вопрос или высказать собственное мнение по обсуждаемой теме. В результате возникшего диалога мы, быть может, скорее достигнем цели — понимания не только постановки и метода решения конкретной задачи, но и самого духа математики.

Глава 1

Головоломки и задачи

Знание только тогда знание, когда оно приобретено усилиями мысли, а не памятью.

Л. Толстой

1.1. Арифметические задачи

1.1.1. Треть равна половине?

П р о ф е с с о р. Рассмотрим следующую задачу: каковы числа, треть от которых равна половине?

П р о с т а к. Я вижу, что это — математическая задача, ответ на которую не очевиден. Сначала, признаюсь, я подумал, что это ноль, но это было бы слишком просто, чтобы походить на истину. Вы же, Профессор, недаром задали нам эту головоломку. Если для решения задачи потребуется составить какое-нибудь уравнение, то я готов увидеть его на бумаге, но сам не хотел бы тратить время на его выписывание. Я согласен, что здесь есть загвоздка, но меня больше интересует правильный ответ, чем возня с его поиском. Просто я многое забыл из школьной математики и сразу мне не вспомнить. Все же непонятно: как может быть, чтобы меньшая часть была равна большей части? Сдается мне, *что эта задача вообще не имеет решения*.

З а н у д а. Оставим эмоции. Следуя духу математики, надо попытаться составить уравнение. Обозначим искомое число через x. Тогда уравнение, соответствующее формулировке задачи, будет иметь следующий вид:

$$x/2 = x/3.$$

Очевидно, если в это уравнение вместо x подставить 0, то получим тождество $0 = 0$ и, следовательно, искомое число равно нулю, о чем и догадывался Простак. Я же не остановился на догадках, а просто вычислил решение, используя математический метод.

Профессор. Действительно, число ноль обладает тем свойством, что любая его часть всегда равна любой его части. Зная это, можно было и не составлять никаких уравнений, чтобы найти решение. Но может быть существуют и другие числа, отличные от нуля, соответствующие решению поставленной задачи?

Зануда. Из моего уравнения таких чисел найти нельзя. Похоже, что их и не существует.

Простак. В твоем уравнении, Зануда, можно сократить левую и правую части на x, получив неверное тождество $1/3 = 1/2$. Так что твой результат $x = 0$ был обнаружен только благодаря тому, что ты начал с подстановок чисел в свое уравнение вместо x, а не попытался прежде упростить само уравнение, как обычно это все делают. Иначе говоря, тебе просто повезло. Тем не менее, мы до сих пор не знаем, существуют ли отличные от нуля числа, треть от которых равна половине.

Зануда. Сокращать на x можно только в том случае, если x не равен 0. Напомню Простаку, что на ноль делить нельзя, поскольку результат деления не определен. Допустим, x не равен 0. Тогда после сокращения на x мое уравнение примет вид $1/3 = 1/2$, что, очевидно, неверно. Отсюда заключаем, что верно обратное, а именно, что $x = 0$. Более того, мы ясно видим, что других вариантов нет. Только ноль удовлетворяет условиям поставленной задачи и является единственным ее решением! Таким образом, я полностью решил вашу задачу, уважаемый Профессор.

Профессор. Я с интересом следил за вашими рассуждениями. Особенно мне понравилось последнее замечание Зануды. Вместе с тем, должен объявить вам, что вы решили не ту задачу, которую я сформулировал. Моя задача решается с помощью уравнения

$$x/3 = 1/2,$$

откуда очевидно следует единственное его решение $x = 3/2$.

Предупреждая ваши возможные недоумения, сразу скажу, что слово "половина" обозначает число 1/2 (т. е. 0,5). Обратите внимание, в задаче не говорилось о половине *чего-то*. Таким образом, требовалось найти числа, треть которых равна просто числу 1/2. Вы же решали задачу не в моей, а в следующей формулировке: каковы числа, треть от которых равна *их* половине? Впредь будьте внимательны при чтении формулировок задач. Естественный язык, на котором мы говорим, позволяет красочно, объемно и лаконично выражать мысли. Вопрос в том, насколько точно та или иная мысль выражена, и однозначно ли она понимается. Недаром говорится, что "мысль изреченная есть ложь". Рассмотренная нами задача относится к головоломкам не потому, что ее трудно решить, а потому, что ее формулировка предполагает некоторую неоднозначность интерпретации. Многие головоломки и почти все анекдоты строятся в расчете на то, что "исходные данные" будут неверно проинтерпретированы из-за невнимательности. Вспомните детскую загадку: А и Б сидели на трубе, А упало, Б пропало, что осталось на трубе? Большинство из нас легко обнаружило, что здесь речь идет о двух объектах, А и Б, а буква "и" является элементом синтаксиса предложения (союзом). Однако допустима и другая интерпретация: "и" обозначает некоторый объект, подобно буквам "А" и "Б".

1.1.2. Когда родился Иван?

Профессор. В 1975 году некто по имени Иван сказал, что ему было n лет в n^2 году. В каком году он родился?

Простак. Составим уравнение:

$$x + n = n^2,$$

где x — искомый год, в котором родился Иван.

В одном уравнении два неизвестных. Это плохо, но обнадеживает то, что величины n и x ограничены:

$$0 < n < 150;$$

$$x + n \leq 1975.$$

Зануда. Я бы ограничил n еще больше, например, так: $0 < n < 120$, хотя навряд ли Ивану могло быть 120 лет. Впрочем, есть же долгожители и постарше. А теперь надо, начиная примерно с года $x = 1855$, перебрать различные сочетания x и n, пока не получится равенство $x + n = n^2$.

Простак. Можно обойтись без перебора, пусть и ограниченного, но все же очень большого количества вариантов. Обратите внимание, что n^2 должно быть не больше 1975 и не слишком далеко от этого значения.

Из того, что $n^2 \leq 1975$, следует $n \leq \sqrt{1975} \approx 44{,}4$. Возьмем ближайшее меньшее целое число $n = 44$. Тогда $n^2 = 1936$ и $x = n^2 - n = 1936 - 44 = 1892$. Итак, Иван родился в 1892 году.

З а н у д а. А может быть, есть и другие подходящие значения x и n?

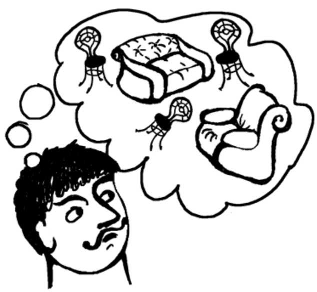
П р о с т а к. Проверим еще одно, следующее, значение $n = 43$. Для него получается $n^2 = 1849$ и $x = 1849 - 43 = 1805$. Но тогда в 1975 году Ивану было бы 170 лет. Кажется, в XX веке таких долгожителей не было, поэтому данный вариант следует отбросить.

Если мы возьмем для пробы $n = 42$, то попадем в XVIII век, поскольку $n^2 = 1754$. Итак, ясно, что условиям нашей задачи отвечает только один вариант.

П р о ф е с с о р. Замечательно. Вот видите, вы обошлись без перебора большого количества вариантов, стоило только немного поразмышлять.

1.1.3. Необычная арифметика

П р о ф е с с о р. Мы все еще со школы знаем основные арифметические операции, такие как сложение, вычитание, умножение и деление чисел. Как и сами числа, эти операции возникли из опыта. Несколько предметов можно считать неразличимыми с некоторой точки зрения и тогда можно подсчитать их общее количество. Например, столы, стулья, диваны и другие предметы обстановки образуют совокупность, называемую одним словом "мебель". Хотя эти предметы могут заметно отличаться друг от друга, мы пренебрегаем этими различиями, подсчитывая их общее количество. Если, скажем, у вас уже есть 2 яблока, а я даю вам еще одно, то вы без особых затруднений сможете вычислить, что теперь у вас стало 3 яблока. При этом вы не обращаете внимания на различия сортов и размеров яблок. Вы просто сопоставляете совокупности различных предметов число — их общее количество.



Объединяя две совокупности, мы применяем операцию сложения чисел, чтобы определить количество предметов в объединении. Выделяя из данной совокупности часть предметов, мы применяем операцию вычитания. Я хочу сказать, что таким действиям над предметами, как образование из них совокупностей, последующее слияние или разъединение этих совокупностей, соответствуют арифметические операции над числами, и наоборот. Но всегда ли это верно? Так, например: смешав 1 литр воды с 1 литром спирта, мы получим не 2, а 1,8 литра спиртового раствора; смешав два равных объема воды — один при температуре 40°, а другой при температуре 60° по шкале Цельсия, — мы не получим удвоенного объема при температуре 100°; соединив в электрической цепи параллельно два резистора с сопротивлениями R1 и R2, получим общее сопротивление R1R2 / (R1 + R2), а не R1 + R2. Итак, мы видим, что операции прибавления (соединения), выполняемой над предметами, не всегда соответствует арифметическая операция сложения соответствующих чисел.

Простак. Очевидно, что арифметические операции нельзя применять к практическим задачам столь прямолинейно. Интерпретация, например, сложения как смешение жидкостей слишком груба, а потому и не приводит к правильным результатам.

Профессор. Очень справедливое замечание! Однако давайте рассмотрим следующую задачу.

Продавец ведет учет эффективности своей торговли, вычисляя отношение количества посетителей магазина, сделавших покупки, к их общему количеству. Например, если за день из пяти посетителей только трое что-то купили, то эффективность торговли в этот день оценивается величиной 3/5. Такое отношение продавец вычисляет каждый день и желает оценить эффективность своей работы сразу за несколько дней. Допустим, в первый день эффективность равна 3/5, а во второй — 7/9. Какова суммарная эффективность торговли за два дня?

Простак. Надо просто сложить дневные эффективности:

$$\frac{3}{5}+\frac{7}{9}=\frac{3\times 9+7\times 5}{5\times 9}=\frac{62}{45}.$$

Зануда. Получается, что за два дня магазин посетили 45 человек, из которых 62 сделали покупки. Какая-то несуразность!

Простак. Я и сам вижу, что это абсурд, и готов исправиться: *суммарная эффективность — это среднее значение дневных эффективностей*, а не просто их сумма. Если требуется найти среднее значение нескольких величин, то их следует сначала сложить, а полученную сумму разделить на количество этих величин. Поэтому средняя эффективность равна

$$\frac{\frac{3}{5}+\frac{7}{9}}{2}=\frac{31}{45}\approx 0,6889.$$

Зануда. Я все равно не согласен с Простаком. За два дня магазин посетили 5 + 9 = 14 покупателей и только 3 + 7 = 10 из них что-то купили. Следовательно, средняя эффективность торговли за два дня равна совсем другой величине:

$$\frac{3+7}{5+9}=\frac{10}{14}\approx 0,7143.$$

Иначе говоря, средняя или суммарная эффективность вычисляется не путем сложения дробей, а как дробь, числитель которой равен сумме числителей, а знаменатель — сумме знаменателей этих частных дробей.

Профессор. Итак, обычная арифметическая сумма двух дробей определяется по формуле:

$$\frac{a_1}{b_1}+\frac{a_2}{b_2}=\frac{a_1\times b_2+a_2\times b_1}{b_1\times b_2}.$$

А Зануда придумал новую операцию сложения двух дробей, которую мы обозначим как $\oplus$, чтобы отличить ее от обычной операции $+$ сложения чисел:

$$\frac{a_1}{b_1}\oplus\frac{a_2}{b_2}=\frac{a_1+a_2}{b_1+b_2}.$$

Простак. Хотя метод, предложенный Занудой, и кажется обоснованным, я все же не вижу причин отказываться от своего способа. Но наши способы дают различные результаты. Так какой же из двух методов правильный?

Профессор. В поисках ответа на этот вопрос давайте рассмотрим еще одну задачу, очень похожую на задачу об эффективности торговли.

Допустим, автомобиль первую сотню километров проехал за 2 часа, а вторую сотню — за 8 часов. С какой средней скоростью автомобиль проехал всю дистанцию длиной 200 км?

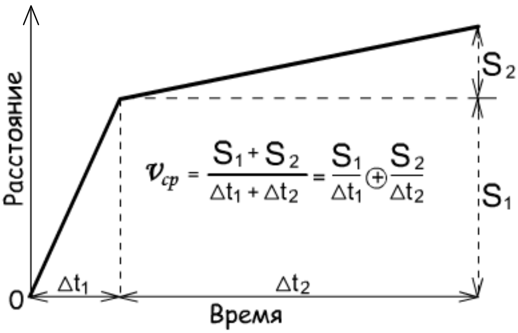
Очевидно, что автомобиль ехал неравномерно: на различных участках пути он имел различные скорости. Нас интересует *средняя скорость, которая есть скорость равномерного движения, при которой автомобиль проедет то же расстояние и за то же время, что и при неравномерном движении*. Тогда в рассматриваемом случае средняя скорость автомобиля равна 20 км/ч. Действительно, с этой скоростью он за 10 часов пройдет весь путь длиной 200 км.

Если бы мы взяли среднее арифметическое скоростей на двух участках, то получили бы неверное значение:

$$\frac{\frac{100}{2}+\frac{100}{8}}{2}=31,25.$$

С такой скоростью за 10 часов автомобиль прошел бы не 200, а 312,5 км. Правильное значение 20 км/ч средней скорости получается с помощью операции

$$\frac{100}{2} \oplus \frac{100}{8} = \frac{100+100}{2+8} = 20.$$



Теперь рассмотрим более общий случай неравномерного движения. Пусть участки пути S_1, S_2, ..., S_n автомобиль проходит за время Δt_1, Δt_2, ..., Δt_n соответственно. Очевидно, скорость на i-м участке равна $S_i / \Delta t_i$. Определим *среднюю скорость на всем пути как взвешенную сумму всех скоростей на отдельных участках*. Это означает, что суммируемые величины скоростей должны быть предварительно умножены на так называемый весовой коэффициент, равный доле общего времени, в течение которого автомобиль двигался с данной скоростью. Временна́я доля i-го участка составляет величину $\Delta t_i / T$, где $T = \Delta t_1 + \Delta t_2 + ... + \Delta t_n$ — общее время в пути. Таким образом, чем меньше по времени автомобиль двигался с данной скоростью, тем меньше вес этой скорости в сумме всех скоростей на отдельных участках, и наоборот. Очевидно, что каждый весовой коэффициент принимает значения в интервале от 0 до 1, а сумма всех таких коэффициентов равна 1.

Итак, взвешенная сумма всех скоростей на отдельных участках равна

$$\frac{S_1}{\Delta t_1} \times \frac{\Delta t_1}{T} + \frac{S_2}{\Delta t_2} \times \frac{\Delta t_2}{T} + ... + \frac{S_n}{\Delta t_n} \times \frac{\Delta t_n}{T} =$$

$$= \frac{S_1 + S_2 + ...S_n}{\Delta t_1 + \Delta t_2 + ... + \Delta t_n} = \frac{S_1}{\Delta t_1} \oplus \frac{S_2}{\Delta t_2} \oplus ... \oplus \frac{S_n}{\Delta t_n}.$$

Таким образом, мы приходим к применению операции $\oplus$ и при более сложном определении средней скорости, которое в конце концов сводится к исходному: средняя скорость равна отношению расстояния к общему времени, затраченному на его прохождение. С другой стороны, это — среднее расстояние, проходимое в единицу времени, или, как еще говорят, среднее расстояние в расчете на единицу времени.

Теперь, полагаю, вам нетрудно будет определить и среднюю эффективность торговли.

Простак. Средняя эффективность торговли аналогична средней скорости движения в пространстве. Ее можно понимать как среднее количество покупателей в расчете на одного посетителя магазина. Тогда правильной будет формула, предложенная Занудой:

$$\frac{a_1}{b_1}\oplus\frac{a_2}{b_2}\oplus\ldots\oplus\frac{a_n}{b_n}=\frac{a_1+a_2+\ldots+a_n}{b_1+b_2+\ldots+b_n}.$$

Профессор. Я согласен с Вами, Простак.

Зануда. Если я правильно понял, для решения задач о средней эффективности и скорости, а может быть и каких-то других, мы используем новую операцию $\oplus$ сложения чисел, а значит, и новую арифметику.

Профессор. Операция сложения $\oplus$ существенно отличается от обычной операции сложения чисел. Прежде всего, операция $\oplus$ была определена для дробей. Разумеется, любое целое число a можно представить в виде дроби $a/1$. В обычной арифметике выполняется равенство $a + b = a/1 + b/1$, а в новой оно нарушается: $a/1 \oplus b/1 = (a + b)/2$. Таким образом, операция $\oplus$, применяемая к двум целым числам, дает в результате их среднее арифметическое.

В обычной арифметике дроби можно сокращать, т. е. делить числитель и знаменатель на одно и то же число, отличное от нуля; при этом результат будет тем же самым числом. Заметим, что дробь в виде отношения a/b еще интерпретируют как отложенное деление числа a на число b. Например, дроби 2/3 и 4/6 выражают одно и то же число, которое в десятичной записи имеет вид 0,666… . В обычной арифметике число (2/3 + 1/4) равно числу (4/6 + 1/4), а в новой арифметике ($2/3 \oplus 1/4$) и ($4/6 \oplus 1/4$) — различные числа.

Таким образом, новая арифметика более специфична и более сложна, чем обычная. Тем не менее, вводя операции, отличные от обычных, мы можем прийти к арифметике, применимой к реальным явлениям. Производимые над числами операции выбираются так, чтобы они соответствовали некоторому интересующему нас классу явлений физического мира. Только опыт может подсказать нам, в каких случаях обычная арифметика применима к тому или иному явлению. Следовательно, мы не можем рассматривать арифметику как свод истин, применимых для описания любых физических явлений. Это заключение относится и к любым новым арифметикам. На эти вопросы в XIX веке обратил внимание выдающийся немецкий естествоиспытатель и математик Г. Гельмгольц.

1.1.4. Параллельная работа с разными производительностями

П р о ф е с с о р. Рассмотрим еще одну задачу, в которой общую работу делают несколько человек, но у каждого своя производительность, или скорость. Допустим, одна секретарша перепечатывает рукопись за 2 часа, а другая — за 3. Весь объем работы можно как-то поделить между секретаршами. Тогда за какое время они перепечатают всю рукопись, если будут работать вместе?

П р о с т а к. Я думаю, что им понадобится больше трех часов.

З а н у д а. Как же так? Одна из них помогает другой, следовательно, время совместной работы будет меньше наименьшего из двух данных времен.

П р о с т а к. Дело в том, что секретарши наверняка будут болтать друг с другом, из-за чего их производительность и снизится почти до нуля.

Шучу, конечно. Вопрос, я думаю, в том, чтобы правильно разделить между секретаршами объем работы. Очевидно, делить пополам не стоит, т. к. более производительная секретарша закончит свою часть работы раньше и будет "простаивать". Следовательно, если ей дать объем работы побольше, то удастся сократить общее время перепечатки всей рукописи. Так что более производительной секретарше надо дать бóльшую часть рукописи. Если обозначить объем работы (например, количество листов рукописи) для более производительной секретарши через x_1, а для менее производительной — через x_2, то отношение этих величин должно быть обратно пропорционально значениям времени, затрачиваемого ими на работу:

$$\frac{x_1}{x_2} = \frac{t_2}{t_1} = \frac{3}{2}.$$

З а н у д а. Хотя гипотеза о том, что должна выполняться данная пропорция, и разумна, она все же остается гипотезой, причем недостаточно обоснованной. А мне бы не хотелось начинать с каких-то, пусть даже правдоподобных, предположений.

Что мы имеем в качестве исходных данных? А то, что некий объем работы (обозначим его через x) первая секретарша выполняет со скоростью x/t_1, а вторая — со скоростью x/t_2, где $t_1 = 2$, а $t_2 = 3$.

Пусть первой секретарше поручили выполнить объем работы x_1, а второй — x_2. Чтобы найти время, затрачиваемое на выполнение работы, надо объем этой работы разделить на скорость (производительность). Далее, когда совместная работа двух секретарш завершится, обе они затратят на это одинаковое время. Таким образом, справедливо следующее равенство:

$$\frac{x_1}{x/t_1} = \frac{x_2}{x/t_2}.$$

П р о с т а к. Нетрудно заметить, что полученное Занудой равенство легко приводится к моему, сформулированному как гипотеза. Так что я был прав!

З а н у д а. Пусть так. Но теперь это равенство вполне обоснованно. Кроме того, оно содержит нужную мне переменную x, которая обозначает теперь весь объем работы.

П р о с т а к. Однако ее можно сократить.

З а н у д а. Можно, но не нужно. Дело в том, что я введу еще одно очевидное равенство:

$$x_1 + x_2 = x.$$

П р о с т а к. Теперь у нас два равенства, но три неизвестных: x_1, x_2 и x. Ничего хорошего в этом я не вижу.

З а н у д а. Позвольте мне продолжить. Итак, у нас есть следующие два равенства, легко получаемые из приведенных ранее с помощью элементарных преобразований:

$$\frac{x_1 t_1}{x} = \frac{x_2 t_2}{x};$$

$$\frac{x_1}{x} + \frac{x_2}{x} = 1.$$

Но что нам требуется найти по условию задачи? А лишь значение левой или правой части первого из равенств, поскольку они выражают время, затраченное на совместную работу. Напомню, что обе секретарши, работая вместе, затрачивают на работу одинаковое время. Так что из двух уравнений требуется сначала найти, например, x_1/x, а затем умножить это значение на t_1.

В результате этих манипуляций получим время, за которое будет напечатана рукопись параллельно работающими секретаршами:

$$\frac{t_1 t_2}{t_1 + t_2} = \frac{2 \cdot 3}{2 + 3} = 1{,}2.$$

П р о ф е с с о р. Что ж, решение верно, хотя и путь к нему не был коротким. Попробуем сократить его, используя другую модель ситуации.

Представьте себе, что секретарши печатают рукописи, так сказать, в противоположных направлениях. Одна из них начинает с первой страницы, а другая — с последней. Если они печатают на компьютере, то такое вполне осуществимо. Разумеется, производительности секретарш не одинаковы, работа в различных направлениях идет с разной скоростью. Но с какой скоростью продвигается работа в целом?

П р о с т а к. Я, кажется, понял, что надо делать! Пусть x — общий объем работы (все расстояние, которое требуется пройти). Секретарши работают со скоростями x/t_1 и x/t_2. С точки зрения первой секретарши, работа идет с суммарной скоростью $x/t_1 + x/t_2$. С такой же скоростью продвигается работа и с точки зрения второй секретарши. Именно с суммарной скоростью растет пачка уже напечатанных страниц, а это и есть скорость совместной работы. В какой-то момент все страницы будут напечатаны. Это произойдет спустя время t_0 после начала совместной работы. Скорость совместной работы двух секретарш есть x/t_0.

Составим уравнение

$$\frac{x}{t_0} = \frac{x}{t_1} + \frac{x}{t_2},$$

откуда, сокращая попутно левую и правую части на x, получаем время совместной работы

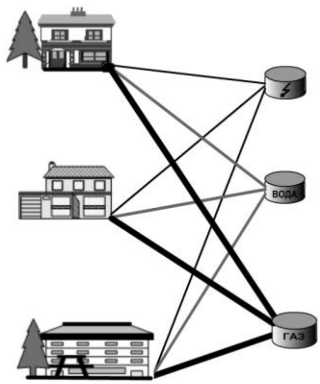
$$t_0 = \frac{t_1 t_2}{t_1 + t_2}.$$

П р о ф е с с о р. Вот видите, немного изменили взгляд на задачу, выбрали другую модель — и решение оказалось более простым!

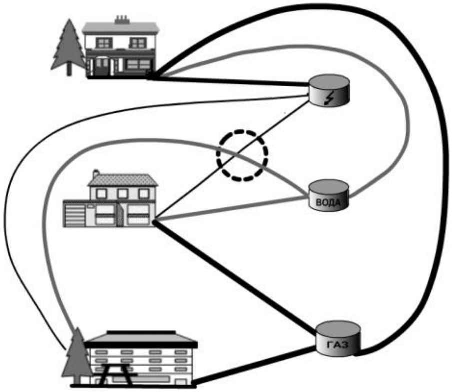
1.2. Геометрическая задача: обустройство дач

П р о ф е с с о р. Теперь рассмотрим задачу, связанную с геометрией. Пусть в дачном поселке слева от дороги находятся три дома, а справа — три колодца с коллекторами (водопроводным, электрическим и газовым). Требуется каждый дом соединить с каждым коллектором, но так, чтобы траншеи с линиями соединений из соображений безопасности не пересекались. Как это сделать?

П р о с т а к. Придется повозиться с чертежами. Похоже, соединительные линии не могут быть прямыми, поскольку при первом рассмотрении прямые линии пересекаются. Если же линии сделать извилистыми, то для всех линий, кроме двух, задачу решить легко, применяя обводки. Но вот развести эти оставшиеся две линии не получается. Мне кажется, что эту задачу вообще нельзя решить.



П р о ф е с с о р. Некоторые задачи могут не иметь решения, но в таких случаях это необходимо доказать. Иначе вы можете только утверждать, что вам пока не удалось найти решения из-за недостатка времени, знаний, ума или еще чего-нибудь.



З а н у д а. Да, я слышал, что графы (множество точек, соединенных линиями) могут иметь или не иметь так называемое свойство планарности. У планарного графа, расположенного в одной плоскости, соединительные линии не пересекаются. Планарные графы применяются при разработке печатных плат для электронных устройств. Действительно, нельзя допустить, чтобы неизолированные проводники электрического тока пересекались. Так что надо доказать возможность построения соответствующего нашей задаче графа, обладающего свойством планарности. Или, наоборот, доказать, что такой граф не возможен. Однако для этого требуются специальные знания, которых у нас нет. Разумеется, я бы мог разыскать подходящий учебник, но там наверняка много комбинаторики, чего я не люблю. Как бы то ни было, данная задача мало похожа на головоломку. Для ее решения больше нужны конкретные знания, чем смекалка и здравый смысл, которых обычно вполне достаточно для разрешения головоломок.

Профессор. Прекрасно, что вы оба решили сначала создать математическую модель ситуации, описанной в условии задачи. Сам собой напрашивается подход: надо заменить дома точками одного сорта, а коллекторы — точками другого сорта. Тогда, как будто, задача сводится к поиску требуемых соединений разносортных точек. Первые эксперименты привели Простака к неудаче. Но, может, он просто провел мало экспериментов? А может быть решения просто не существует, что, однако, нельзя доказать сколь угодно длительными неудачными экспериментами. Недостаток ваших специальных знаний из теории графов я компенсирую тем, что подскажу: для принятой модели планарный граф не существует. Но что из этого следует?

Простак. Ваш вопрос, Профессор, по меньшей мере странен. Вы же только что сами на него ответили: задача не имеет решения.

Зануда. Как-то не верится, что с виду незамысловатая инженерная задача, наверняка уже встречавшаяся проектировщикам дачных поселков, неразрешима. Что же получается, мы должны отказать заказчику в выполнении проекта?

Профессор. Не спешите с окончательным ответом. Что мы установили? Только то, что задача неразрешима в рамках принятой модели. Интуиция подсказывает нам, что исходная практическая задача должна иметь решение — ведь разнородные коммуникации как-то разводят, причем даже в более сложных случаях. Например, часто требуется добавить еще телефонные и телевизионные кабели, а также выделенные линии связи с Интернетом, не говоря уже о канализации.

Простак. Я кажется понял, в чем дело. Надо прокладывать линии на разных уровнях по глубине или высоте. Странно, что я раньше об этом не догадался.

Зануда. Но это уже не математическая задача.

Профессор. Уважаемый Простак сразу же решил задачу практически, но, надо признать, после некоторых околоматематических мытарств. А Вы, уважаемый Зануда, как будто потеряли интерес к задаче, едва почувствовав, что ее решение находится вне математики. Давайте еще раз зафиксируем то, что мы получили в попытках решить задачу с помощью математики. Мы приняли некоторую математическую модель и убедились, что в ее рамках задача неразрешима. А какова наша модель? Кроме условного изображения домов точками одного сорта, коллекторов — точками другого, а траншей между ними — линиями, мы еще неявно предположили, что результат решения должен располагаться в одной плоскости. Если вполне естественно допустить, что дома и коллекторы находятся в одной плоскости, то разве в условии задачи требовалось, чтобы траншеи проходили в этой же плоскости? Как и в случае первой задачи о числах, вы невнимательно отнеслись к формулировке условия. Если бы явно требовалась планарность графа соединений домов с коллекторами, то ваша модель была бы вполне подходящей к данной практической задаче, и ее решение выглядело бы несколько необычно для обычных людей: решение не существует. В подобных случаях, как правило, заказчика просят как-то изменить условия задачи, чтобы она была, по крайней мере, разрешима. Теперь, я надеюсь, вы согласитесь, что рассмотренная задача все же относится к головоломкам, поскольку требует для своего решения только внимания и немного смекалки.

1.3. Алгоритмические задачи

1.3.1. О песочных часах

Профессор. Предположим, у нас есть двое песочных часов, позволяющих отмерить соответственно 7 и 11 минут. Возможно ли с их помощью отмерить временной интервал в 15 минут? Если да, то как?

Простак. К сожалению, мы не можем отмерить с помощью одних часов половину или какую-нибудь другую часть того интервала, на который они рассчитаны. Поэтому мы можем отмерить только 7, 11 и 7 + 11 = 18 минут.

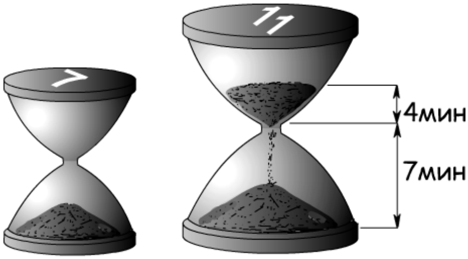
Зануда. Не только. Кроме того, мы можем отмерить еще и интервал, равный 11 – 7 = 4 минутам. Для этого достаточно запустить двое часов одновременно, и как только весь песок в 7-минутных часах высыпется, в 11-минутных его еще останется на 4 минуты. В этот момент вы можете положить часы набок, чтобы прекратить утечку песка, а затем, когда понадобится, снова поставить их в вертикальное положение. При этом остаток песка на 4 минуты, разумеется, должен находиться в верхней полуколбе часов.

Простак. Но этот 4-минутный остаток песка находится в 11-минутных часах. Следовательно, мы не сможем сначала отмерить 11 минут и вслед за этим еще 4 минуты, чтобы в сумме получилось 15.

Зануда. Значит необходимо "перенести" этот 4-минутный остаток в 7-минутные часы.

Простак. А как это сделать?

Зануда. Просто запустить 7-минутные часы заново, когда в 11-минутных часах в верхней колбе останется песка на 4 минуты. Как только весь песок в 11-минутных часах окажется в нижней полуколбе, в 7-минутных часах внизу будет песка ровно на 4 минуты. Можете положить эти часы набок, чтобы сохранить их состояние до нужного момента. Теперь все готово к тому, чтобы отмерить 15 минут. Для этого следует сначала запустить 11-минутные часы, а после, когда весь их песок окажется внизу, привести в действие 7-минутные часы с остатком песка в верхней полуколбе на 4 минуты. Вот и все!



Профессор. Отлично! Теперь я знаю, что вам не составит особого труда отмерить 10 минут. Попробуйте на досуге разработать соответствующий алгоритм.

1.3.2. Поиск фальшивой монеты: одна из восьми

Профессор. Пусть среди восьми с виду одинаковых монет находится одна фальшивая, которая легче настоящих. Как выявить фальшивую монету, используя рычажные весы не более двух раз?


Простак. Очевидно, что взвешивая пару монет (на каждой чаше весов по одной монете), мы можем ограничиться всего лишь одним взвешиванием, но только если нам повезет: фальшивая монета окажется в выбранной для взвешивания паре. В худшем случае нам потребуется четыре взвешивания. Я где-то слышал о решении похожей задачи, которая, кажется, называется "поиск льва в пустыне". Там вся пустыня разбивалась на две одинаковых области и по рыку льва определялось, в какой из областей он находится. Эта область, откуда разносился рык, делилась пополам и снова определялась подобласть, где обретается лев. Данная процедура продолжалась до тех пор, пока область разбиения не оказывалась настолько малой, чтобы мы могли с уверенностью сказать: "Вот лев." Не кажется ли вам, что между поисками льва и фальшивой монетой есть аналогия?


Зануда. Если воспользоваться аналогией Простака, то требуется разбить множество из восьми монет на два подмножества из четырех монет и сравнить их веса. Выбрать самое легкое подмножество (именно там находится фальшивая монета) и разбить его пополам, чтобы затем сравнить эти половины на весах. Аналогичным образом следует поступать до тех пор, пока на чашах весов не окажется по одной монете. На этом последнем этапе взвешивание и выявит фальшивую, более легкую, монету. Нетрудно подсчитать, что данная стратегия потребует трех взвешиваний, а не двух, как указано в задаче. Так что либо "поиск льва в пустыне" — плохой метод, либо поставленная задача неразрешима.


Простак. Я вспомнил о методе поиска льва потому, что он был самым быстрым. Поэтому я склонен считать, что либо двумя взвешиваниями нельзя определить одну фальшивую монету из восьми, либо аналогия между задачами о льве и взвешиваниях недостаточно полна, чтобы лучший метод для первой задачи оказался лучшим и для второй.


Профессор. Я вижу, что рассмотрение предыдущих задач не прошло для вас даром.


Зануда. Я не могу отказаться от использования, хотя бы частичного, аналогии с задачей о льве. Может быть, попробовать делить множество монет не на два, а на большее количество подмножеств? Что будет, если разбить множество монет на три подмножества?

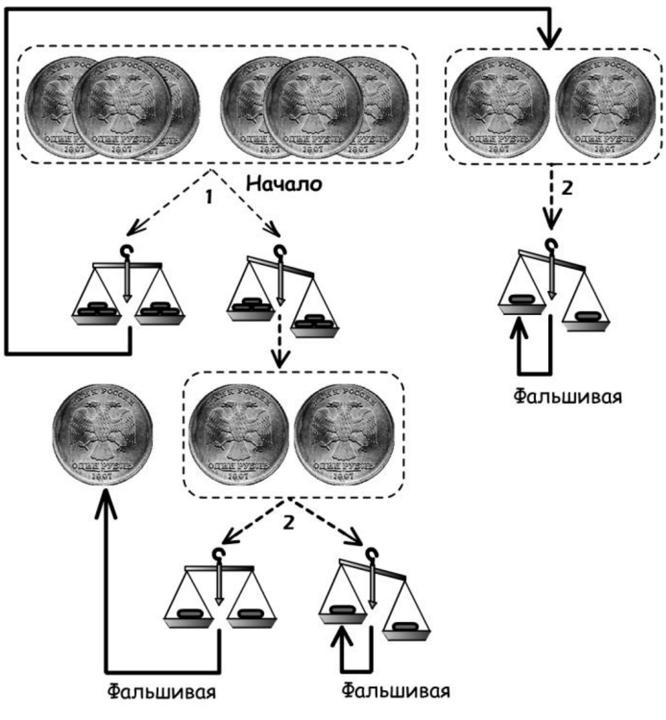

Простак. Конечно, можно попробовать. Но сравнивать по весу мы можем только два подмножества. Более того, сравнивать надо только подмножества, содержащие одинаковое количество монет. При этом сравниваемые множества могут оказаться и равными по весу, в отличие от случая деления пополам. Наконец, разбить множество из восьми монет на три подмножества можно несколькими способами. Какой из них выбрать?

Зануда. Способов разбиения восьми монет на три подмножества не так уж и много. Не перебирая все возможные варианты, следует выбрать тот из них, при котором количества монет в подмножествах наиболее близки друг к другу. Очевидно, что первые два подмножества должны содержать по три монеты, а последнее — оставшиеся две (3 + 3 + 2 = 8).

Простак. Хорошо, но что взвешивать на первом этапе? Фальшивая монета может оказаться в любом из трех множеств, а результат взвешивания может ничего не дать: сравниваемые множества могут оказаться одинаковыми по весу.

Зануда. Если сравниваемые множества одинаковы по весу, то это означает, что ни в одном из них нет фальшивой монеты и, следовательно, она находится в третьем множестве. А разве это не положительный результат, не информация для планирования последующих действий?

Простак. Действительно, я был опрометчив в своих оценках. Тем не менее, с чего начать? Следует ли сначала сравнить на весах две монеты из третьего множества, или же сравнить первые два множества, содержащие по три монеты?



З а н у д а. Для меня очевидно, что начать надо с двух множеств, содержащих по три монеты. Впрочем, ты можешь поступить иначе, чтобы убедиться в возможности варианта, когда двух взвешиваний тебе не хватит. Итак, сравним два множества, по три монеты в каждом. Тогда возможны два варианта:

1. Оба множества весят одинаково, т. е. в них нет фальшивой монеты. В данном случае переходим к третьему множеству из двух монет, сравнение которых на весах и укажет фальшивую. В итоге мы нашли фальшивую монету посредством двух взвешиваний.
2. Трехмонетные множества различаются по весу. Тогда выбираем наиболее легкое из них и взвешиваем любые две монеты. Это (второе по счету) взвешивание сразу укажет фальшивую монету, независимо от того, равны ли по весу сравниваемые монеты. Действительно, если сравниваемые монеты имеют разный вес, то самая легкая из них — фальшивая, а если ни одна из них не перевешивает другую, то фальшивой является третья монета, которую мы не клали на чашу весов. Таким образом, и в этом варианте мы обошлись всего двумя взвешиваниями.

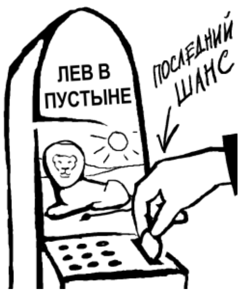
1.3.3. Поиск фальшивой монеты: одна из девяти

П р о ф е с с о р. Браво, Зануда! Вы блестяще модифицировали метод или алгоритм, применить который предлагал Простак. Теперь я несколько усложню задачу. Пусть количество внешне одинаковых монет равно 9. Можно ли в этом случае с помощью двух взвешиваний определить более легкую фальшивую монету?

П р о с т а к. Для решения задачи с девятью монетами следует попробовать применить тот же алгоритм, что и в случае с восемью монетами. Разбиваем множество всех монет на три количественно равных подмножества, т. е. по три монеты в каждом. Далее сравниваем на весах любые два из них. Если вес одинаков, то фальшивая монета в третьем множестве, которое мы не клали на весы. А что с ним делать, мы уже знаем из предыдущей версии задачи. Чтобы выявить в нем фальшивку, требуется только одно взвешивание. Всего же необходимо два взвешивания. Если же при первом взвешивании выяснилось, что сравниваемые трехмонетные подмножества имеют разный вес, то переходим к более легкому из них. Как нам уже известно, в этом случае потребуется только одно взвешивание, а в общем итоге — два. Так что алгоритм работает и применительно к задаче о девяти монетах.

З а н у д а. Пока Простак решал задачу по моему алгоритму, я уже успел выяснить, что в случае 10 монет понадобятся минимум 3 взвешивания.

П р о ф е с с о р. Вы, сдается мне, попробовали применить тот же самый алгоритм? Но, быть может, для 10 монет найдется другой, более эффективный алгоритм, который позволит обойтись не более, чем двумя взвешиваниями?



З а н у д а. Признаюсь, я применил тот же алгоритм, что и для задач о 8 и 9 монетах. А о возможности существования другого, более эффективного алгоритма я не подумал.

П р о с т а к. Ничего не могу сказать о существовании более эффективного алгоритма для случая 10 и более монет, но в случае 9 и менее монет, я убежден, алгоритма лучшего, чем только что рассмотренный, не найти.

1.3.4. Поиск фальшивой монеты: попытка обобщения

П р о ф е с с о р. Обратите внимание, что вопрос о самом эффективном алгоритме становится особенно важным, если сформулировать задачу следующим образом: каково наименьшее количество взвешиваний, с помощью которых можно выявить фальшивую монету среди N монет? В такой формулировке задачи, даже при конкретном значении N, решение должно быть иным. Заметьте, что, вообще говоря, не требуется разрабатывать алгоритм группировки монет и их взвешиваний. Но если вы все же пошли по такому пути, то необходимо еще доказать, что не существует более эффективного алгоритма.

З а н у д а. Я, пожалуй берусь доказать, что для достаточно больших количеств всех монет N не обойтись одним взвешиванием.

П р о с т а к. Но это почти и так очевидно, безо всяких выкладок.

З а н у д а. Но если я это докажу, то смогу убедить вас, что мой алгоритм для случая $N = 8$ или $N = 9$ самый эффективный.

П р о ф е с с о р. Попробуйте, уважаемый Зануда, это интересно.

З а н у д а. В самом деле, взвешивание состоит в том, что на чаши весов кладутся два подмножества монет, причем монет должно быть поровну, поскольку в противном случае результат взвешивания ничего не скажет нам о наличии или отсутствии фальшивой монеты. Итак, сравниваемых подмножеств должно быть минимум два. Допустим, эти подмножества содержат по m монет. Разумеется, $m \neq 0$ и $2m \leq N$. Очевидно, возможны два варианта:

1. $2m = N$. Нетрудно заметить, что N является четным числом. Если $N = 2$, то $m = 1$ и, следовательно, достаточно одного взвешивания. Если $N > 2$ (например, 4, 6, 8, …), то $m > 1$ и результат взвешивания приведет к необходимости рассматривать подмножество из нескольких монет, следовательно, понадобится, по крайней мере, еще одно взвешивание. Таким образом, для четных $N > 2$ одного взвешивания недостаточно.
2. $2m < N$. Нетрудно заметить, что в данном случае $N > 2$ (например, 3, 4, 5, …). Если $N = 3$ или $N = 4$, то $m = 1$ и, следовательно, одного взвешивания будет достаточно. Если $N > 4$, то либо $m > 1$, либо $N - 2m > 1$. В первом случае при взвешивании не исключено, что фальшивая монета окажется среди $m > 1$ монет, и тогда потребуется хотя бы одно дополнительное взвешивание. Во втором случае не исключено, что при первом взвешивании вес подмножеств окажется одинаковым и придется рассматривать подмножество из $N - 2m > 1$ монет, а тогда без дополнительного взвешивания не обойтись. Итак, при $N > 4$ одного взвешивания недостаточно.

Объединяя результаты рассмотрения двух частных случаев, но исчерпывающих все пространство возможностей, заключаем, что при общем количестве монет $N \geq 4$ требуется не меньше двух взвешиваний. Поскольку ранее я изобрел алгоритм, обеспечивающий для $N = 8$ и $N = 9$ только что полученное минимальное, равное двум, количество взвешиваний, он наилучший.

П р о с т а к. Признаюсь, я не ожидал от тебя, Зануда, таких аналитических способностей. Тем не менее, теперь мне ясно, что ты не изобрел свой алгоритм в результате какого-то наития, подобно поэту или композитору, а пришел к нему чисто логическим путем, т. е. посредством скучных, рутинных операций. Впрочем, это не умаляет твоих заслуг.

З а н у д а. Честно говоря, на меня повлияло твое, Простак, упоминание о поиске льва в пустыне методом деления ее участков пополам. Я тогда подумал: а что если делить не пополам, а на три или больше частей? Но теперь я и сам вижу, что мой алгоритм не зависит от этой идеи. Она была моей "музой" — скорее стимулировала мое воображение, чем указывала путь. Алгоритм решения был почти полностью предопределен условием задачи: взвешивайте группы монет на рычажных весах. Оставалось неочевидным: что сравнивать с помощью весов — две произвольно выбранные монеты или два произвольных подмножества монет? То, что количества монет на разных чашах весов должны быть одинаковы, и так понятно. А сколько монет следует выбирать для одного взвешивания, я вычислил, рассуждая логически.

Простак. Мне кажется, что подобным образом можно доказать, что для одних значений N потребуется не менее трех, а для других — не менее четырех взвешиваний. Хорошо бы найти формулу, определяющую минимальное количество взвешиваний для заданного N.

Зануда. Не ручаюсь за общий случай, но, как мне кажется, я бы смог доказать, что при $N \geq 10$ потребуется не менее трех взвешиваний.

Профессор. Попробуйте это сделать на досуге, сейчас не будем тратить на это время.

Простак. Действительно, мы и так слишком задержались на этой задаче, не имеющей важного практического значения. Но меня сейчас больше волнует вопрос: может быть, это не единственный алгоритм, обеспечивающий обнаружение фальшивой монеты с помощью минимального количества взвешиваний?

Профессор. Да, таких алгоритмов может быть несколько. Например, при $N = 6$ можно разбить множество всех монет на 2 подмножества по 3 монеты либо на 3 подмножества по 2 монеты. В любом случае, потребуется 2 взвешивания. А поскольку для $N = 6$ по "теореме" Зануды это минимальная оценка, то оба алгоритма считаются наилучшими. Замечу, что здесь алгоритмы различаются только способом исходного разбиения множества монет на подмножества.

1.3.5. Обмен валют: постановка задачи

Профессор. Завершая тему фальшивой монеты, замечу, что есть много других, но аналогичных задач. Например, задача обмена валют, так или иначе решаемая с помощью программного обеспечения компьютера в обменном пункте. Валюты представлены различными наборами купюр (бумажных денег, банкнот). Например, есть купюры достоинством 5, 10, 50, 100, 500, 1000 и 5000 рублей, но нет купюр достоинством 3 и 4 рубля. Наборы купюр других валют могут отличаться. Кроме того, в обменном пункте в данный момент может не оказаться купюр какого-то определенного достоинства, а количество имеющихся очень ограниченно. Клиент собирается обменять некоторую денежную сумму, например, в рублях, на другую валюту по курсу. Очевидно, что возможна ситуация, когда предъявленная сумма не может быть полностью переведена в другую валюту из-за ограниченного набора купюр в обменном пункте. Чтобы все же выполнить операцию обмена, клиент либо должен согласиться получить сдачу в рублях, либо добавить некоторую сумму в рублях к предъявленной для обмена. Естественно желать, чтобы эта дополнительная сумма сдачи или добавки была как можно меньше. Можно потребовать, чтобы эта сумма была минимальной. Именно в такой постановке задача может заинтересовать математиков.

П р о с т а к. Я смутно ощущаю, что задача об обмене валют связана с задачей о взвешиваниях, но пока не могу понять, каким именно образом.

З а н у д а. Очевидно, что сначала надо с помощью коэффициента, выражающего переводной курс валют, вычислить сумму, которую обменный пункт должен, в первом приближении, отдать клиенту.

П р о с т а к. За вычетом комиссионных, разумеется.

З а н у д а. Да, но сейчас это не столь важно. Далее следует подобрать купюры так, чтобы сложившаяся сумма оказалась как можно ближе к вычисленной по курсу.

П р о с т а к. Лучше сказать, чтобы разница между вычисленной суммой и суммой номиналов купюр была минимальной. Поскольку, вообще говоря, одну и ту же сумму можно получить с помощью различных комбинаций купюр, придется рассматривать множество вариантов.

З а н у д а. А вот и аналогия с задачей о взвешиваниях! На одну чашу рычажных весов мы кладем вычисленную сумму денег, а на другую — купюры требуемой валюты. Нетрудно заметить, что купюры здесь играют роль гирь. Требуется так подобрать гири разных номиналов и их количества, чтобы наклон рычага весов был минимальным.

П р о ф е с с о р. Обнаруженная вами аналогия хороша. Однако должен заметить, что ни один обменный пункт не решает задачу обмена в такой постановке (я имею в виду требование минимальности). Дело в том, что, как показали специальные математические исследования, длина пути поиска такого решения слишком велика.

Простак. В обменные пункты следует просто поставить более производительные компьютеры. В конце концов, то, что нам недоступ-но сейчас, станет обычным завтра.

Профессор. К сожалению, загвоздка имеет принципиальный, а не технологический характер. Дело в том, что рассматриваемая задача с требованием минимальности расхождения вычисленной и составленной из купюр сумм имеет, как говорят специалисты, экспоненциальную сложность. Это значит, что сложность (например, количество шагов или время вычисления) экспоненциально зависит от объема исходных данных. Так, если n — объем исходных данных, то сложность экспоненциально сложных задач определяется как e^n, т. е. очень быстро возрастает с увеличением n. Например, если $n = 50$, то e^n выражает число, превосходящее количество атомов в видимой части Вселенной. Только перечисление их с помощью самого быстродействующего компьютера потребует практически неприемлемо долгого времени. Относительно нашей задачи доказано, что она разрешима и, более того, найден алгоритм ее решения. Вместе с тем, доказано, что эта задача имеет экспоненциальную сложность, что делает соответствующий алгоритм в практически интересных случаях практически (но не теоретически) нереализуемым. При очень малых объемах данных он вполне реализуем и способен достаточно быстро давать результаты, но это может быть неинтересно с практической точки зрения. Таким образом, существуют теоретически разрешимые задачи с практически нереализуемыми алгоритмами.

Простак. Ваши слова звучат как парадокс, Профессор. Но ведь формулировка задачи для пункта обмена валют выглядит вполне разумно, не говоря уж о ее практической важности. Быть может, ее следует немного подкорректировать, сохраняя смысл?

Профессор. Неплохая мысль.

Зануда. Мне кажется, что нужно убрать требование минимальности как слишком жесткое.

Простак. Но что же тогда останется? Ты хочешь, чтобы пункт обмена решал свою задачу как попало? Например, клиент хочет обменять 1000 рублей на доллары, а пункт предлагает ему 1 доллар, дополняя остальное рублями в виде сдачи.

Профессор. Убрать требование минимальности — хорошая идея. Мы можем построить алгоритм решения задачи обмена валют, который не обеспечивает минимальную разницу между вычисленной суммой и составленной из купюр во всех случаях, а лишь стремится уменьшить эту разницу. Возможно, иногда ему удастся достичь минимального результата, но гарантировать это во всех случаях нельзя. Мы надеемся, что такое поведение алгоритма обеспечит если не наилучший, то хороший с практической точки зрения результат. Замечу, что во многих случаях так и поступают.

З а н у д а. Как это формальный алгоритм может стремиться к чему-либо? Он же не живой. Любой алгоритм просто что-то делает, а что-то — нет, он либо решает задачу, либо — нет, без всяких эмоций.

1.3.6. Поиск кратчайшего пути

П р о ф е с с о р. Я поясню свою мысль об алгоритме, который стремится к наилучшему результату, на следующем примере. Пусть из пункта *A* в пункт *B* можно добраться различными путями по дорогам с перекрестками, на которых установлены стрелочные указатели расстояний до следующего перекрестка. Как, начиная движение из пункта *A*, дойти до пункта *B* кратчайшим путем?

Самый простой по своей идее алгоритм решения этой задачи основывается на методе прямого перебора всех возможных путей достижения пункта *B* из пункта *A* и выбора из них наименьшего по своей протяженности. Однако количество всех путей может оказаться слишком огромным, чтобы их перебор был практически осуществим.

Однако можно поступить и иначе. На каждом перекрестке будем выбирать то направление движения, в котором следующий перекресток ближе. Если случится забрести в тупик, то можно вернуться на предыдущий перекресток и выбрать другое направление. Очевидно, что на каждом перекрестке мы разумно выбираем направление, используя всю доступную на данном этапе движения информацию. Если бы выбор направления движения на каждом перекрестке производился случайно, то посторонний наблюдатель не мог бы сказать, что мы стремимся сократить маршрут. Поэтому мы говорим, что наш алгоритм стремится к сокращению маршрута, хотя и не гарантирует, что весь пройденный путь будет кратчайшим. Выбор наилучшего локального решения не гарантирует, что глобальное решение, составленное из локальных, будет наилучшим. Действительно, сэкономив в начале пути, мы можем лишить себя возможности таких последующих решений, которые вкупе с предыдущими оказались бы лучше.

Итак, рассмотренный алгоритм лишь стремится сократить весь маршрут, но не гарантирует его минимальной протяженности. Зато он не требует полного перебора всех возможных путей, а, следовательно, работает быстрее. Впрочем, я должен заметить, что существуют эффективные алгоритмы поиска кратчайших маршрутов, основанные не на полном переборе. Наиболее известным из них является алгоритм Дейкстры, но сейчас мы не будем его рассматривать.

Простак. Мне кажется, что простой алгоритм, который лишь стремится сократить маршрут, не гарантируя при этом наилучшего результата, можно усовершенствовать. Так, прежде чем окончательно выбрать направление движения на данном перекрестке, можно сначала проверить все направления из него. Иначе говоря, следует сбегать на все смежные перекрестки, чтобы по указателям на них узнать расстояния до следующих перекрестков. Получив эту информацию и вернувшись на исходный перекресток, можно принять решение о движении с учетом протяженности не одного, а двух переходов. Разумеется, эта стратегия потребует больше времени, но зато конечный результат будет лучше.

Зануда. Аналогично, можно заглянуть не только на непосредственно смежные, но и на последующие перекрестки. Если продолжить эту аналогию, то мы придем в конце концов к полному перебору всех маршрутов и выбору самого короткого из них.

Профессор. Вы оба мыслите в верном направлении. Но давайте оставим эту задачу. Подведем итоги обсуждения так называемых алгоритмических задач. Обратите внимание: их головоломный характер обусловлен тем, что решение требует изобретения плана действий, т. е. алгоритма. Однако попытки решить задачу в более общей постановке, или даже только некоторая модификация условия задачи, обнаруживают необходимость математического анализа. Возможно, вам будет небезынтересно узнать, что рассмотренные нами задачи так или иначе были использованы в теории оптимального кодирования и теории информации, в так называемой транспортной задаче оптимального планирования перевозок и ряде других областей. Некоторые алгоритмы, придуманные для решения частных задач, оказывались применимыми и для других.

1.4. Логические задачи

Профессор. Рассмотренные ранее задачи требовали некоторых логических рассуждений. И, как мы убедились, до сих пор логика только помогала нам. Теперь же я предлагаю вам рассмотреть исключительно логические задачи.

1.4.1. Почему трезвенник был уволен за пьянство?

П р о ф е с с о р. В некоторой судоходной компании действовало строгое правило: на судах во время рейса под страхом увольнения было запрещено употреблять алкоголь. Но капитан одного из судов этой компании нередко пренебрегал данным правилом, и это сходило ему с рук, поскольку никто из членов экипажа не доносил на него. И вот однажды старший помощник капитана, слывший среди членов экипажа трезвенником, решил "подсидеть" своего шефа, воспользовавшись его слабостью к алкоголю. Будучи дежурным и застав капитана изрядно выпившим, он записал в вахтенном журнале: "Сегодня капитан был мертвецки пьян". На следующий день, заступив на вахту, капитан просмотрел журнал и в бешенстве набросился на своего помощника: "Зачем ты это сделал? О, я знаю, каналья, ты давно метишь на мое место!" — "По долгу службы я обязан заносить в журнал все события на судне, и то, что я записал, — чистая правда. Какие могут быть ко мне претензии?" "Ладно, за это я тебе отомщу", — подумал капитан, но промолчал. Заканчивая свое дежурство, он записал в вахтенном журнале: "Сегодня старший помощник был трезв". По возвращении домой оба, капитан и его старший помощник, были списаны на берег по одной и той же статье — за пьянство. Хотя комментировать анекдот — верный способ убить его, все же ответьте, почему уволили обоих?

П р о с т а к. Капитан, понятно, был уволен, поскольку в вахтенном журнале прямо отмечалось нарушение им закона компании. А необычная, согласитесь, запись капитана о состоянии своего помощника наводит на мысль, что последний был пьяным каждый день плавания, кроме одного, когда была сделана эта запись.

З а н у д а. Как бы то ни было, с формальной точки зрения, получается, что две противоположных записи в журнале, а значит, и два противоположных события, привели к одному и тому же результату. Капитан был уволен за то, что однажды был пьян, а его помощник — за то, что однажды был трезв. Однако за трезвость по правилам компании не увольняют.

П р о ф е с с о р. Вы только что верно подметили, Зануда, что помощник был уволен за то, что *однажды* был трезв. В этом-то "однажды" все и дело. В вахтенный журнал не заносят банальные сведения, укладывающиеся в обычный распорядок на судне, — наоборот, там отмечают все чрезвычайное. Не записывать же туда, например, что сегодня взошло солнце! Впрочем, данное сообщение банально в средних широтах и весьма информативно для приполярных областей. Если действие происходит в высоких широтах, то данная запись извещает нас о том, что полярная ночь закончилась, в противном случае она не несет никакой информации. Аналогичная ситуация и в анекдоте о капитане с помощником. Руководство компании рассуждало, по-видимому, так. Коль скоро в вахтенном журнале однажды появилась банальная запись, то соответствующее событие (помощник был трезв) тоже имело место всего один раз и, следовательно, данная запись уже не банальна, что вполне соответствует общей презумпции небанальности всех записей в журнале.

Большинство, если не все, анекдотов построены таким образом, чтобы противопоставить презумпцию слушателей и буквальный смысл описываемой ситуации.

1.4.2. Кому достанется киевский престол?

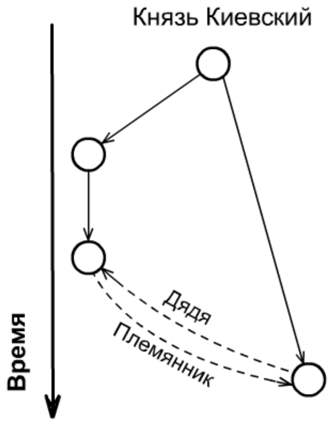
Профессор. Давным-давно, на Руси главный (т. е. киевский) престол наследовался князьями Рюриковичами по старшинству. Впрочем, это же правило применялось и к наследованию уделов. Правило было простым и как будто естественным: наследником являлся старший по происхождению и по возрасту. Старшинство по происхождению определялось так: отец, дед, прадед и т. д. старше соответственно сына, внука, правнука и т. д. Иначе говоря, родитель старше своих детей, детей своих детей и т. д. Старшинство по возрасту определялось естественным образом: кто раньше по времени родился, тот и старше. Братья, очевидно, были равны по происхождению, но не равны по возрасту (даже близнецы рождаются друг за другом, а не одновременно). Впрочем, появление близнецов-первенцев у царствующей особы обычно вызывало изрядное беспокойство, поскольку возникал повод оспорить первенство рождения, а это было чревато гражданской войной. Если верить Александру Дюма-старшему, французская королева Анна Австрийская родила двойню. Один из мальчиков стал впоследствии королем Людовиком XIV, а второй провел жизнь в тюрьме под железной маской. И все это ради того, чтобы не дать повод к открытой войне между противоборствующими политическими партиями, каждая из которых отстаивала бы своего короля. Разумеется, партии воевали бы за свои корпоративные интересы, а не за истинного короля и торжество закона. Король в данном случае использовался бы как знамя, объединяющее бойцов. Ради гарантии гражданского мира закон престолонаследования должен был быть сформулирован достаточно точно.

Зачем правило престолонаследования в древней Руси формулировалось с учетом двух порядков старшинства?

Простак. Ответ очевиден. Предположим, у князя несколько сыновей. Все они равны с точки зрения старшинства по происхождению, поскольку являются братьями. Наследником может быть только один из них. Поэтому для исключения возможных коллизий добавляется еще одно условие: наследником признается старший из братьев по возрасту.

Профессор. Именно так и думали наши предки, формулируя закон престолонаследования. Впрочем, если верить Н. Карамзину, древние русичи переняли этот закон у варягов-викингов вместе с их вождями, приглашенными на княжение. Но достаточно ли учета двух порядков старшинства, чтобы избежать неопределенности?

Допустим, у князя два сына, а у старшего из них есть свой сын. Таким образом, в княжеском семействе есть дядя и племянник. Если князь умирает, то ему наследует старший из сыновей, поскольку он остается старшим и по происхождению, и по возрасту. Но что произойдет, если умрет и вступивший в наследство старший сын князя, — кому следует передать престол?



Простак. А кто из оставшихся в живых старше по происхождению? Ясно — младший сын самого первого князя. Следовательно, он и должен занять престол.

Зануда. Но старший по происхождению может оказаться младшим по возрасту. Второй сын мог родиться у старого князя значительно позже того, как его первый сын обзавелся собственными детьми. Таким образом, вполне возможно, чтобы племянник оказался старше своего дяди.

Если и сейчас так бывает, то в старину, когда дети рождались чуть ли не до глубокой старости своих отцов, — такое и подавно могло быть.


Простак. Дядя — старший по происхождению и лишь потом — младший по возрасту. Поэтому право на престол за ним.


Зануда. С какой стати ты решил, что сначала следует учитывать старшинство по происхождению и только затем — по возрасту?

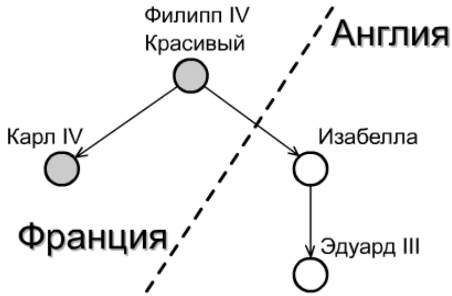

Простак. Насколько я помню, правило наследования формулировалось так: наследником является старший по происхождению и по возрасту. Здесь сначала упоминается происхождение, а затем — возраст.


Зануда. Сложные условия, состоящие из двух и более простых, нельзя произнести или записать иначе, как последовательно друг за другом. В данном случае ничто, кроме порядка упоминания в одной фразе, не указывает на то, что одно условие должно быть главным, а второе дополнительным, принимаемым в расчет при прочих равных обстоятельствах. Наконец, истинность или ложность предложения, состоящего из двух других, соединенных союзом "и", определяется только истинностью или ложностью последних и не зависит от их порядка в составном предложении. Так что в данном случае никто из претендентов не обладает свойствами, удовлетворяющими закону о престолонаследовании. Следовательно, ситуация тупиковая или, на языке шахмат, патовая.

Профессор. Примерно так и оспаривали дядья с племянниками свои права на престол, что по обыкновению заканчивалось войнами. Если закон не помогает, побеждает сила.

1.4.3. Почему началась столетняя война?


Профессор. Возможно, коллизий можно было бы избежать, если бы закон о престолонаследовании формулировался примерно так: наследником престола является старший по перворождению в старшей линии рода, а переход престола к другой линии происходит только в случае, когда старшая линия рода прервалась (последний ее представитель умер, не оставив детей). Европейские страны пошли по этому пути, но и им не удалось избежать конфликтов при дележе наследства. Поводом для начала Столетней войны (1337—1453 гг.) послужили притязания английского короля Эдуарда III на французскую корону после смерти французского короля Карла IV.



После смерти Филиппа IV Красивого, знаменитого разгромом могущественного ордена Тамплиеров, французский трон достался его старшему сыну, который спустя некоторое время умер, не оставив сыновей. Тогда корона перешла ко второму, среднему сыну Филиппа, но тот умер, также не имея детей мужского пола, а потому королем стал младший из трех братьев — Карл IV.

Дочь Филиппа IV Изабелла, сестра Карла IV, вышла замуж за английского короля Эдуарда II и от этого брака у них родился сын — будущий король Англии Эдуард III. Таким образом, Эдуард III приходился внуком Филиппу IV.

Но вот Карл IV умирает, не оставив сыновей, как и все его братья. Его род по мужской линии пресекается. В данной ситуации по обычаям того времени судьбу французской короны должны были решить Генеральные штаты (парламент). Рассматривались два претендента: двоюродный брат Филипп Валуа и племянник Эдуард III Английский, имевший во Франции огромные ленные владения. Это были владения не Англии как государства, а ее короля как частного лица (феодала). При этом само владение основывалось на феодальном отношении вассал-сеньор (хозяйствующих субъектов). Так что английский король как владелец половины земель Франции являлся вассалом французского короля, но при этом сама Англия и юридически, и политически не зависела от Франции. Если бы Эдуард III добился французской короны, то его вассальная зависимость исчезла бы, он стал бы сеньором сам себе, а на карте Европы появилось бы новое государство, объединяющее Англию и Францию. Понятно, что Эдуард III решил достичь этой цели во что бы то ни стало.

Французская партия не хотела этого допустить. Однако ей требовалось юридическое обоснование отказа Эдуарду в его слишком далеко идущих притязаниях. Следует заметить, что в то время использовалось так называемое обыкновенное право, т. е. право, основанное на обычаях предков. Для обоснования чего-либо достаточно было сослаться на тот или иной старинный обычай или найти исторический прецедент, который был в свое время признан соответствующим обычаем. Авторитетным сводом обычаев тогда была так называемая салическая, или франконская правда. Так вот, французы нашли в этой правде обычай передавать корону по наследству только отпрыскам мужского пола. Поскольку линия Эдуарда III являлась женской, идущей от Филиппа IV через его дочь Изабеллу, французы сочли притязания Эдуарда на французский престол беспочвенными. Англичане оспорили заявление французов, ссылаясь на другие прецеденты.

Юридическая база спора была нечеткой, договориться не удалось и началась длительная и кровопролитная война, в горниле которой, как у французов, так и у англичан, впервые выкристаллизовалось национальное самосознание. Война началась как дело сеньоров-феодалов, а закончилась как дело наций. Английская корона потеряла почти все свои владения во Франции. Шанс объединения французов и англичан в единой европейской супердержаве был упущен.

Простак. Да, неясность в законах всегда оставляет лазейки для их нарушения, а также таит в себе потенциальные конфликты. Но причем тут математика?

Профессор. Не кажется ли вам, что закон следовало бы предварительно исследовать как математическую задачу? В данном случае закон выражался в некотором правиле переходов между вершинами древовидного графа (генеалогического дерева). Как и большинство других законов, он должен был работать, когда все идет хорошо, естественным порядком. Однако законы, как и математические задачи, особенно всякого рода алгоритмы, очень важно исследовать в так называемых крайних, исключительных случаях, когда возможны противоречия и тупики. Вспомните хотя бы школьную задачу о нахождении корней алгебраического уравнения второй степени вида $ax^2 + bx + c = 0$.

Простак. Корни это числа, подставляя которые в уравнение вместо переменной x, мы получаем тождество $0 = 0$. Есть формула, выражающая значения корней x_1, x_2 через коэффициенты a, b, c данного уравнения:

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Нужно просто подставить эти коэффициенты в формулу, чтобы получить искомый результат. Это — типичная вычислительная задача и никакой премудрости в ней нет.

Зануда. Однако нетрудно заметить, что данная формула не работает, если коэффициент $a = 0$, поскольку всем известно, что на 0 делить нельзя. Кроме того, если окажется, что $b^2 - 4ac < 0$, то корни будут не действительными числами, а комплексными. Более того, я думаю, не исключена ситуация, когда корней у данного уравнения вообще может не быть. В этом случае формула, очевидно, тоже не работает.

П р о ф е с с о р. Вот именно. Возможны различные ситуации, определяемые конкретными значениями коэффициентов a, b, c, но формула из справочника применима не ко всем из них. Она применима только тогда, когда корни для данного уравнения в самом деле существуют.

Я обращаю ваше внимание на то, что мы часто относимся к заданному нам вопросу или поставленной задаче так, как будто ответ или решение уже существуют, а нам лишь требуется только найти их. В предположении существования решения мы интересуемся общим методом (формулой) его поиска или вычисления. Иначе говоря, нам нужен результат. При этом мы нередко забываем, что исходное предположение о существовании решения или принципиальной возможности получения результата само нуждается в обосновании. Придумав общий алгоритм решения задачи, мы должны задуматься о границах его применимости, а значит — вернуться к формулировке той задачи, которая явилась мотивом для разработки этого алгоритма. Быть может, наш алгоритм решает лишь частный случай первоначально поставленной задачи, или, наоборот, — позволяет решить более общий ее вариант. Может статься, что наш алгоритм решает совсем другую задачу.

То, о чем я сейчас говорю, наиболее актуально проявляется в программировании, т. е. в создании компьютерных программ. Представьте себе, что поставлена задача о вычислении корней любого квадратного уравнения. Это означает, что пользователь программы может задать любые численные значения коэффициентов a, b, c. Написать программный код на каком-нибудь языке программирования для вычислений по указанной ранее формуле не представляет особого труда. Кстати, начинающие программисты, как правило, начинают именно с перевода на язык программирования взятой из справочника подходящей формулы. Тем не менее, основная задача программиста состоит в том, чтобы написать код, анализирующий исходные данные (значения a, b, c), выявляя особые частные случаи, в том числе и случай применимости названной формулы. Часто этот анализ занимает больше половины всего кода программы. И только если данная формула окажется применимой, программа должна выполнить соответствующие вычисления.

Итак, тщательный анализ крайних случаев, исследование границ применимости является непременным условием использования общих правил, алгоритмов и законов, наконец.

1.4.4. Казнить или помиловать?

П р о ф е с с о р. Давным-давно один губернатор решил пропускать в свои владения только правдивых людей, для чего на пропускном пункте границы области своей юрисдикции установил виселицу и издал закон: "Всякий входящий должен объявить, куда и зачем он идет; кто скажет правду, того пропускать, а кто солжет, того тут же и повесить". Каждый входящий должен присягнуть говорить правду, только правду и ничего, кроме правды. Сначала все шло своим чередом: прохожие рассказывали о цели своего прибытия, пограничники и судебные как-то оценивали их правдивость и действовали сообразно с законом. Но однажды некто заявил, что он прибыл сюда, чтобы его вздернули на этой виселице, а других целей у него нет. Каким должен быть приговор судьи, который обязан в точности исполнить закон?


Простак. Я думаю, что несмотря на необычность заявления прохожего, суд должен удовлетворить его просьбу. Впрочем, я не вполне ручаюсь за правильность такого решения. Ведь должен же здесь быть какой-то подвох.


Зануда. Прежде всего, следует разобраться, говорил прохожий правду или лгал.


Простак. А как это выяснить? Чужая душа — потемки, не средневековые пытки же применять. Быть может, он действительно искал смерти и решил уйти из жизни с посторонней помощью. Какой смысл ему лгать? Ведь если об этом дознаются судьи, то непременно казнят его во имя закона.


Зануда. По-твоему, Простак, выходит, что прохожий не лгал и, значит, должен быть пропущен. Однако сначала ты предположил обратное, т. е. что следует удовлетворить его желание и казнить. Прежде чем выносить вердикт, надо было хорошенько подумать.


Простак. Да, я поторопился. В самом деле, если прохожий действительно желал быть повешенным, то для достижения своей цели ему следовало бы солгать, т. е. сказать, что он следует в губернию за чем-то, но только не для того, чтобы его казнили тут же на границе. Иначе говоря, прохожий должен солгать, чтобы достичь своей цели. Что ж, это вполне жизненная ситуация. Однако мотив лжи только подтверждает, что в сущности прохожий был правдив. Если он действительно хотел быть повешенным, то ему надлежало лгать. Если же, наоборот, он в самом деле хотел перейти границу, то ему лучше было сказать что-нибудь другое, но не шутить на предмет виселицы.


Профессор. Я согласен, что шутки со смертью плохи. Любая шутка предполагает в своей основе некоторую ложь. Мы смеемся над шуткой, лишь разоблачив спрятанную в ней ложь и поняв, почему она была так похожа на правду. Тем не менее, из рассуждений Зануды вытекает, что чудаковатый прохожий должен был либо говорить ложь, внешне похожую

на правду, либо говорить с виду правду, оставаясь внутренне лжецом. Впрочем, я, кажется, совсем вас запутал.

1 2 3

П р о с т а к. Думаю, проблема в том, чтобы достоверно установить, правду или ложь говорил прохожий. Каков критерий истинности высказывания?

П р о ф е с с о р. Ваши рассуждения, на первый взгляд, кажутся разумными, т. е. не лишенными некоторой логики. Однако они основываются на реконструкции возможных мыслей и желаний прохожего. Чувствуется, что вы не желаете, чтобы пострадал невинный (т. е. правдивый), и в то же время вы не хотите упустить нарушителя закона. Что ж, это вполне жизненная ситуация. Вы повели следствие даже так, чтобы докопаться до возможных мотивов подследственного: чего он хотел в действительности. В конце концов, ваш вердикт будет выглядеть примерно так: поскольку прохожий скорее говорил правду, чем ложь, то его следует пропустить через границу. Не исключаю, что ваше решение будет иметь прямо противоположный характер. Тем не менее, в условии задачи требуется, чтобы закон был соблюден в точности. А вы даже не попытались проверить, возможно ли это. Другими словами, в задаче требуется принять такое решение, чтобы восторжествовал закон. Поверьте, что решение данной задачи не нуждается в выяснении, что же именно сказал прохожий — правду или ложь. Вопрос надо ставить так: прохожий должен быть либо пропущен через границу, либо казнен; какое из двух решений соответствует закону. В математике часто начинают анализ по схеме: допустим, что первая альтернатива верна, тогда посмотрим, что из этого получится; аналогично допускается истинность второй альтернативы и изучаются ее возможные следствия. Попробуйте порассуждать таким способом.

З а н у д а. Допустим, что прохожий должен быть пропущен через границу. Но если позволить ему идти дальше, то тогда получается, что он нарушил присягу и, следовательно, должен быть казнен. Пройдя мимо виселицы целым и невредимым, прохожий, как выясняется, солгал относительно своей цели. Следовательно, по закону его необходимо казнить. Я рад, что для принятия такого решения мне не пришлось копаться в глубинных намерениях чудака-прохожего. Вместо этого просто посмотрите, что он будет делать. Если пойдет дальше, значит — лгун и по закону должен быть повешен. Очевидно противоречие с первоначальным вердиктом суда!

Простак. Постойте! Но как же его казнить? Ведь он так и заявил, что только для того и пришел, чтобы его повесили, т. е. говорил правду. Значит, по закону его следует оставить в живых. Первоначальный вердикт суда оказывается правильным!

Зануда. Мы снова пришли в исходную точку, с которой я начал: если его пропустить, то скоро мы обнаружим, что его надо повесить ради торжества закона, но в то же время по тому же закону его надо оставить в живых. Впрочем, я не рассмотрел еще одну альтернативу. Допустим с самого начала, что прохожий должен быть казнен…

Простак. Да и без дальнейших рассуждений я уже вижу, что результат будет аналогичным. Как только будет решено казнить, так тут же выяснится, что надо помиловать, но помиловать нельзя. Короче говоря, возникает порочный круг.

Профессор. Действительно, какое бы решение ни было принято, закон будет нарушен. Адвокат и обвинитель могут спорить в суде сколь угодно долго. Если судья не проявит слабость и останется верным закону, он не примет никакого решения, поскольку его просто не существует. Возможность возникновения порочных кругов, противоречий коренится в особенностях нашего мышления, нашей логики. Задачи, связанные с логическими противоречиями, называют парадоксами, антиномиями и апориями. Они были обнаружены еще древними греками и в настоящее время их известно очень много, хотя часть из них являются различными версиями одного и того же прототипа.

Замечу, друзья, что данная задача была описана четыре века тому назад Сервантесом в его знаменитом романе "Дон Кихот". Эта история там произошла с Санчо Пансой, когда он временно был назначен губернатором и подвергся испытанию, чтобы быть осмеянным.

1.4.5. Исполнение приговора — сюрприз для осужденного

Профессор. А вот вам другая логическая задача, впервые обсуждавшаяся в серьезном философском английском журнале "Mind" в середине XX века.

Один ужасный разбойник, на счету которого было немало кровавых злодеяний, был наконец-то пойман и осужден на смертную казнь. Приговор суда был таков: ты будешь повешен в один из семи дней на следующей неделе, причем ты не будешь знать заранее, в какой именно день это произойдет; тебе сообщат об этом в день казни утром, до полудня. Понятно, что судья хотел заставить злодея мучиться неопределенностью, ведь говорят, что ожидание смерти хуже, чем сама смерть. Однако адвокат разбойника шепнул ему наедине: приговор не может быть исполнен, ибо противоречив. В самом деле, в последний день недели казнь не может состояться, так как, будучи живым в предпоследний день после полудня, разбойник легко догадается, что казнь назначена на следующий, единственный день, когда еще может свершиться правосудие. Таким образом, разбойник знал бы о дне казни раньше, чем ему об этом сообщили, а исполнение казни в этот день противоречит приговору. Аналогично, казнь не может состояться и в шестой, предпоследний, день: поскольку осужденный дожил до полудня пятого дня, то, стало быть, он уже знает, что казнить его могут только в шестой день (в седьмой день это не может произойти, как мы уже выяснили). По той же причине казнь не может состояться в пятый и другие дни недели. Разбойник успокоился благодаря логически безукоризненным доводам адвоката. Однако в один из дней недели, на которой по решению суда должна была состояться казнь, рано утром в камеру разбойника явился тюремщик вместе с палачом и заявил, что казнь назначена на сегодня после полудня. "Но это невозможно, поскольку противоречит приговору!" — удивленно вскричал разбойник. Казнь, тем не менее, была произведена в полном соответствии с приговором. В чем тут загвоздка?

П р о с т а к. Скорее всего, адвокат в чем-то ошибся.

З а н у д а. Нет, адвокат рассуждал вполне логично.

П р о с т а к. Если казнь состоится, то это будет противоречить приговору в той его формулировке, в которой он был объявлен судьей. По приговору разбойник не должен заранее знать, когда его казнят; он должен узнать об этом в день казни утром от тюремщика. Адвокат же убедил разбойника, что тот может узнать о дне казни накануне без объявления ему об этом, а чисто логическим путем. При этом не важно, на какой именно день назначена экзекуция. Таким образом, разбойник получил предварительное знание о дне казни и именно поэтому казнь не может состояться. Перечитайте приговор. В нем требуется, чтобы осужденный узнавал о дне казни не иначе как посредством объявления.

З а н у д а. Вроде бы все верно, но предварительное знание, добытое с помощью адвоката логическим путем, есть знание не о конкретном дне казни, а о том, что приговор противоречив. Впрочем, можно считать, что это знание и о дне, хотя оно и выглядит несколько странно: день казни логически невозможен.

П р о с т а к. Быть может, следует пересмотреть дело в суде высшей инстанции и сформулировать приговор так, чтобы не возникало парадокса?

П р о ф е с с о р. В данном случае есть аргументы в пользу решения о том, что исполнение приговора суда первой инстанции не противоречит его формулировке. Действительно, логические рассуждения своего адвоката разбойник принял в качестве бесспорного доказательства невозможности своей казни и успокоился. Теперь он знал, что ни в один из семи дней казнь не может произойти. Когда утром, еще до полудня, в камеру зашли тюремщик и палач, узник заранее знал, что казнь не может произойти и в этот день. Когда же ему объявили о повешении сегодня после полудня, он был удивлен (что, кстати, свидетельствует об отсутствии предварительных знаний об услышанном). Таким образом, все формальности были соблюдены и оставалось только дождаться полудня, чтобы уже через минуту, если хотите, привести приговор в исполнение.

П р о с т а к. Все же меня не оставляет смутное ощущение, что адвокат просто усыпил бдительность своего подопечного казуистическими рассуждениями, чем и воспользовались власти.

П р о ф е с с о р. Сейчас я вам докажу, что накануне совершения казни разбойник действительно не знал, что завтра его повесят. Предположим, что казнь была назначена на какой-то из семи дней. Обозначим номер этого дня через x. Что же именно знал или мог знать разбойник накануне, т. е. в $(x - 1)$-й день после полудня? Он мог рассуждать следующим образом. Поскольку, по логике адвоката, в день $(x + 1)$ и последующие дни казнь невозможна, то остается день x — единственный последний день, когда еще возможно исполнение правосудия. Этот несложный вывод есть не что иное, как знание разбойника, хотя и не все доступное ему знание. Однако он тут же обнаруживает, что полученное утверждение противоречит приговору. Чтобы снять противоречие, разбойник заключает, что в день x казнь невозможна. Данный вывод составляет еще одну часть полученного разбойником знания. В итоге накануне дня x полное знание разбойника относительно следующего дня можно сформулировать так: казнь возможна и казнь невозможна. Очевидно противоречие. Согласно логике, противоречивое высказывание может быть только ложным, а ложное высказывание не выражает никакого знания. Таким образом, получается, что казнить разбойника можно в любой день, поскольку накануне он не будет располагать знанием о дне казни.

Чтобы вам было понятнее, я рассмотрю частный случай, связанный с последним днем недели, когда приговор еще, вроде бы, можно исполнить. Итак, тюремщик приходит утром в камеру осужденного и спрашивает: "Когда состоится казнь?" — "Сегодня, но я об этом знал еще вчера, а потому казнь сегодня не состоится", — отвечает разбойник. "Ты не знал и не знаешь дня казни, поскольку не можешь точно указать день, а потому ты будешь повешен сегодня после полудня", — сказал тюремщик.

Простак. Ничего не понимаю, оказывается казнить разбойника можно в любой, даже последний день недели, но тогда получается, что в рассуждениях адвоката таится ошибка. В чем же она?

Профессор. Давайте разберемся, на чем основывались рассуждения адвоката и к чему они привели. Адвокат исходил из того, что приговор не противоречив, а результатом его логических рассуждений явилось заключение, что приговор все же противоречив и именно поэтому не может быть выполнен. Однако из противоречия можно вывести что угодно. Противоречие могут использовать обе стороны судебного процесса, адвокат и прокурор, в пользу своей победы. Вопрос лишь в том, кто раньше сдаст свои позиции или, лучше сказать, чью сторону примет судья в затянувшемся споре сторон. Мое же рассуждение о дне казни x также опирается на предположение, что приговор не противоречив, но в отличие от адвоката я пришел к противоречию не приговора, а знания разбойника о дне казни, и поэтому приговор может быть исполнен.

Наличие противоречия всегда связано с неким порочным кругом, по которому наши рассуждения могут бродить сколь угодно долго. Вот вам пример, в котором порочный круг хорошо виден. Итак, тюремщик приходит в последний день утром в камеру осужденного и спрашивает: "Когда состоится казнь?" — "Сегодня, но я об этом знал еще вчера", — отвечает разбойник. Заметьте, он ничего не говорит о невозможности казни, избегая противоречивости своих знаний. "Я знаю формулировку приговора, и поскольку твоя осведомленность не позволяет его исполнить, то казнь не состоится. Но раз так, то твое знание о дне казни ложно или, иначе говоря, ты не знал о дне казни. Следовательно, казнь произойдет сегодня. Но тогда ты знал об этом и, следовательно, казнь не состоится…", — так бесконечно долго стал рассуждать тюремщик, ибо не мог разрешить противоречие. Это противоречие между знанием разбойника и приговором. Но информация о возможном знании или незнании разбойника включена, если вы еще помните, в саму формулировку приговора. Так что это противоречие, как будто, самого приговора.

Зануда. Из последнего вашего рассуждения, профессор, я понял, что приговор все-таки не может быть приведен в исполнение хотя бы потому, что тюремщик будет бесконечно долго рассуждать, как ему быть. Однако причину этого я усматриваю в том, что разбойник просто не сообщил тюремщику полного своего знания. Ведь благодаря адвокату он знал, что приговор противоречив, а это знание больше, чем то, которое содержится в фразе "Казнь будет сегодня, но я об этом знал еще вчера".

Профессор. Действительно, адвокат сообщил разбойнику, используя логические рассуждения, нечто большее, а именно, что приговор не может быть исполнен. Но для исполнения приговора требовалось лишь, чтобы разбойник не знал о дне казни накануне. Что же знал разбойник о дне казни? А только то, что ни в один из семи дней недели казнь не может состояться. Но если он скажет об этом в ответ на вопрос тюремщика "Когда состоится казнь?", последний должен будет констатировать, что разбойнику неизвестен день казни ни сегодня, ни накануне, ни в любой другой день. Так что исполнить приговор можно.

Простак. А все же у разбойника был шанс спастись! Для этого ему нужно было, не мудрствуя лукаво, просто написать в записке произвольный день предполагаемой казни и хранить ее на всякий случай. С вероятностью 1/7, небольшой, но все же не нулевой, он совпал бы с назначенным днем. Представьте, в пятницу утром, например, в тюремную камеру входит палач и объявляет, что сегодня он исполнит приговор, а осужденный достает записку, в которой написано примерно следующее: "Сегодня, в понедельник, я уже знаю, что казнь назначена на пятницу".

Профессор. Пожалуй, так разбойник мог бы миновать расплаты за свои злодеяния. Я хотел бы обратить ваше внимание не на то, можно ли в действительности исполнить приговор, а на трудности, которые могут встретиться при решении логических задач.

Отмечу, что впервые эта задача, как новый парадокс, была опубликована Майклом Скривеном. Затем ее разбирали весьма маститые математики, такие как Д. О'Коннор, Т. О'Бейрн, У. Куайн и др.

Глава 2

О логике

Трепет охватывает при мысли, какого труда требуют поиски истины, даже малой ее части.

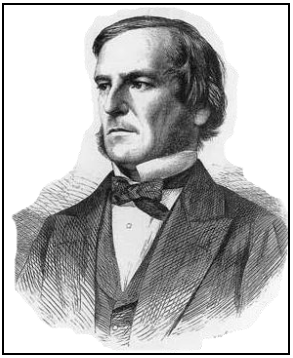
Стендаль

П р о ф е с с о р. Теперь давайте рассмотрим некоторые основные понятия и приемы, обычно используемые математиками. Мы увидим, что в жизни их применяют не только математики, но и почти все, кто считают себя мыслящими разумно. Речь пойдет о законах логики, определениях объектов изучения и доказательствах утверждений о свойствах этих объектов.

2.1. Высказывания

П р о ф е с с о р. Вот несколько типичных примеров высказываний: "Москва — столица России", "Саша любит Машу", "2 меньше, чем 3", "На Марсе есть жизнь", "Если завтра будет хорошая погода, то мы поедем на пикник". Сейчас мы абстрагируемся (т. е. отвлечемся) от содержания, смысла высказывания, приняв во внимание лишь то, что оно может быть истинным или ложным.

Такой подход к высказываниям составляет основу математической двузначной логики. Его в середине XIX века разработал англичанин Джорж Буль, поэтому математическую двузначную логику часто называют в его честь булевской (англ. boolean). Существуют и другие логики, но в этой беседе мы их не будем рассматривать.



Джордж Буль (1815—1864)

З а н у д а. Если я Вас правильно понял, речь пойдет о высказываниях или утверждениях, выраженных на естественном языке, относительно которых нас интересует только истинны они или ложны. Некоторые утверждения можно считать заведомо истинными, а другие — заведомо ложными. Однако в большинстве случаев мы не можем приписать истинность или ложность утверждению вне зависимости от конкретного контекста или положения вещей. Так, можно придумать утверждение, которое в один момент истинно, а в другой — ложно. Например, утверждение "Идет дождь" истинно только тогда, когда действительно идет дождь, и ложно в ясную погоду; в одном месте дождь идет, а в другом нет. А вот еще один пример. Пусть сейчас светит солнце и температура воздуха равна 27 градусам. Тогда утверждение "Сейчас хорошая погода" для кого-то будет истинным, а для кого-то, любящего прохладу, — ложным.

П р о ф е с с о р. Сейчас нам не важно, каким образом оценивается истинность утверждений. Более того, нам не важно, истинным или ложным является само конкретное высказывание в данный момент и в данном месте или по чьему мнению. Не важно также, осмысленно ли это высказывание. Имеет значение только то, что оценить истинность высказывания принципиально возможно. Разумеется, чтобы оценить истинность высказывания на практике, иногда требуются дополнительные раздумья и исследования, но это — вопрос применимости теоретической модели. Модели для того и придумываются, чтобы освободиться от чрезмерного количества нюансов, затемняющих суть. Правда, создавая модель, мы можем и чрезмерно абстрагироваться или, как еще говорят, "выплеснуть вместе с водой и ребенка".

Итак, утверждение или высказывание может быть истинным или ложным. Очевидно, что из любого высказывания можно получить другое высказывание путем отрицания первого. Точнее говоря, если A — некоторое высказывание, то "не A" — высказывание, полученное отрицанием высказывания A. Очень часто вместо "не A" пишут $\overline{A}$, обозначая чертой сверху оператор отрицания. Вы можете читать выражение $\overline{A}$ как "неверно, что A".

Обозначим через *И* всегда и для всех заведомо истинное высказывание, а через *Л* — заведомо ложное высказывание. В любой теории сразу же после ввода в обиход каких-то сущностей (в нашем случае — высказываний) пытаются сравнивать их между собой. Прежде всего, необходимо определить равенство между высказываниями. Они могут быть различными по форме выражения, но в то же время равными с некоторой определенной точки зрения. Например, высказывания "$2 \times 2 = 4$" и "$2 + 2 = 4$" внешне (по форме) различны, но с точки зрения арифметики равносильны в том смысле, что они оба истинны. Аналогично, высказывания "$1 = 0$" и "Петр I — первый в мире космонавт" равносильны в том смысле, что оба ложны.

Условимся использовать знак равенства между высказываниями следующим образом: выражение $A = B$ означает, что высказывания A и B имеют одинаковое значение истинности, не важно какое именно. Они одновременно либо оба истинны, либо оба ложны. В частности, выражение $A = И$ означает, что высказывание A истинно. Очевидно, что в этом случае A не может быть ложным, поскольку *И* истинно по определению. Аналогично, выражение $A = Л$ означает, что высказывание A ложно. Подчеркну еще раз, что равенство $A = B$ может выполняться даже для высказываний, имеющих совершенно различный смысл. Так например, равенство "2 меньше 3" = "Эта фраза содержит пять слов" совершенно различных по смыслу высказываний верно, ибо участвующие в нем высказывания оба истинны. Равенство "$2 + 2 = 5$" = "Эта фраза содержит шесть слов" также верно, поскольку оба участвующих в нем высказывания ложны. Чтобы не запутаться в трактовке знака равенства между высказываниями, выражение $A = B$ читают как "A равносильно B". Замечу попутно, что оператор отрицания можно понимать не только как приставку *не*, но и как "неверно", "не есть" и т. п.

З а н у д а. Нетрудно заметить, что справедливы следующие равенства: $И = И$ и $Л = Л$, а также $A = A$ для любого A. Это тривиально.

П р о ф е с с о р. Далее, если $A = И$, то $\overline{A} = Л$, и, наоборот, если $A = Л$, то $\overline{A} = И$. Иначе говоря, отрицание высказывания A приводит к другому высказыванию $\overline{A}$, значение истинности которого противоположно значению истинности исходного высказывания. Это положение соответствует интуиции мере большинства, если не всех, людей. В таком случае верно, что для любого высказывания A верно отношение $\overline{\overline{A}} = A$. Здесь с помощью двух черточек обозначена операция двойного отрицания. Если дважды отрицать высказывание, то получится это же высказывание. Например, высказывание "Неверно, что 4 не равно 2×2" равносильно высказыванию "4 равно 2×2".

Очевидно, что из высказываний A и B можно получить более сложные, составные высказывания, соединив их союзом (логической связкой) И или ИЛИ. Обозначим эти союзы символами & и $\vee$ соответственно. Тогда выражения $A \,\&\, B$ и $A \vee B$ являются высказываниями. Замечу попутно, что операции & и $\vee$ соединения двух высказываний в специальной литературе еще называют конъюнкцией и дизъюнкцией, или логическим умножением и логическим сложением соответственно. Значения истинности составных высказываний определяются в зависимости от значений истинности элементарных высказываний, обозначенных как A и B.

Точнее, высказывание $A \,\&\, B$ истинно тогда и только тогда, когда оба высказывания A и B истинны; в противном случае (т. е. когда ложно хотя бы одно из них — A или B) высказывание $A \,\&\, B$ ложно. Высказывание $A \vee B$ истинно тогда и только тогда, когда истинно хотя бы одно из высказываний — A или B; в противном случае (т. е. когда ложны оба высказывания — A и B) высказывание $A \vee B$ ложно. Возможно, сказанное будет понятнее, если использовать так называемую таблицу истинности:

A	***B***	***A & B***	***A ∨ B***
И	*И*	*И*	*И*
И	*Л*	*Л*	*И*
Л	*И*	*Л*	*И*
Л	*Л*	*Л*	*Л*

Если угодно, то таблицу истинности для $A \,\&\, B$ и $A \vee B$ можно считать определениями операций & и $\vee$. Очевидно, что данные операции коммутативны, т. е. если высказывания в выражении с этими операциями поменять местами, то получим равносильное высказывание.

Обратите внимание на следующие важные частные случаи:

И & И = И *И ∨ И = И*

Л & Л = Л *Л ∨ Л = Л*

И & Л = Л *И ∨ Л = И*

И & А = А *И ∨ А = И*

Л & А = Л *Л ∨ А = А*

Если значения истинности обозначить числами 1 и 0 для значений "Истина" и "Ложь" соответственно, операции & и $\vee$ можно понимать как обычное арифметическое умножение и как специальным образом определенное сложение (через *x* здесь обозначено произвольное значение — 1 или 0):

$1 \times 1 = 1$	$1 + 1 = 1$
$0 \times 0 = 0$	$0 + 0 = 0$
$1 \times 0 = 0$	$1 + 0 = 1$
$1 \times x = x$	$1 + x = 1$
$0 \times x = 0$	$0 + x = x$

Теперь мы готовы выразить законы противоречия и исключения третьего:

- $A\ \&\ \overline{A} = Л$ — закон противоречия;
- $A \vee \overline{A} = И$ — закон исключения третьего.

Эти формулы легко получить с помощью таблицы истинности, если считать, что $B = \overline{A}$, и, кроме того, учесть, что значение истинности $\overline{A}$ противоположно значению истинности *A*.

2.1.1. Закон противоречия

Профессор. В соответствии с законом противоречия ни одно высказывание и его отрицание не могут быть одновременно истинными. Другими словами, каким бы ни было высказывание *A*, составное высказывание $A\ \&\ \overline{A}$ всегда ложно. В этом легко убедиться, построив таблицу истинности для высказывания $A\ \&\ \overline{A}$.

A	$\overline{A}$	$A\ \&\ \overline{A}$
И	*Л*	*Л*
Л	*И*	*Л*

Привести рассуждение к противоречию означает показать, что, допуская истинность какого-нибудь высказывания, мы приходим в результате к его ложности, или, другими словами, начав с какого-то утверждения, мы в конце концов приходим к его отрицанию.

Простак. Рассматривая в предыдущей беседе логические задачи, мы постоянно сталкивались с противоречием, которое и мешало получить решение. Противоречие лежит в основе любого парадокса!

З а н у д а. Однако противоречивые высказывания нередко встречаются и в обыденной речи, причем они вполне осмысленны и не вызывают ощущения нарушения законов логики. Например, высказывание "Эта таблетка и полезна, и вредна" легко трактуется как утверждение, что лекарство приносит пользу, но имеет и противопоказания, т. е. в одних случаях она полезна, а в других вредна. Тем не менее, формально данное высказывание содержит два взаимно исключающих высказывания, соединенных союзом И: "Эта таблетка полезна" И "Эта таблетка вредна".

П р о ф е с с о р. А действительно ли эти элементарные высказывания, содержащиеся в составном высказывании, исключают друг друга?

П р о с т а к. Они исключали бы друг друга, если бы относились к одной и той же ситуации. Трактовать сложное высказывание как истинное возможно только тогда, когда его компоненты соотносятся с различными ситуациями. Так, в одной из них таблетка полезна, а в другой вредна. Попробуем рассмотреть данное высказывание в контексте одной и той же ситуации. Например, применительно к случаю, когда врачу требуется решить, дать эту таблетку конкретному больному, или нет. Если этому больному таблетка полезна, то ее следует дать, иначе (если таблетка вредна) — нет. В этой ситуации сказать, что таблетка и полезна, и вредна, означает солгать.

З а н у д а. Не согласен! Я, конечно, понимаю проблему врача, которому необходимо на что-то решиться. Однако для конкретного больного одна и та же таблетка может быть одновременно и полезна, и вредна. Например, она может помочь его сердцу, но навредить печени.

П р о с т а к. Просто я не достаточно четко определил ситуацию. Возможно, требуется учесть еще какие-то дополнительные обстоятельства, надо хорошенько подумать. Я чувствую, что выход есть. Ведь не сомневаемся же мы, что высказывание "Дважды два равно четырем и дважды два не равно четырем" ложно!

П р о ф е с с о р. Друзья, вы забыли, что в математической логике отвлекаются от конкретного содержания высказываний, оперируя только их значениями истинности: любое высказывание может быть либо истинным, либо ложным. Элементарные высказывания можно комбинировать друг с другом, применяя союзы И и ИЛИ. При этом нас интересует лишь значение истинности результирующей комбинации. Поэтому мнение, что комбинация вида $A \,\&\, \overline{A}$ выражает противоречие, может быть выражено только как утверждение, что это ложное высказывание.

П р о с т а к. Тогда выходит, что математическая логика плохо применима к жизни, в которой помимо оценок истинности важен еще и смысл высказываний. Например, высказывание "Мы не одни во Вселенной" осмысленно, но вряд ли кто-нибудь отважится однозначно приписать ему одно из двух значений истинности. Я бы скорее сказал "не знаю", чем категорично утверждал, что оно истинно или, напротив, ложно.

Профессор. Вопрос о применимости к жизни математической логики, как и любой другой теории, находится за пределами самой теории. Не претендуя на универсальность, теория лишь моделирует действительность в той мере, в какой способна на это. Тем не менее, во многих случаях она помогает нам разобраться в сложных ситуациях. Например, допущение, что высказывание "Дважды два равно четырем и дважды два не равно четырем" ложно, довольно легко принимается нашим разумом, соответствует нашей интуиции. Здесь ситуация, в контексте которой или применительно к которой рассматривается данное высказывание, достаточно определенна. Это теория чисел или, иначе говоря, арифметика. Существуют и другие области знания, в которых математическая логика успешно работает.

Возвращаясь к примеру о таблетке, замечу, что трудность решения задачи в целом обусловлена трудностью определения истинности элементарных высказываний "Эта таблетка полезна" и "Эта таблетка вредна". Пусть врач решит на основе специальных исследований, какое из этих двух высказываний истинно применительно к данной ситуации. Далее, он должен принять или отвергнуть допущение, что эти элементарные высказывания противоположны друг другу в том смысле, что второе из них является равносильной перефразировкой высказывания "Эта таблетка не полезна" ("Не верно, что эта таблетка полезна"). Например, врач может исходить из принципа выбора наименьшего зла. Если он принимает такое допущение, то должен также принять, что одновременно истинными противоположные элементарные высказывания быть не могут.

Итак, в математической логике противоречие, заключающееся в одновременной истинности высказывания A и его отрицания $\overline{A}$, выражается в том, что составное высказывание $A \mathbin{\&} \overline{A}$ ложно.

2.1.2. Закон исключения третьего

Профессор. Закон исключения третьего говорит, что высказывание и его отрицание не могут быть ложными одновременно: либо истинно A, либо истинно $\overline{A}$. Используя таблицу истинности для определения $A \vee B$, построим частную таблицу истинности для определения $A \vee \overline{A}$:

A	$\overline{A}$	$A \vee \overline{A}$
И	*Л*	*И*
Л	*И*	*И*

Высказывание и его отрицание, соединенные союзом ИЛИ, всегда образуют истинное высказывание. Абсолютно (тождественно) истинное высказывание не информативно. Помните, в истории про Буратино народный лекарь Богомол констатировал: "Или пациент жив, или он умер". Очевидно, что независимо от пациента и его действительного состояния данное высказывание истинно. Аналогично, высказывание "На Марсе есть жизнь или ее там нет" воспринимается как истинное. Тождественно истинные высказывания можно конструировать сколько угодно. Их использование в обычной речи, особенно в спорах, может создать иллюзию аргументированности умозаключений. Однако это просто пустословие.

Простак. А я не понял, почему в названии этого закона говорится об исключении какого-то третьего?

Профессор. При формулировке данного закона предполагалось, что значения истинности высказываний A и $\overline{A}$ противоположны: если A истинно, то $\overline{A}$ ложно, и наоборот. Если высказывание $A \vee \overline{A}$ истинно, то возможны только два варианта:

- если истинно высказывание A, то высказывание $\overline{A}$ ложно;
- если истинно высказывание $\overline{A}$, то высказывание A ложно.

Третьего и других вариантов нет. Поэтому закон исключения третьего можно сформулировать и так: "Либо A, либо $\overline{A}$, и третьего не дано". Если же быть более точным, то этому утверждению следует придать такой вид: "Истинно A либо $\overline{A}$, но не оба вместе".

А теперь рассмотрим задачу. Что вы можете сказать о значении истинности высказывания "Слова *Абракадабра* нет в этой книге"?

Простак. Чтобы ответить на Ваш вопрос, необходимо внимательно просмотреть эту книгу. Если в ней нет данного слова, то высказывание истинно, в противном случае — ложно.

З а н у д а. Чтобы дать правильный ответ, совсем не нужно просматривать всю книгу, поскольку уже и так видно, что искомое слово находится в самом вопросе, а он напечатан в данной книге. Следовательно, рассматриваемое высказывание ложно. А раз так, то должно быть истинным его отрицание, т. е. высказывание "Слово *Абракадабра* есть в этой книге".

П р о ф е с с о р. Вы легко справились с этой задачей. Хотя задача и была несложной. Я лишь хотел обратить ваше внимание на то, что о каком бы слове *x* ни шла речь в высказывнии вида "Слово *x* не содержится в этой книге", само это высказывание всегда будет ложным. И наоборот, противоположное высказывание "Слово *x* содержится в этой книге" всегда будет истинным, т. е. тавтологичным.

П р о с т а к. Таким образом, мы не можем в некотором тексте сформулировать на том же языке истинное высказывание о том, что какого-то слова в этом тексте нет.

П р о ф е с с о р. Я бы не торопился со столь категоричным утверждением.

З а н у д а. Действительно, достаточно сформулировать высказывание как "Слова *Абракадабра* нет в этой книге, за исключением данного предложения", чтобы оно стало истинным, если, конечно, это слово и в самом деле далее нигде в тексте не встречается.

П р о ф е с с о р. Пожалуй, это выход.

2.1.3. Совместное использование НЕ, И и ИЛИ

П р о ф е с с о р. Теперь мы остановимся на двух важных правилах применения отрицания к составным высказываниям, образованным с помощью союзов И и ИЛИ. Вот эти правила, называемые законами Моргана:

$$\overline{A \& B} = \overline{A} \vee \overline{B};$$

$$\overline{A \vee B} = \overline{A} \& \overline{B}.$$

Примечание

Доказать эти формулы можно с помощью таблицы истинности для операций & и $\vee$. Попробуйте это сделать самостоятельно на досуге.

Законы Моргана позволяют увидеть связь между законами противоречия и исключения третьего: отрицание высказывания $A \mathbin{\&} \overline{A}$ равносильно высказыванию $A \vee \overline{A}$ и, наоборот, отрицание высказывания $A \vee \overline{A}$ равносильно высказыванию $A \mathbin{\&} \overline{A}$. Иначе говоря, верны следующие равенства:

$$\overline{A \mathbin{\&} \overline{A}} = A \vee \overline{A};$$

$$\overline{A \vee \overline{A}} = A \mathbin{\&} \overline{A}.$$

Как видите, отрицание противоречия равносильно высказыванию, соответствующему закону исключения третьего, и, наоборот, отрицание исключения третьего равносильно противоречию.

З а н у д а. Для доказательства этих равенств необходимо, кроме прочего, использовать правило двойного отрицания $\overline{\overline{A}} = A$.

П р о ф е с с о р. Совершенно верно. И добавим, что от перемены мест элементарных высказываний значение истинности составного высказывания не изменяется. Этот факт тоже можно доказать, используя таблицу истинности операций & и $\vee$.

Как я уже отмечал, законы противоречия и исключения третьего дополняют друг друга. Отрицание высказывания, соответствующего одному из них, равносильно высказыванию, соответствующему другому закону, и наоборот.

Давайте в качестве примера рассмотрим высказывание сказочного Колобка "Я не низок, не высок". Это высказывание понимается вполне правильно и легко допускается его истинность, хотя свойство "быть низким" противоположно свойству "быть высоким". Мы ведь сразу понимаем, что Колобок хотел сказать, что он средних размеров. Предлагаю вам рассмотреть это высказывание с точки зрения математической логики высказываний.

П р о с т а к. В символическом виде фразу Колобка можно переписать так: $\overline{\textit{я\_низок}} \mathbin{\&} \overline{\textit{я\_высок}}$. Тогда по второму правилу Моргана это высказывание равносильно следующему: $\overline{\textit{я\_низок} \vee \textit{я\_высок}}$, т. е. "Неверно, что я низок или высок". А почему бы и нет? Здесь само собой напрашивается предположение, что Колобок какого-то среднего роста, т. е. не низкий и не высокий.

З а н у д а. А как же быть с законом противоречия? По этому закону высказывание *я_низок* & *я_высок* должно быть ложным, поскольку $я\_высок = \overline{я\_низок}$, а $я\_низок = \overline{я\_высок}$. Ведь свойство "быть низким" противоположно свойству "быть высоким".

П р о ф е с с о р. Противоположные свойства и понятия не обязательно исключают друг друга. В данном примере можно представить себе шкалу размеров: низкий, ниже среднего, средний, выше среднего, высокий. Очевидно, что если Колобок не низкий, то он либо ниже среднего, либо выше среднего, либо высокий. Если Колобок не высокий, то он может быть выше среднего, средним, ниже среднего или же низким. Таким образом, для рассмотренной модели неверно, что $я\_высок = \overline{я\_низок}$, а также неверно, что $я\_низок = \overline{я\_высок}$. Как мы видим, противоречия нет.

З а н у д а. Действительно, я поторопился, допустив, что противоположные понятия исключают друг друга.

П р о ф е с с о р. Итак, мы рассмотрели операции НЕ, И и ИЛИ, которые определяются своими таблицами истинности. Этих операций вполне достаточно, чтобы реализовать любую таблицу всевозможных сочетаний значений истинности трех высказываний *A*, *B* и *C*. Поэтому набор данных операций называют полным.

П р о с т а к. А разве есть еще какие-то осмысленные логические операции?

П р о ф е с с о р. Конечно, есть. Например, операция исключающего ИЛИ. В отличие от рассмотренной до сих пор операции ИЛИ, исключающее ИЛИ образует из двух высказываний третье, истинное только тогда, когда истинно только какое-нибудь одно из исходных высказываний, но не оба сразу. Обозначим эту операцию как XOR. Тогда ее определение можно дать с помощью следующей таблицы истинности:

A	***B***	***A* XOR *B***
И	*И*	*Л*
И	*Л*	*И*
Л	*И*	*И*
Л	*Л*	*Л*

Нетрудно проверить, что выполняется равенство, показывающее, как операцию XOR можно выразить через НЕ, И и ИЛИ:

$$A \text{ XOR } B = (A \,\&\, \overline{B}) \vee (\overline{A} \,\&\, B).$$

Существуют и другие важные операции, которые можно выразить через НЕ, И и ИЛИ, но мы рассмотрим их позже.

2.1.4. Перестановка, группировка и распределение высказываний

Профессор. Теперь мы рассмотрим важные свойства операций И и ИЛИ. Думаю, вы согласитесь, что верны следующие законы, называемые законами перестановочности, или коммутативности:

$$A \,\&\, B = B \,\&\, A;$$

$$A \vee B = B \vee A.$$

Зануда. Хотя данные законы кажутся самоочевидными, все же хотелось бы иметь их доказательства.

Профессор. Убедиться в правильности законов коммутативности нетрудно с помощью анализа таблиц истинности для операций & и $\vee$.

Зануда. И как я упустил из виду, что все законы мы обосновываем с помощью таблиц истинности, являющихся определениями логических операций!

Профессор. Аналогичным способом доказывается выполнимость законов группировки, или ассоциативности:

$$(A \,\&\, B) \,\&\, C = A \,\&\, (B \,\&\, C);$$

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

Круглые скобки особенно полезны, когда рассматриваются сложные высказывания с применением операций & и $\vee$. Для таких высказываний выполняются законы распределения, или дистрибутивности:

$$A \,\&\, (B \vee C) = (A \,\&\, B) \vee (A \,\&\, C);$$

$$A \vee (B \,\&\, C) = (A \vee B) \,\&\, (A \vee C).$$

2.1.5. Логическое следование и эквивалентность

Профессор. Как я уже отмечал, набор операций НЕ, И и ИЛИ является полным в том смысле, что с помощью данных операций можно реализовать любую таблицу истинности, а значит, и любую логическую операцию. Поясню это подробнее. Допустим, мы придумали какую-то новую логическую операцию, которую обозначили как *. Определение этой операции можно представить в виде некоторой таблицы истинности. Какой бы ни была эта таблица, операцию * можно выразить через операции НЕ, И и ИЛИ. Ранее мы рассмотрели в качестве примера определение через данные операции исключающего ИЛИ (XOR).

Теперь мы рассмотрим две часто используемые логические операции — логическое следование одного высказывания из другого и эквивалентность двух высказываний. Мы увидим, что эти новые операции могут быть выражены через уже известные нам операции НЕ, И и ИЛИ.

В рассуждениях, особенно претендующих на статус доказательства или логического вывода следствий из посылок, часто используются предложения вида "Если A, то B". Здесь высказывание A называют посылкой (условием), а высказывание B — следствием (заключением). Точнее, такого рода предложение следует понимать следующим образом: "Если истинно A, то истинно B". Вместе с тем, это предложение можно рассматривать как высказывание, которое с точки зрения булевской логики само по себе может быть либо истинным, либо ложным, независимо от того, существует ли в действительности некая причинно-следственная или иная смысловая связь между высказываниями A и B. А как определить значение истинности высказывания "Если A, то B" формально, не анализируя причинно-следственных или каких бы то ни было других смысловых, глубинных связей между A и B? В булевской логике значение истинности этого высказывания определяется только в зависимости от значений истинности высказываний A и B, подобно тому, как это делалось для определения значений истинности высказываний A & B и $A \vee B$. Для краткости обозначим высказывание вида "Если A, то B" как A => B.

П р о с т а к. Если я правильно Вас понял, то достаточно построить таблицу истинности для высказывания A => B. Если поступить чисто формально, то с этой целью достаточно просто для каждой из четырех возможных комбинаций значений истинности высказываний A и B назначить подходящее значение истинности для составного высказывания "Если A, то B".

П р о ф е с с о р. Верно. Вопрос только в том, чтобы выбрать подходящее значение истинности. Очевидно, что выбор не может быть произвольным. Мы хотели бы, чтобы он соответствовал, насколько это возможно, нашей интуиции или обычному словоупотреблению высказывания типа "Если A, то B".

З а н у д а. Это не так уж сложно сделать, поскольку количество возможных вариантов невелико, а именно 4.

П р о ф е с с о р. Разумеется, мы могли бы рассмотреть все возможные варианты. Но перебор можно сократить, если учесть соображение общего порядка. А именно: высказывание *A* => *B* ложно только тогда, когда посылка *A* истинна, а следствие *B* — ложно; в остальных случаях это высказывание истинно. Давая такое определение, мы учитываем следующие обстоятельства:

- если посылка ложная, то из нее может следовать как истинное, так и ложное заключение (из лжи следует что угодно);
- из истинной посылки может следовать только истинное заключение.

Таким образом, высказывание *A* => *B* ложно только тогда, когда посылка *A* истинна, а заключение (следствие) *B* ложно; в остальных случаях *A* => *B* истинно.

Обобщая вышеизложенное, можно построить следующую таблицу истинности, которая и является определением операции => логического следования, моделирующей в булевской логике союз "если ..., то ...":

A	***B***	***A* => *B***
И	*И*	*И*
И	*Л*	***Л***
Л	*И*	*И*
Л	*Л*	*И*

В математике операция => называется импликацией (англ. implication — следствие, вывод), точнее, материальной импликацией. Обратите внимание: значение истинности импликации *A* => *B* никак не обусловлено ни смысловой, ни причинно-следственной связью высказываний *A* и *B*, ни содержанием каждого из них в отдельности. Например, высказывание "Если $2 \times 2 = 4$, то $2 + 3 = 5$" истинно потому, что истинны посылка и следствие. В данном случае смысловая связь между посылкой и следствием действительно может быть установлена в арифметике, но ее наличие не является необходимым условием, чтобы принять истинность самой импликации. Так, высказывание "Если $2 \times 2 = 10$, то $2 + 3 = 5$" является истинным, поскольку истинно следствие, хотя посылка явно ложная. По той же причине мы принимаем, что высказывание "Если я папа римский, то Лев Толстой — автор романа «Война и мир»".

П р о с т а к. А как насчет высказывания "Если в огороде бузина, то в Киеве дядька"?

З а н у д а. Все зависит от значений истинности посылки "в огороде бузина" и следствия "в Киеве дядька". Хотя между ними нет никакой причинно-следственной связи, данную импликацию следует признать ложной, только если в огороде есть бузина, а дядьки в Киеве нет. Или, иначе говоря, импликация истинна, если выполняется хотя бы одно из условий:

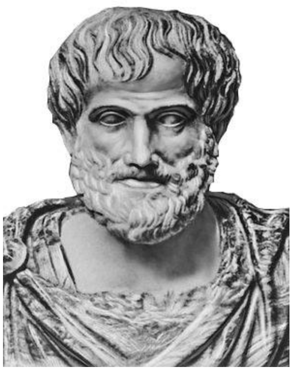
- высказывание "в Киеве дядька" истинно;
- высказывание "в огороде бузина" ложно.

П р о ф е с с о р. Правильно. Импликация не вполне соответствует использованию условных утверждений "если …, то …" в обычной речи. При обычном словоупотреблении обращается внимание на значение истинности следствия, а не условного утверждения в целом. Конечно, если само утверждение "Если *A*, то *B*" истинно и, кроме того истинна посылка *A*, то истинно и следствие *B*. Это правило вывода, называемое еще силлогизмом, можно символически записать так:

$$\frac{A, A => B}{B}.$$

Здесь над чертой указаны посылки силлогизма (правила вывода), принимаемые как истинные высказывания, а под чертой — выводимое следствие.

Данное правило вывода (силлогизм) введено в формальную логику Аристотелем, учителем Александра Македонского. Оно получило специальное название modus ponens, поскольку в логике используются и другие правила вывода. Однако modus ponens — основное, наиболее важное правило логического вывода, называемое еще правилом отделения, или гипотетическим силлогизмом.



Аристотель (384—322 гг. до н. э.)

Чтобы не путаться, я рекомендую читать высказывание вида $A => B$ как "A влечет B" или "Из A следует B". Это особенно полезно при осмыслении условного предложения, содержащего импликацию $A => B$. Например, предложение "Если верно, что если A, то B, то ..." нелегко понять; предложение "Если верно, что из A следует B, то ..." кажется более понятным и благозвучным.

Операцию импликации =>, служащую для моделирования высказываний вида "Если A, то B", "A влечет B" или "Из A следует B", можно выразить через ранее рассмотренные операции отрицания и дизъюнкции (ИЛИ):

$$A => B = \overline{A} \vee B.$$

Данное равенство легко доказать на основе сравнения таблиц истинности для $A => B$ и $\overline{A} \vee B$. Я рекомендую использовать высказывание $\overline{A} \vee B$ в тех случаях, когда требуется доказать истинность или ложность высказывания $A => B$. Так, например, многие теоремы в математике формулируются в виде "Если A, то B". Чтобы доказать подобную теорему, достаточно доказать истинность равносильного утверждения $\overline{A} \vee B$, т. е. что следствие B истинно, или посылка A ложна. Если теорема верна, то для ее доказательства достаточно установить истинность чего-нибудь одного: либо $\overline{A}$, либо B. А если теорема не верна, то потребуется удостовериться, что ложны оба утверждения — и $\overline{A}$, и B.

С помощью таблиц истинности нетрудно установить следующее равенство:

$$A => B = \overline{B} => \overline{A}.$$

Это равенство лежит в основе доказательства утверждений методом от противного, который очень часто применяется, когда прямое доказательство не удается найти: оно оказывается либо слишком трудоемким, либо невозможным принципиально. Далее я приведу схему доказательства методом от противного теоремы вида "Если A, то B".

Формулировка теоремы понимается так: если верно (имеет место, истинно) утверждение A, то верно (имеет место, истинно) утверждение B. В доказательстве сначала допускают противное, а именно — что утверждение B ложно (т. е. $\overline{B}$ истинно). Затем каким-то образом (сейчас не важно, каким именно) показывают, что при таком допущении утверждение A также оказывается ложным (т. е. $\overline{A}$ истинно). Но это невозможно, поскольку в условии теоремы предполагалось, что A истинно. Иначе говоря, допустив ложность B мы пришли к противоречию. Следовательно неверно, что B ложно, а это означает, что B, истинно! В дальнейшем мы неоднократно будем применять этот метод доказательства утверждений.

П р о с т а к. Насколько я помню еще со школы, многие теоремы формулируются с указанием вида условий. Например, условия могут быть необходимыми, достаточными, а также одновременно и необходимыми, и достаточными. Кроме того, нередко встречаются формулировки вида "… если и только если …", "… тогда и только тогда …". Не могли бы Вы, Профессор, пояснить это?

П р о ф е с с о р. В обычной речи вообще и в математических теоремах в частности нередко утверждается, что какое-то высказывание *B* верно не само по себе, а при выполнении некоторого условия *A*. Условия могут быть необходимыми и достаточными для того, чтобы высказывание *B* было верным (истинным). Условие *A* является необходимым для *B*, если без его выполнения высказывание *B* оказывается ложным. Условие *A* достаточно для *B*, если при его выполнении высказывание В оказывается истинным.

Рассмотрим, например, высказывание "Целое число делится на 2 без остатка". Возьмем в качестве условия высказывание "Десятичная запись числа не оканчивается на 3". Выполнение этого условия для числа совсем не гарантирует, что данное число делится на 2 без остатка (например, 15). С другой стороны, если условие не выполняется, то число, оказываясь нечетным, не делится на 2 без остатка. Таким образом, рассмотренное условие является необходимым для делимости числа на 2. Возьмем теперь условие "Десятичная запись числа оканчивается на 0". Очевидно, что при его выполнении число делится на 2 без остатка, каким бы оно ни было. Поэтому данное условие является достаточным. В то же время оно не является необходимым, поскольку при его невыполнении (например, 14) число может делиться на 2.

Итак, для рассмотренного примера условные утверждения о делимости целых чисел пополам можно сформулировать следующим образом:

- "Чтобы целое число делилось на 2 без остатка, *необходимо*, чтобы его десятичная запись не оканчивалась на 3";
- "Чтобы целое число делилось на 2 без остатка, *достаточно*, чтобы его десятичная запись оканчивалась на 0".

П р о с т а к. Если объединить оба условия, то можно получить необходимое и достаточное условие делимости целых чисел пополам без остатка. Однако легко заметить, что мы не получим определения всех таких чисел. Например, числа 5, 7 и многие другие не удовлетворяют такому определению.

З а н у д а. Выходит, что условие, являющееся одновременно и необходимым и достаточным, не может выполнять роль определения чего либо?

П р о ф е с с о р. Не торопитесь с выводами, друзья. Объединить какие-то необходимое и достаточное условия союзом И еще не означает получить единое условие, являющееся одновременно и необходимым, и достаточным. Прежде всего, я хотел бы обратить ваше внимание на то, что необходимое условие задает, вообще говоря, не все свойства определяемого объекта. Кроме того, этому условию, может статься, удовлетворяют и другие объекты, а не только определяемый. На практике стремятся найти такое необходимое условие, которому удовлетворяло бы возможно меньшее количество объектов. Для примера о делимости пополам можно было бы составить такое необходимое условие: "Число должно быть 2, или 4, или 6, или 8, или 10 и т. д.". Очевидно, что для подобной формулировки потребуется бесконечно много слов. Однако не в этом суть. Объем формулировки можно сделать конечным, всего в несколько строк, которые бы определяли четные числа. Дело в том, что необходимое условие определяет не только то, что определяется другой частью условного предложения. В противном случае это условие было бы и необходимым, и достаточным.

С другой стороны, достаточному условию могут удовлетворять не все определяемые объекты. В принципе, достаточное условие можно было бы сформулировать так, чтобы ему удовлетворяло только какое-нибудь одно число. Например: "Чтобы целое число делилось на 2 без остатка, *достаточно*, чтобы это было число 2". На практике стараются найти такое достаточное условие, которому удовлетворяло бы возможно большее количество объектов. Если удается сформулировать достаточное условие так, чтобы оно определяло только то, что определяется другой частью условного предложения, то это условие становится к тому же и необходимым.

Итак, для определения объекта достаточное условие стараются возможно более расширить, а необходимое условие, наоборот, сузить. Области охвата, задаваемые этими условиями, постепенно сближаются до тех пор, пока в них не окажется только определяемый объект.

П р о с т а к. А каково же необходимое и достаточное условие делимости целых чисел пополам без остатка?

П р о ф е с с о р. Среди нескольких возможных определений, выражаемых через необходимое и достаточное условие, приведу, например, такое: "Чтобы целое число делилось на 2 без остатка, *необходимо и достаточно*, чтобы последняя *цифра* в его десятичной записи делилась на 2 без остатка."

Простак. А как насчет условных предложений вида "... если и только если ...", "... тогда и только тогда ...", "... в том и только том случае ..."?

Профессор. Все эти предложения имеют тот же смысл, что и предложение вида "... необходимо и достаточно ...".

В предложении вида "Если верно A, то верно B", записываемом символически как импликация $A => B$, истинность высказывания B является необходимым условием для истинности высказывания A, а истинность A является достаточным условием для истинности B. Рассмотрим, почему это так.

Предполагается, что импликация $A => B$ истинна. Но тогда если A истинно, то и B истинно. Таким образом, истинность высказывания A достаточна для того, чтобы B было истинным. Теперь допустим, что B ложно, тогда и A ложно. Следовательно, истинность B необходима, чтобы высказывание A было истинным.

Простак. А как выглядит случай, когда истинность A и необходима, и достаточна для истинности B?

Профессор. Чтобы ответить на Ваш вопрос, сначала определим операцию эквивалентности, или логического тождества двух высказываний, которую обозначим как <=>. С помощью этой операции из двух высказываний, A и B, получается третье высказывание $A <=> B$, которое истинно только тогда, когда высказывания A и B имеют одинаковые значения истинности. Точнее, эта операция определяется следующей таблицей истинности:

A	***B***	***A <=> B***
И	*И*	*И*
И	*Л*	*Л*
Л	*И*	*Л*
Л	*Л*	*И*

Нетрудно доказать, что справедливо следующее равенство:

$$A <=> B = (A => B) \ \& \ (B => A).$$

Утверждение об эквивалентности двух высказываний равносильно тому, что из первого следует второе и, наоборот, из второго следует первое. Формулировка теоремы об эквивалентности A и B имеет следующий вид: "Для того чтобы имело место A, необходимо и достаточно, чтобы имело место B". Разумеется, вместо слов "необходимо и достаточно" можно подставить слова "если и только если", "тогда и только тогда" и т. п. Доказательство такой теоремы разбивается на две части:

- доказательство того, что из истинности A следует истинность B (A => B); на данном этапе мы убеждаемся, что истинность A достаточна для истинности B, а истинность B необходима для истинности A;
- доказательство того, что из истинности B следует истинность A (B => A); на данном этапе мы убеждаемся, что истинность B достаточна для истинности A, а истинность A необходима для истинности B.

В итоге мы получаем доказательство того, что истинность A необходима и достаточна для истинности B, и наоборот. Таким вот образом происходит объединение необходимого и достаточного условий, в результате которого выясняется, что все определяемое высказыванием A определяется и высказыванием B и наоборот. Другими словами, выясняется, что высказывания A и B логически эквивалентны, хотя при этом они могут иметь различные формы выражения.

З а н у д а. Я заметил, профессор, что понятие эквивалентности высказываний (<=>) является тем же, что и введенное в самом начале понятие их равносильности (=). Быть может, я просто не заметил какие-то тонкие различия?

П р о ф е с с о р. Эквивалентность (<=>) является, как и рассмотренные ранее операции (отрицание, дизъюнкция и конъюнкция), операцией алгебры логики, с помощью которой из двух высказываний A и B получается третье высказывание A <=> B. Если хотите, этому новому высказыванию мы можем присвоить специальное обозначение, например, C, и написать равенство C = A <=> B, выражающее равносильность двух высказываний. Высказывания в алгебре логики интересуют нас только с точки зрения их значений истинности. Поэтому высказывание C рассматривается либо как истинное, либо как ложное. Равносильность (=) это не операция, а отношение между высказываниями, которое применяется в рассуждениях о высказываниях. Так, например, если мы обозначили высказывание A <=> B как C, то теперь можем утверждать, что высказывания A <=> B и C равносильны, что символически записывается как A <=> B = C. Рассуждения о высказываниях не принадлежат алгебре логики, они находятся вне ее. Так, утверждение, что два сравниваемых высказывания имеют одинаковые значения истинности, не важно какие именно, есть утверждение на естественном языке о равносильности этих вы-

сказываний. В сложных высказываниях их составные части могут быть заменены равносильными высказываниями, чем и объясняется польза утверждений о равносильности. Например, импликацию $A => B$ можно заменить равносильным высказыванием $\overline{A} \vee B$. Однако выражение $A => B = \overline{A} \vee B$ не является высказыванием алгебры логики, а есть утверждение на внешнем языке — естественном языке рассуждений о высказываниях алгебры логики. Так, символическое утверждение $A => B = \overline{A} \vee B$ переводится на обычный естественный язык как "$A => B$ равносильно $\overline{A} \vee B$" или, точнее, как "Верно, что высказывание 'из A следует B' равносильно высказыванию 'не A или B'". Обратите внимание, что высказывая утверждение $A => B = \overline{A} \vee B$, мы подразумеваем, что оно истинно. Напротив, высказывание $A => B <=> \overline{A} \vee B$ может быть как истинным, так и ложным. Приводя данную формулу, мы ничего не утверждаем о значении ее истинности. Вместе с тем, в алгебре логики мы хотим использовать, среди прочих, и высказывания, которые в какой-то степени моделируют высказывания о равносильности. Пожалуйста, делайте это с помощью подходящих операций алгебры логики. Так, утверждение "$A => B$ равносильно $\overline{A} \vee B$" моделируется в алгебре логики как высказывание $(A => B) <=> (\overline{A} \vee B)$. Здесь круглые скобки служат для того, чтобы разделить члены операции эквивалентности.

П р о с т а к. Поясните, пожалуйста, изложенное Вами на более конкретных примерах.

П р о ф е с с о р. Рассмотрим два высказывания "Саша любит Машу" и "Маша любит Сашу". Сейчас нас не интересует, истинны данные высказывания или ложны, хотя мы знаем, что приписать значение истинности этим высказываниям мы, так или иначе, сможем. Из данных высказываний мы можем сконструировать новое высказывание "Саша любит Машу" <=> "Маша любит Сашу". Что можно сказать о значении истинности этого составного высказывания? Используя определение операции <=> посредством таблицы истинности, мы заключаем, что составное высказывание истинно только в двух случаях:

❒ "Саша любит Машу" — истинно и "Маша любит Сашу" — истинно;

❒ "Саша любит Машу" — ложно и "Маша любит Сашу" — ложно.

В остальных случаях данное составное высказывание ложно. Вот и все, что мы можем сказать о составном высказывании.

Теперь рассмотрим выражение "Саша любит Машу" = "Маша любит Сашу". Что можно сказать о значении истинности этого составного высказывания? А только то, что оно истинно. Точнее, данное предложение, использующее символ = вместо слова "равносильно", утверждает (т. е. допускает истинность) того, что значения истинности элементарных высказываний "Саша любит Машу" и "Маша любит Сашу" одинаковы. Это означает, что оба элементарных высказывания либо истинны, либо ложны. Другими словами, если верно, что Саша любит Машу, то верно, что Маша любит Сашу, и, наоборот, если Маша любит Сашу, то верно, что Саша любит Машу. Вообще говоря, в жизни не всегда имеет место симметричность любовных отношений. Однако в данном случае эта симметричность утверждается предложением о равносильности высказываний "Саша любит Машу" и "Маша любит Сашу".

Обобщим сказанное. Выражение $A <=> B$ есть высказывание, значение истинности которого вычисляется в зависимости от значений истинности членов A и B в соответствии с определением операции <=>. Имея дело с высказыванием $A <=> B$, мы интересуемся значением его истинности. Напротив, выражение $A = B$ есть утверждение о том, что значения истинности членов A и B одинаковы. Утверждение о высказывании X это высказывание, что X — истинно. Имея дело с выражением $A = B$, мы не интересуемся значением его истинности (оно истинно в силу своего существования, констатации), но используем его в анализе путем подстановок в выражения высказываний, имеющих сложную структуру.

З а н у д а. Верно ли, что выражение $A = B$ имеет тот же смысл, что выражение "$A <=> B$ — истинное высказывание"?

П р о ф е с с о р. Да, можно сказать и так.

Я рекомендую вам в качестве упражнения самостоятельно доказать истинность и равносильность некоторых высказываний:

$$A => A$$

$$A => (A \vee B)$$

$$A => (B => A)$$

$$(A => B) => ((B => C) => (A => C))$$

$$(A \vee B) = (A => B) => B$$

$$A <=> B = (A \& B) \vee (\overline{A} \ \& \ \overline{B})$$

$$A \text{ XOR } B = \overline{A <=> B}\,.$$

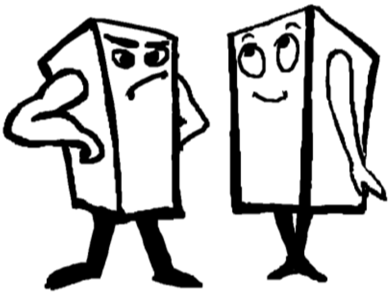
2.2. Предикаты, или высказывания с аргументами

Профессор. Высказывания, рассмотренные ранее, могут выражать суждения о свойствах и отношениях конкретных объектов. Например, в высказывании "Яблоко спелое" выражается суждение о свойстве "быть спелым" не для произвольного объекта, а для яблока; в высказывании "Саша любит Машу" выражается суждение об отношении любви между именно Сашей и Машей, а не просто произвольными людьми. Однако одним и тем же свойством могут обладать или не обладать различные объекты. Аналогично, одно и то же отношение может связывать или не связывать различные

объекты. Иначе говоря, свойства и отношения можно рассматривать сами по себе, не в связи с конкретными объектами. Выразительные возможности математической логики существенно расширяются, если использовать выражения с аргументами (переменными), например, "x является красным", "x любит y", "$5x^2 + 10x - 15 = 0$". В этих выражениях через x и y обозначены переменные, вместо которых можно подставить объекты (точнее, их уникальные имена, идентификаторы). Само по себе выражение с переменными не является высказыванием, подобно тому как числовая функция, заданная формулой, не имеет значения, пока все ее переменные не будут замещены конкретными числами. По той же причине выражение "Нечто является красным" также не может принять конкретного значения истинности, пока мы не определим, что понимается под словом "Нечто". Как только в выражение вместо переменных будут подставлены идентификаторы конкретных объектов, оно превратится в высказывание, имеющее одно из двух значений истинности — ИСТИНА или ЛОЖЬ.

Выражение с переменными, принимающее одно из двух значений истинности после замещения переменных объектами, называется предикатом (англ. predicate — утверждение), или пропозициональной функцией (англ. proposition — предложение, суждение). Совокупность всех объектов (значений переменных предиката) называется универсумом.

Простак. Понятно, что выражение "x является красным" есть предикат, а не высказывание. Если слово "Яблоко" однозначно идентифицирует конкретный объект, например, данное яблоко, то выражение "Яблоко является красным" есть высказывание, которое истинно или ложно в зависимости от действительного цвета данного яблока. Как Вы сказали,



Профессор, на объекты не накладывается никаких ограничений. Но тогда возможны некоторые недоразумения. Подставим в предикат "*x* является красным" вместо переменной *x* число 2, тогда получится высказывание "2 является красным". Но ведь это абсурд, поскольку ни одно число не имеет цвета, запаха, вкуса и других подобных свойств. Иначе говоря, число не является ни красным, ни некрасным, и поэтому данное высказывание нельзя считать ни истинным, ни ложным. Аналогичным образом дело обстоит и с высказыванием "Кирпич имеет мягкий характер". Если еще можно согласиться с тем, что "Кирпич имеет твердое тело", то очевидно, что высказывание "Кирпич имеет нежную душу" совершенно абсурдно.

З а н у д а. Я согласен с Простаком: есть высказывания, которые трудно или даже невозможно оценить в терминах истинности и ложности. Скорее всего, это бессмысленные высказывания, либо поэтические метафоры, вроде "зимнее лето" и "летняя зима", "крылатое выражение", "глаз — алмаз" и т. д. Может быть, исключить такого рода высказывания из поля зрения математической логики?

П р о с т а к. К чему столь кардинальные меры, Зануда? Мне кажется, что достаточно для каждого предиката определить свой универсум объектов, к котором данный предикат применим.

П р о ф е с с о р . Во-первых, если высказывание бессмысленно или абсурдно по своему содержанию, то это совсем не значит, что оно и не истинно, и не ложно. Например, высказывания "2 является красным" и "Кирпич имеет нежную душу" можно считать ложными, опираясь на то, что числа не имеют цвета, а кирпич — неодушевленный предмет. Во-вторых, если мы для каждого предиката определим свой универсум объектов, к которым он применим, то тогда предикат будет не двузначной, а трехзначной пропо-зициональной функцией. Для объектов из универсума применимости такой предикат будет иметь значение ИСТИНА или ЛОЖЬ, а для всех остальных — какое-нибудь другое значение, например, NULL. В этом случае мы выйдем за границы классической булевской логики. Я предлагаю вам сейчас не делать этого. Наконец, рассматривая какой-то набор предикатов, мы можем для них всех как-то ограничить универсум, подобно тому как ограничивают предметную область или тему разговора. Например, занимаясь высказываниями о числах,

уместно считать, что универсум — это множество всех или только некоторое подмножество всех чисел.

Будем рассматривать только двузначные предикаты, которые могут принимать одно из двух возможных значений истинности после замены их переменных объектами из единого универсума. Мы также отвлечемся от трудностей, которые могут возникнуть при реальном определении значения тех или иных предикатов для каких-то конкретных объектов. В конце концов, мы можем не рассматривать абсурдные высказывания, подобно тому как в обычной речи мы стараемся избегать непонятных или чересчур замысловатых, "темных" выражений.

Итак, поскольку предикат — это двузначная функция, обозначим его как $P(x)$, где в скобках указана переменная, которая принимает значения из универсума объектов. Разумеется, вместо буквы P мы можем использовать любые символы или их последовательности. Например, предикат "x красный" в наших обозначениях можно записать как *Красный*(x). Предикат с одной переменной называют одноместным. Одноместные предикаты выражают свойства объектов. Если некоторый объект, обозначаемый, например, как a, обладает свойством P, то высказывание $P(a)$, получающееся в результате подстановки объекта a в выражение $P(x)$, является истинным. Если объект a не обладает свойством P, то высказывание $P(a)$ является ложным. Другими словами, предикат $P(x)$ для объекта a принимает значение истинности, совпадающее со значением истинности высказывания $P(a)$.

П р о с т а к. Если я правильно понял Вас, предикаты играют в логике такую же роль, что и обычные функции в области чисел. Например, $\sin(x)$ — функция, которая принимает некоторое значение только после подстановки вместо переменной x некоторого числа. Так, значение $\sin(\pi/2)$ равно 1, а значение выражения x^2 при $x = 3$ равно 9. Аналогично, предикат $P(x)$ есть некоторая форма, шаблон какого-то высказывания. Чтобы этот шаблон стал высказыванием, необходимо заместить переменную каким-нибудь ее значением.

П р о ф е с с о р. Верно. Именно в связи с данной аналогией предикаты еще называют пропозициональными функциями.

Далее, предикат, выражающий какое-нибудь отношение между двумя объектами, называется двухместным или бинарным. Примеры бинарных предикатов — выражения "x любит y" и "$x < y$". Символически бинарный предикат можно записать как функцию двух переменных $P(x, y)$. Например, *Любит*(x, y), $<(x, y)$. Способ записи предиката выбирается, исходя из удобства. Например, отношение "меньше, чем" обычно записывают как $x < y$, а не как $<(x, y)$.

Предикаты могут иметь и больше двух переменных. Предикат с тремя переменными называют тернарным. В случае $n > 3$ переменных предикаты называют n-местными или n-арными.

Предикаты могут быть равносильными или неравносильными. Два предиката $P(x)$ и $Q(x)$ равносильны ($P(x) = Q(x)$), если для любого объекта универсума они принимают одинаковые значения истинности. Например, предикаты "$x^2 = 1$" и "$(x - 1)(x + 1) = 0$" равносильны на универсуме вещественных чисел, поскольку их выражения могут быть преобразованы одно к другому. Надеюсь, вы помните, что $(x - 1)(x + 1) = x^2 - 1$. Если бы в каком-нибудь универсуме оказалось, что все красные объекты являются круглыми и наоборот, то для такого универсума предикаты (свойства) *Красный*(x) и *Круглый*(x) были бы равносильными.

Над предикатами можно выполнять такие же логические операции, как и над высказываниями: отрицание, конъюнкцию, дизъюнкцию, импликацию, эквивалентность. В результате применения операции к предикатам получаются другие предикаты. Так, операция отрицания, примененная к предикату $P(x)$, приводит к предикату $\overline{P(x)}$. Для любого объекта x если $P(x) = И$, то $\overline{P(x)} = Л$ и наоборот, если $P(x) = Л$, то $\overline{P(x)} = И$. Операции конъюнкции (&), дизъюнкции ($\vee$), импликации (=>) и эквивалентности (<=>) определяются, как и в случае высказываний, таблицами истинности:

P(x)	**Q(x)**	**P(x) & Q(x)**	**P(x) ∨ Q(x)**	**P(x) => Q(x)**	**P(x) <=> Q(x)**
И	*И*	*И*	*И*	*И*	*И*
И	*Л*	*Л*	*И*	*Л*	*Л*
Л	*И*	*Л*	*И*	*И*	*Л*
Л	*Л*	*Л*	*Л*	*И*	*И*

Например, конъюнкция (т. е. соединение логическим союзом И) предикатов "x твердый" и "x красный" образует предикат "x твердый и x красный", или, более лаконично, "x твердый и красный"; импликацию "x твердый => x красный" можно представить выражением "если x твердый, то он красный".

Как в обычной речи, так и в математической логике часто используются так называемые общеутвердительные и частноутвердительные высказывания. Предложение "все люди млекопитающие" является примером общеутвер-

дительного высказвания, а предложение "некоторые люди — женщины" — пример частноутвердительного высказывания. Заметим, что оба эти высказывания являются истинными. Разумеется, мы могли бы придумать и какие-нибудь ложные высказывания. В этих высказываниях, как нетрудно заметить, используются слова "все" и "некоторые", несущие некоторую функциональную нагрузку. Они указывают, насколько велика область значений переменных, где преди-кат является истинным. Поэтому данные слова называются *кванторами*. Слово "все" означает, что предикат истинен для всех (каждого, любого) элементов универсума, а слово "некоторые" — что предикат истинен по крайней мере для какого-нибудь одного из них. В математической логике эти кванторы обозначаются символами $\forall$ и $\exists$ и называются кванторами всеобщности и существования соответственно. Кванторы записываются перед предикатами с указанием переменной, к которой они относятся.

Выражение $\forall x P(x)$ понимается как высказывание, что для всех элементов универсума предикат $P(x)$ истинен. Это выражение можно прочитать следующими способами: "Для всех x верно $P(x)$", "Для каждого x верно $P(x)$", "Для любого x верно $P(x)$ ", "Для всякого x верно $P(x)$".

Выражение $\exists x P(x)$ означает, что предикат $P(x)$ истинен для некоторого элемента универсума. Его можно прочитать следующим образом: "Существует x, такой что верно $P(x)$", "Для некоторого x верно $P(x)$".

Обратите внимание: выражения $\forall x P(x)$ и $\exists x P(x)$ уже не зависят от значения x; они являются высказываниями и, следовательно, могут быть как истинными, так и ложными. Например, высказывание $\forall x(\sin(x)=1)$ ложно, а высказывание $\exists x(\sin(x)=1)$ истинно. Говорят, что квантор *связывает* переменную, к которой он относится. Так, в выражении $\forall x P(x)$ квантор всеобщности связывает переменную x.

З а н у д а. А чем является выражение $\forall x P(x, y)$ — высказыванием или предикатом?

П р о ф е с с о р. В данном выражении переменная x связана квантором, а переменная y — нет, она остается, как говорят, *свободной*. Таким образом, данное выражение зависит от одной переменной и, следовательно, является предикатом.

Далее, в случае конечного универсума объектов, состоящего из объектов a_1, a_2, …, a_n, высказывание $\forall x P(x)$ равносильно высказыванию $P(a_1)$ & $P(a_2)$ &...& $P(a_n)$. Аналогично, высказывание $\exists x P(x)$ равносильно высказыванию $P(a_1) \vee P(a_2) \vee ... \vee P(a_n)$. Нетрудно заметить, что если истинно высказывание $\forall x P(x)$, то истинно и высказывание $\exists x P(x)$, хотя обратное в общем случае неверно (если верно частное утверждение, то соответствующее общее утверждение может и не быть верным). Например, утверждение "Некоторые люди умны" по-видимому, истинно, но утверждение "Все люди умны" скорее ложно, чем истинно.

До сих пор мы рассматривали предикаты только с одной переменной. Однако предикаты могут содержать и несколько переменных. Например, предикат $x^2 = y$ предствляет собой равенство двух числовых выражений с двумя переменными — x и y. Следующее выражение является истинным высказыванием:

> $\forall x \exists y (x^2 = y)$ — для всех x существует y, такой что выполняется равенство $x^2 = y$.

Изменение типа квантора в выражении может изменить его значение истинности. Так например, следующее высказывание ложно:

> $\exists x \forall y (x^2 = y)$ — существует x, такой что для всех y выполняется равенство $x^2 = y$.

Однако для любого предиката $P(x, y)$ если истинно высказывание $\exists y \forall x P(x, y)$, то истинно и высказывание $\forall x \exists y P(x, y)$. Иначе говоря, перестановка кванторов всеобщности и существования вместе с соответствующими им переменными не изменяет значения истинности высказывания.

Пусть $P(x)$ — некоторый предикат. Тогда справедливы следующие равенства:

$$\overline{\forall x P(x)} = \exists x \overline{P(x)},$$

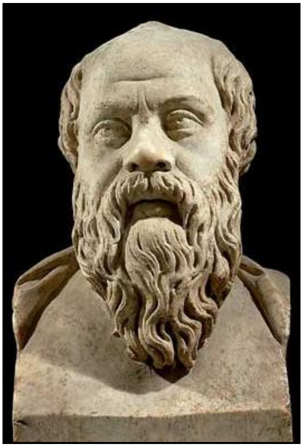
$$\overline{\exists x P(x)} = \forall x \overline{P(x)}.$$

Пусть, например, P обозначает свойство "быть красным". Тогда первое из двух указанных выше равенств означает, что высказывание "Неверно, что все x красные" равносильно высказыванию, что "Некоторые x не красные". Второе равенство утверждает равносильность высказываний "Неверно, что есть красный объект" и "Все объекты не красные".

В учебниках логики часто приводится в качестве примера следующее аристотелевское правило вывода (силлогизм): "Все люди смертны. Сократ — человек. Следовательно, Сократ смертен". Пусть *Человек*(x) и *Смертен*(x) — предикаты, выражающие свойства "быть человеком" и "быть смертным" соответственно. Тогда рассматриваемый силлогизм можно записать в такой форме:

$$\frac{\forall x(\text{\textit{Человек}}(x) => \text{\textit{Смертен}}(x)), \text{\textit{Человек}}(\text{\textit{Сократ}})}{\text{\textit{Смертен}}(\text{\textit{Сократ}})}.$$

Здесь предполагается, что посылки над горизонтальной чертой и следствие под ней являются истинными высказываниями.



Сократ (около 470—399 гг. до н. э.)

В повседневной речи часто встречаются выражения вида "Вообще говоря, …", в которых многоточие заменяется предложением о каком-нибудь свойстве или отношении между объектами, например, "Вообще говоря, сахар сладкий", "Вообще говоря, люди хорошие" и т. п. Я хотел бы обратить ваше внимание на то, что слова "вообще говоря" нередко употребляются неправильно. Что, например, вы думаете по этому поводу?

П р о с т а к. Я считаю, что назначение слов "вообще говоря" состоит в указании, что последующее предложение истинно всегда и при всех обстоятельствах. Они играют роль квантора всеобщности. Так, предложение вида "Вообще говоря, имеет место $P(x)$" означает, что истинно высказывание $\forall\, xP(x)$.

З а н у д а. Мне кажется, Простак слишком прямолинеен в своих суждениях. Я думаю, что мы говорим "Вообще говоря, …", когда хотим обобщить что-то конкретное или сказать, что данное высказывание истинно в подавляющем большинстве случаев, хотя возможны и исключения. Так, например, высказывание "Вообще говоря, люди хорошие", вполне осмысленное и приемлемое в обычной речи, является ложным, если его формализовать как $\forall\, x(\text{Человек}(x) \Rightarrow \text{Хороший}(x))$, поскольку существуют и плохие люди. Истинно выражение $\overline{\forall x(\text{Человек}(x) \Rightarrow \text{Хороший}(x))}$, равносильное выражению $\exists x\overline{(\text{Человек}(x) \Rightarrow \text{Хороший}(x))}$ — существует человек, который не является хорошим. Однако по своей форме это высказывание совсем не похоже на общеутвердительное высказывание.

П р о ф е с с о р. Вводные слова "вообще говоря" допускают несколько интерпретаций в обычной речи. Но, пожалуй, основной из них является та, при которой следующее за ними предложение оказывается истинным высказыванием. Выражение этого высказывания не обязательно должно содержать квантор всеобщности. Например, правильно было бы сказать "Вообще говоря, не все люди хорошие". Если вы провозгласите фразу "Вообще говоря, все люди хорошие", то скорее всего услышите в ответ какой-нибудь контрпример типа "Но есть и плохие люди, например, преступники". Заметьте, что стоит нам высказаться более или мене общо, как сразу же найдется оппозиционер, который укажет случай, когда наше высказывание неверно. Иногда наши обобщения действительно неправомерны, но чаще всего словами "вообще говоря" мы просто заменяем более точные словосочетания "в большинстве случаев", "как правило" и т. п. Так, фразу "Вообще говоря, все люди хорошие" следует понимать как "В большинстве своем люди хорошие".

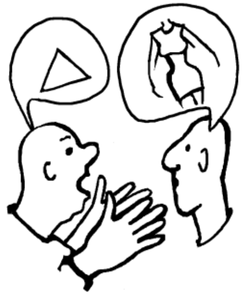
2.3. Определения

П р о ф е с с о р. Логический вывод одних утверждений из других предполагает, что есть некоторые исходные утверждения, называемые определениями и аксиомами, истинность которых либо самоочевидна, либо допускается в качестве правдоподобной гипотезы. Если из принятых гипотез логически вытекают следствия, истинность которых не вызывает сомнения (например, они подтверждены практикой), то мы получаем оправдание принятых допущений и приобретаем некоторую дополнительную уверенность в их истинности. Однако нет никакой гарантии, что однажды наши гипотезы не будут опровергнуты. В практике познания мира данное обстоятельство не вызывает особого удивления или недоумения. Действительно, наши теории со временем корректируются и даже заменяются новыми, более приспо-

собленными к реалиям жизни, но в математике мы не можем этим утешиться. Математика занимается конструкциями нашего разума. Разум творит свои миры, черпая вдохновение и факты как из внешнего, так и из внутреннего мира, созданного при рождении и дополненного в результате соприкосновения с внешним окружением.

Математика может быть полезна, о чем свидетельствует огромное количество наших достижений прежде всего в технической области. Описание движения планет и космических кораблей, предсказания солнечных и лунных затмений, телефония, радиосвязь, телевидение, компьютеры, атомные электростанции и многое другое не стало бы нашим достоянием без математики. Но, кроме того, математика интересна как феномен умственной деятельности человека. Почему разум помогает выжить, завоевать пространства — и в то же время порождает чудовищ в виде противоречий?

Люди, чем бы они ни занимались, всегда испытывают особый интерес к определениям используемых в речи терминов и понятий. Если ожесточенный спор заходит в тупик, то первое, что начинают делать спорщики, искренне желающие прийти к истине, а не просто победить или посрамить противника, — так это уточняют значения используемых терминов. Дело в том, что в разговоре два человека могут вкладывать в одни и те же слова различные значения и смысл. Различия все же могут остаться, даже если собеседники как будто договорились об одинаковом понимании содержания слов и фраз. Поэтому, говоря вообще, "мысль изреченная есть ложь". Однако у нас нет другого способа обмениваться мыслями, кроме как посредством речи и, следовательно, мы должны быть внимательны к употреблению слов и использованию понятий.



Исходные утверждения, как бы они ни были сформулированы, некоторым образом характеризуют объект исследования, выделяя его из всех остальных. При этом желательно, чтобы определение характеризовало только интересующий нас объект и ничто другое. Существует много видов определений. Среди них особую роль играют так называемые *прямые* определения, которые могут быть даны как через ближайший род и видовое отличие, так и путем построения.

В соответствии с принципом определения объекта через ближайший род и видовое отличие следует указать наименьший класс (род), содержащий данный объект, и признаки, по которым данный объект отличается от других объектов данного класса. Например, "Человек это млекопитающее, которое обладает такими-то свойствами [выделяющими его среди других млекопитающих]".

Во многих случаях используется еще один вид определений — через аксиомы или постулаты. Аксиомы (постулаты) констатируют некоторые свойства интересующего нас объекта и принимаются без доказательства. Последнее возможно, если аксиомы либо самоочевидны, либо достаточно оправданны. При введении такого вида определения мы обычно уже знаем, что определяемый объект принадлежит некоторому классу (роду), но для задания специфического различия используем некоторое соотношение или условие, которому должен удовлетворять определяемый объект. Эти соотношения, или условия, и являются аксиомами. Так, величину x можно определить прямо — с помощью некоторого уравнения, например, $x = f(y)$, а можно определить неявно — с помощью уравнения вида $F(x, y) = 0$. Во втором случае мы имеем дело с аксиомой, утверждающей, что определяемый объект x удовлетворяет соотношению $F(x, y) = 0$.

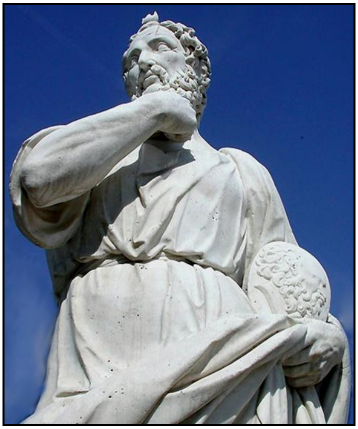
Однако определение через аксиомы будет действительно определением, если удастся доказать, что определяемый объект существует. Вы можете попытаться построить объект, удовлетворяющий всем аксиомам определения. В общем же случае существование объекта означает отсутствие противоречий в аксиомах.

Далее я предлагаю вам рассмотреть некоторые определения, их использование в доказательствах теорем, а также кое-какие важные сопутствующие вопросы.

2.3.1. Определение человека

П р о ф е с с о р. Давным давно Платон, ученик Сократа, дал определение человека, приведшее в восторг его учеников: "Человек это двуногое без перьев". Диоген, оппонент Платона в то время, вылез из своей бочки и принес ощипанного петуха, воскликнув: "Вот человек Платона!" Однако Платон не растерялся, тут же добавив: "… и с плоскими ногтями". Что вы думаете об этом?

П р о с т а к. На мой взгляд, определение Платона с уточнением формы ногтей довольно остроумно и кратко. Мне не приходит на ум никакой контрпример. Похоже, нет других живых существ, отличных от человека, которые были бы двуногими, без перьев, да еще с плоскими ногтями. Впрочем, и неживых тоже.



Платон
(428—328 гг. до н. э.)

Диоген в бочке
(около 400—325 гг. до н. э.)

З а н у д а. Соглашаясь с Простаком, я должен заметить, что определение Платона отделяет человека от других объектов универсума, но не дает представления о его сущности. Если бы на лбу всех людей было написано, например, "гуманоид", то Платон дал бы такое определение: "Человек это тот, кто помечен как "гуманоид"". Правда, в этом случае он не вызвал бы восторга ни у кого, даже у собственных учеников.

П р о с т а к. Это уж чересчур. А если бы на лбу людей было написано "человек", то определение становилось бы еще нелепее: "Человек это тот, кто помечен как "человек"". Что-то вроде масла масляного.

З а н у д а. Доведение идеи до абсурда — неплохой метод выяснения ее сути. Я хотел лишь подчеркнуть, как мало дает информации о человеке определение Платона. Мне кажется, что определение человека должно было бы включать понятия о его разуме, способности передавать мысли, пользоваться орудиями труда и т. д.

Например, христиане должны выполнить обряд крещения новорожденного. Чтобы это стало возможным, необходимо признать, что новорожденный является человеком.

П р о с т а к. Но разве одного факта рождения от человеческой женщины еще не достаточно?

З а н у д а. Как оказалось, нет. Человек, по давным-давно укоренившемуся мнению, это не просто отпрыск человека. Это еще и существо, обладающее душой. Душой обладают только люди, поэтому животных и не крестят. Обладает ли новорожденный душой, сразу не выяснить. Надо подождать, наблюдая за ним в течение нескольких дней. Вопрос становится особенно актуальным, если ребенок имеет врожденные отклонения от принятой нормы, например, шесть пальцев, не имеет ручек или ножек и т. п. Теперь мы, благодаря науке, знаем, что физические недостатки не определяют однозначно умственную недостаточность или отсутствие души. Однако всякий раз, встречаясь с подобным явлением, мы все же задаем себе этот сакраментальный, а может, и крамольный вопрос: человек ли перед нами? Раньше вопрос о корректном определении человека был совсем не праздным. От его решения зависел статус человека в обществе: крещеный или нет. Так что, определить, что такое человек, непросто, но важно.

П р о ф е с с о р. Зануда прав. Еще в XVII веке немец Готфрид Вильгельм Лейбниц и англичанин Джон Локк уделили этому немало внимания. Интересно, что первым написал свою книгу Джон Локк, а Лейбниц написал свою уже после — в качестве ответа Локку. Таким образом, получился заочный диалог между двумя выдающимися мыслителями. И доныне он занимателен и поучителен. Во-первых, речь этого диалога естественна, понятна и красива. Во-вторых, диалог построен так, что мы ощущаем стремление к истине более важным, чем сама истина. Мы видим анализ жизни на лабораторном столе, если позволено мне будет так сказать. Другими словами, мы видим сам процесс рождения истины, даже если в действительности, особенно с учетом нашей современной осведомленности, она так и не родится.

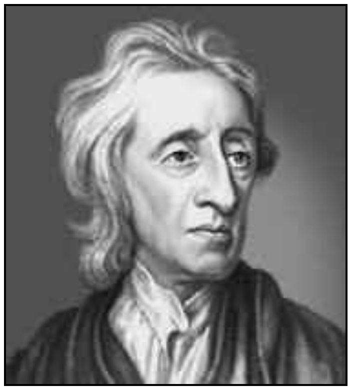
Итак, один и тот же объект может иметь различные определения в зависимости от точки зрения на него. Каждое определение чему-нибудь да служит. Например, определение Платона может служить той же цели, что и индекс УДК или книжный ISBN (систематизация и последующий быстрый поиск книги по ее индексу). Однако вряд ли подобное определение можно использовать в качестве посылки для логического вывода каких-то других свойств, например, человека.

В математике определения стараются выбрать таким образом, чтобы их можно было применить в доказательствах теорем. Теорема, как известно, представляет собой утверждение, имеющее доказательство или, по крайней мере, требующее такового. Теоремам надо от чего-то оттолкнуться или на что-то сослаться. Это что-то должно быть истинным. Самыми первыми истинами являются утверждения, сформулированные в определениях. Другими словами, определения это не только какие-то соглашения об употреблении слов (названий), но и постулирование истинности каких-то утверждений. Далее мы попробуем порассуждать на тему определений.

2.3.2. Доказательство того, что 2 + 2 = 4

П р о ф е с с о р. Как ни странно, обычно труднее всего доказать то, что практически самоочевидно. Чем очевиднее высказывание, тем глубже находятся его содержание и смысл в нашем подсознании и тем труднее вывести их на поверхность сознания, где возможно использование речи и логики. Иногда совершенно очевидные всем свойства и соотношения принимают в качестве истин (аксиом), не требующих доказательства, а другие суждения выводят из них логическим путем. Доказать логически, что $2 + 2 = 4$, пытался еще великий немецкий философ и математик Готфрид Вильгельм Лейбниц, имевший случай побеседовать с великими мира сего, в том числе, с русским царем Петром I.

Готфрид Лейбниц (1646—1716)



Джон Локк (1632—1704)

Не важно, что равенство $2 + 2 = 4$ общеизвестно. Нас интересует, как прийти к нему логическим путем, исходя из более простых с точки зрения интуиции вещей, таких как число 1 и операция прибавления 1 к заданному целому числу.

Предположим, что каким-то образом определены число 1 и операция $x + 1$ прибавления числа 1 к заданному числу x.

Определим конкретные числа 2, 3 и 4 посредством следующих равенств:

$$1 + 1 = 2; \qquad (1)$$

$$2 + 1 = 3; \qquad (2)$$

$$3 + 1 = 4. \qquad (3)$$

Определим операцию прибавления $x + 2$ следующим равенством:

$$x + 2 = (x + 1) + 1, \tag{4}$$

где x — любое целое число из определенных ранее, т. е. 2, 3 или 4.

Введя данные определения, выполним несложные преобразования:

- по определению (4) имеем:

  $2 + 2 = (2 + 1) + 1$
- по определению (2) имеем:

  $(2 + 1) + 1 = 3 + 1$
- по определению (3) имеем:

  $3 + 1 = 4$
- В результате получаем равенство:

  $2 + 2 = 4$

Именно это нам и требовалось доказать.

Зануда. Уважаемый Профессор, мне кажется, что в Вашем доказательстве не заданное явным образом определение числа 1 и операции прибавления $x + 1$ не играет никакой роли, даже в определении (4), поскольку $(x + 1)$ в правой части равенства есть просто форма, в которую можно подставить только ранее определенные числа 2, 3 и 4.

Профессор. Ваша наблюдательность, Зануда, выше всяких похвал! Действительно, определения числа 1 и операции $x + 1$ прибавления 1 к любому числу x нигде не используются в последующих рассуждениях. В том-то все и дело! Проведенное мною доказательство большинство математиков не принимают в качестве такового. Это просто элементарное приведение частных формул к заданному определению, или проверка того, что частное утверждение (в данном случае $2 + 2 = 4$) удовлетворяет принятым определениям. Так, используя подстановочную формулу (4) и определения (1), (2), (3), мы выполняли дозволенные определениями преобразования выражения $2 + 2$, которые в конце концов привели к требуемому равенству $2 + 2 = 4$.

Другое дело — доказательство не частного, а общего высказывания. Например, доказательство того, что $x + a$ равно тому числу, которое получается при сложении чисел x и a. Далее мы рассмотрим пример определения, позволяющего сделать это.

2.3.3. Определения сложения и умножения целых чисел

Профессор. Рано или поздно дети научаются считать. Их родители, а затем и профессиональные учителя, прикладывают к этому некоторые усилия. В процессе обучения арифметике и учитель, и ученик обычно заинтересованы лишь в конечном результате — научиться перечислять и сравнивать объекты в их количественном отношении. При этом внутренняя логика, психологизм и интуитивные основы процесса обучения нередко остаются без внимания. Давайте рассмотрим, как можно определить операции сложения и умножения целых чисел.

Сначала займемся сложением. Предположим, что операция $x + 1$ прибавления 1 к числу x уже как-то определена. Каким бы ни было это определение, оно не будет играть никакой роли в последующих рассуждениях. Теперь требуется определить операцию $x + a$, заключающуюся в прибавлении числа a к числу x.

Допустим, что определена операция

$$x + (a - 1).$$

Тогда операция x + a будет определена равенством

$$x + a = (x + (a - 1)) + 1.$$

Зануда. Таким образом, мы знаем, что такое $x + a$, если знаем, что такое $x + (a - 1)$, а последнее мы знаем, если известно, что такое $x + (a - 2)$ и так далее, вплоть до $x + 1$, что известно нам по определению.

Профессор. Совершенно верно, последовательными рекуррентными (т. е. возвращающимися к началу) операциями можно определить операции $x + 2$, $x + 3$, $x + 4$ и т. д. Обратите внимание, равенство

$$x + a = (x + (a - 1)) + 1$$

содержит в себе бесконечное количество различных определений, каждое из которых имеет смысл только тогда, когда известно предшествующее ему определение.

С помощью данного определения можно доказать важные свойства операции сложения:

- $a + (b + c) = (a + b) + c$ — ассоциативность;
- $a + b = b + a$ — коммутативность.

Не приводя доказательства данных равенств, отмечу лишь, что определение операции $x + a$ сформулировано так, что позволяет это сделать.

Аналогичным образом можно определить операцию умножения:

$$a \times 1 = a,$$

$$a \times b = (a \times (b - 1)) + a.$$

Второе из этих равенств, как и в определении сложения, содержит в себе бесконечное количество определений. Поскольку $a \times 1$ определено, можно определить $a \times 2$, $a \times 3$ и т. д.

Используя определение операции умножения, можно доказать ее свойства:

- $(a + b) \times c = (a \times c) + (b \times c)$ — дистрибутивность;
- $a \times b = b \times a$ — коммутативность.

2.3.4. Непредикативные определения

Профессор. В любом естественном языке имеются прилагательные — слова, выражающие свойства объектов, например, красный, мягкий, сладкий, большой и т. п. Но прилагательное как слово языка может обладать или не обладать тем свойством, которое оно описывает, например: прилагательное "русский" является русским, а прилагательное "английский" является русским, а не английским; прилагательное "многосложный" является многосложным, а прилагательное "односложный" является многосложным, а не односложным. Иначе говоря, прилагательное может быть применимо или неприменимо к самому себе.

Дадим следующее определение указанным разновидностям прилагательных:

> Прилагательные, применимые к самим себе, назовем ***автологическими***, а прилагательные, неприменимые к самим себе, — ***гетерологическими***.

Вопрос: является ли прилагательное "гетерологический" гетерологическим?

Зануда. Допустим, что прилагательное "гетерологический" гетерологично. Тогда по определению оно неприменимо к самому себе, т. е. утверждение "прилагательное *гетерологический* гетерологично" ложно. Следовательно, прилагательное "гетерологический" автологично, что противоречит допущению.

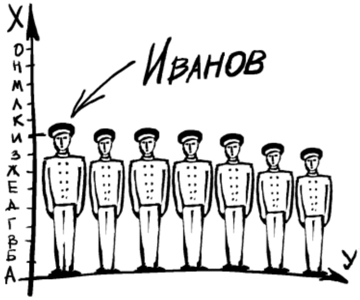
Простак. Допустим, наоборот, прилагательное "гетерологический" автологично. Тогда по определению оно применимо к самому себе, т. е. вопреки допущению верно утверждение "прилагательное *гетерологический* гетерологично". Но что же тогда получается? Допустив, что прилагательное "гетерологический" гетерологично, мы приходим к выводу, что оно, напротив, автологично. Предположив обратное, что прилагательное "гетерологический" автологично, мы приходим к тому, что оно гетерологично. Очередной парадокс!

З а н у д а. Выходит, что какое бы допущение мы ни приняли, оно приведет нас к противоречию.

П р о ф е с с о р. Данный парадокс был впервые сформулирован в 1908 г. К. Греллингом и Л. Нельсоном. В символической форме его можно записать так: x гетерологичен, если x есть "не x". Причина данного парадокса заключается в том, что рассматриваются две одноименные вещи: понятие "гетерологический" и конкретное прилагательное "гетерологический". Попытка применить рассмотренное нами определение к конкретному прилагательному есть не что иное, как попытка описать его посредством самого себя. Понятие имеет некоторый объем — класс всех тех объектов, которые соответствуют данному понятию (подпадают под него, являются его частными случаями или примерами). Определение гетерологических прилагательных задавало класс конкретных прилагательных, которому могло принадлежать или не принадлежать прилагательное "гетерологический". Давая определение, мы не задумывались об этом, поскольку предполагали, что если прилагательное "гетерологический" не принадлежит классу гетерологических, то принадлежит классу автологических. Как бы то ни было, определение неявно описывало "гетерологический" через "гетерологический" и тем самым создавало предпосылку блуждания по порочному кругу.

Определения, описывающие объекты через самих себя, выдающийся французский математик А. Пуанкаре назвал *непредикативными*. Непредикативные определения могут, хотя и не обязательно, привести к противоречию. Вместе с тем, непредикативные определения используются очень широко. Вот примеры непредикативных и непротиворечивых определений: "Иванов — лучший ученик в своем классе", "Это предложение — короткое", "Число 3 — максимальное среди чисел 1, 2, 3", "Иванов — рядовой 2-го взвода 1-й мотострелковой роты 216-го отдельного батальона". Так, в последнем определении солдат Иванов определяется через взвод, который в конечном счете определяется списком его личного состава, содержащего среди прочих и Иванова.

Однако в настоящее время у нас нет критерия, с помощью которого можно было бы распознать, приводит данное определение к противоречию или не приводит. Поскольку всегда существует опасность, что вновь введенное

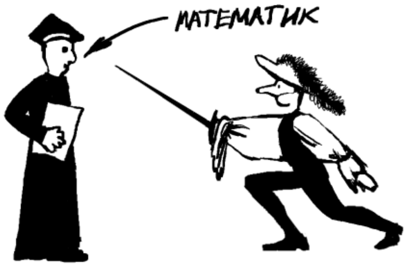


непредикативное определение противоречиво, А. Пуанкаре предложил избегать в математике непредикативных определений. Но в уже существующем здании математики их имеется огромное множество, например, определение максимума функции на заданном числовом интервале: максимальное значение функции на заданном интервале — наибольшее из значений, принимаемых этой функцией на заданном интервале. Здесь максимальное значение определяется через класс, содержащий максимальное значение. Выдающийся немецкий математик Г. Вейль безуспешно пытался переформулировать некоторые определения так, чтобы избавиться от их непредикативности.

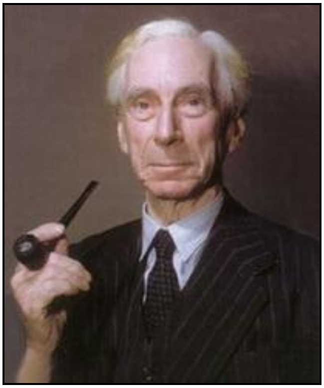
Анри Пуанкаре (1854—1912)

Герман Вейль (1885—1955)

П р о с т а к. Мне кажется, что можно переформулировать определение гетерологических и автологических прилагательных так, чтобы оно стало предикативным. Для этого к нему достаточно добавить, что оно не относится к прилагательным "гетерологический" и "автологический". Другими словами, следует особо оговорить исключения.



П р о ф е с с о р. Что ж, я лично согласен с Вами. Однако замечу, что не все математики приемлют такой подход. В частности, выдающийся английский философ, математик и логик Б. Рассел по данному поводу выразился следующим образом: "Всякий волен в беседе с человеком, у которого длинный нос, заметить: "Говоря о носах, я отнюдь не имею в виду слишком длинные носы". Вряд ли, однако, такое замечание можно считать успешной попыткой обойти больной вопрос".



Бертран Рассел (1872—1970)

А вот еще один пример. Широко известно высказывание "Каждое правило имеет исключения". Данное высказывание само по себе является правилом и, следовательно, для него можно найти по крайней мере одно исключение. Однако это означает, что существует правило без каких бы то ни было исключений. Подобные высказывания содержат ссылку на самих себя и при этом отрицают самих себя. Данное противоречие отвергалось некоторыми логиками на том основании, что само высказывание объявлялось лишенным смысла. Ведь существуют грамматически правильные предложения, поясняли они, лишенные смысла, т. е. ложные, как, например, высказывание "Это предложение состоит из ста слов".

Тем не менее, произнося фразу "Каждое правило имеет исключения", мы должны подразумевать, во избежание противоречий, что она сама не относится к тем правилам, о которых в ней говорится. Другими словами, это не просто правило, а *метаправило* (правило о правилах).

Наконец, я предлагаю вам рассмотреть еще одно непредикативное определение, ведущее к противоречию. Оно было предложено Б. Расселом в 1918 г. в виде следующей задачи, получившей название "Парадокс брадобрея":

Один брадобрей объявил, что он бреет только всех тех, кто не бреется сам.

Разумеется, брадобрей не бреет тех людей, которые бреются сами. Казалось бы, здесь нет ничего особенного. Описанная ситуация банальна, но до тех пор, пока не возник вопрос: должен ли брадобрей брить самого себя? Если он не бреется сам, то должен себя брить. Если же он бреется сам, то не должен себя брить. Таким образом, налицо противоречие, подобное тем, которые мы рассмотрели ранее. В отличие от предыдущих разборов, давайте более тщательно проанализируем данную ситуацию, используя для ясности и краткости изложения некоторую несложную символику.

Итак, мы имеем следующее определение брадобрея: брадобрей — это человек, который бреет только тех людей ("человеков"), которые не бреются сами. Уточним данное определение, введя следующие обозначения:

- x — брадобрей, он же является человеком;
- y — человек;
- x бреет y тогда и только тогда, когда y не бреет y.

Очевидно, что данное определение равносильно такому:

> x бреет y тогда и только тогда, когда неверно, что y бреет y.

Здесь логическая связка "тогда и только тогда" (<=>) предполагает, что выполняются следующие два высказывания:

- если x бреет y, то y не бреет y;
- если y не бреет y, то x бреет y.

Примечание

Напомню, что высказывание $A \Leftrightarrow B$ равносильно двум высказываниям: $A \Rightarrow B$ и $B \Rightarrow A$.

Вопрос заключается в том, истинно ли высказывание "x бреет x"? Для ответа на него достаточно в только что упомянутых двух высказываниях заменить y (человека) на x (человека, являющимся брадобреем), чтобы получить следующее:

- если x бреет x, то x не бреет x;
- если x не бреет x, то x бреет x.

Теперь противоречие очевидно! Брадобрей был определен через класс людей, который содержит в качестве своего элемента брадобрея. А это и есть непредикативное определение, приведшее в данном случае к противоречию.

Простак. Я хотел бы попробовать переформулировать определение брадобрея, чтобы избавиться от противоречивости. Брадобрей не должен рассматривать себя в качестве своих возможных клиентов. Если x — брадобрей, а y — клиент брадобрея, то необходимо потребовать, чтобы $x \neq y$. Не вводя в обиход понятие клиента, брадобрея, обозначенного через x, можно определить так:

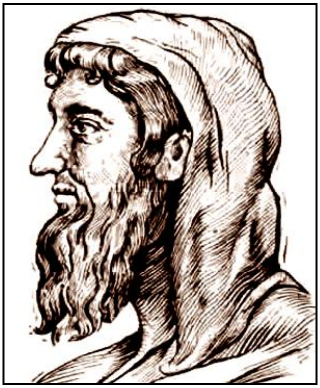
x бреет y тогда и только тогда, когда y не бреет y и $x \neq y$.

Профессор. Что ж, на мой взгляд, это приемлемый способ разрешить данное противоречие. Я надеюсь, его анализ показал вам, что и рассмотренные ранее противоречия имеют сходную природу. Существуют и другие способы их разрешения, но о них мы поговорим позже *(см. разд. 3.3)*.

2.3.5. Аксиоматические определения

Профессор. В математике нередко объект изучения определяется посредством одного или нескольких утверждений, выражающих его свойства и/или отношения с другими объектами. Эти утверждения, принимаемые как истинные без доказательства, называются аксиомами. Также довольно часто используют термин "постулат". По существу это тоже аксиома. Однако иногда между постулатом и аксиомой вводят некоторое тонкое различие. Так, *аксиомами* называют утверждения об объекте, истинность которых самоочевидна. От *постулата* же не требуется самоочевидность, он вводится из каких-то требований нашего разума, для обеспечения некоторого технологического удобства. Хотя постулат не обязан быть самоочевидным, он должен быть достаточно обоснованным. Так, например, независимость скорости света от скорости его источника совсем не очевидна, но она постулируется в специальной теории относительности, без чего было бы невозможно объяснить ряд физических эффектов. Указанное различие между аксиомами и постулатами для нас не важно. Важно другое — изучаемый объект это то, что удовлетворяет аксиомам. Старейшей аксиоматической системой является система аксиом

геометрии Евклида. Она устанавливает отношения между точками, прямыми и плоскостями.



Евклид (III в. до н. э.)

Зануда. Но может случиться, что одним и тем же аксиомам удовлетворяют несколько различных объектов, и наоборот, ни один из объектов не удовлетворяет хотя бы одной из аксиом.

Профессор. Если несколько различных объектов удовлетворяют всем аксиомам, то последние определяют класс объектов. Иногда аксиомы, придуманные в расчете на какие-то более или менее отчетливые интуитивно понятные объекты, оказываются применимыми и к совершенно другим образам. Тем и хороши аксиоматические определения, что они определяют не только тот предмет, на котором сосредоточено внимание их автора. А если не найдется ни одного объекта, отвечающего всем аксиомам, то этому может быть несколько причин. Во-первых, Вы просто плохо искали, и эту причину мы не будем обсуждать. Во-вторых, система аксиом может быть противоречивой, и тогда удовлетворяющий ей объект просто не существует — его нельзя ни найти, ни построить. Таким образом, существование объекта, определяемого системой аксиом, понимается как непротиворечивость последней.

Простак. А как узнать, противоречива или нет данная система аксиом?

Профессор. Установить непротиворечивость аксиом не всегда легко. Прежде всего, выясняют, совместимы ли аксиомы между собой.

Две аксиомы несовместимы, если принятие истинности одной из них требует признать ложность другой; в противном случае аксиомы совместимы. Далее, если из аксиом посредством логического вывода мы получаем некоторое высказывание A, а также его отрицание $\overline{A}$, то данная система аксиом противоречива. Это очень плохо, поскольку из противоречивой системы аксиом можно вывести что угодно — как истинные, так и ложные высказывания. К чему нам такие аксиомы? Они ничего не определяют или, наоборот, определяют все. Определение всего без каких бы то ни было исключений ничего не определяет.

Простак. А если некоторые аксиомы можно логически вывести из других аксиом той же системы?

Профессор. В этом случае говорят, что аксиомы не являются независимыми и возможно сокращение их количества. Удаленные из системы аксиомы становятся теоремами, т. е. утверждениями, выводимыми из остальных аксиом. Нередко бывает, что математик, опираясь на свой опыт и интуицию, формулирует некоторую теорему, надеясь на то, что она окажется истинной. Далее он должен доказать ее, опираясь на принятые аксиомы. Однако не всегда это удается. Например, знаменитые теоремы Ферма и Гольдбаха до сих пор остаются лишь гипотезами, не получившими в математике ни подтверждения, ни опровержения.

Простак. Я предчувствую возможные трудности создания аксиоматической системы для чего-либо. При этом возникает соблазн ввести какую-нибудь аксиому, которая пусть даже не очевидна, но нужна для каких-то "технологических потребностей", в качестве строительных лесов, например, для легкого вывода заведомо правильной теоремы или во избежание противоречий. Другими словами, в аксиомах мы можем потребовать не только то, что самоочевидно, но и то, что нам потребуется в дальнейшем, несмотря на то, что эти утверждения не только не самоочевидны, но и могут на первый взгляд просто противоречить нашей интуиции.

Профессор. Затронутая Вами тема очень важна и привлекает внимание многих математиков и философов. По данному поводу я ограничусь лишь тем, что процитирую Б. Рассела (книга "Введение в математическую философию", 1919 г.): "Метод постулирования того, что нам требуется, обладает многими преимуществами, но такими же преимуществами обладает воровство перед честным трудом".

Итак, в качестве примера аксиоматического определения я предлагаю вам рассмотреть определение натуральных чисел.

П р о с т а к. Натуральные числа это 1, 2, 3, 4 и т. д. Иначе говоря, это целые положительные числа, наименьшим из которых является 1. Они известны нам с начальных классов школы и не требуют какого бы то ни было специального определения. Но если Вы настаиваете, то я попытаюсь их определить:

1. 1 есть натуральное число.
2. Если x — натуральное число, то $x + 1$ — тоже натуральное число.

Начиная с числа 1, я могу получить любое натуральное число. Правда, нужно еще определить операцию прибавления 1 к числу (думаю, это несложно) и добавить соответствующее определение в качестве третьей аксиомы. По крайней мере, мое определение дает достаточно отчетливое представление о натуральных числах.

П р о ф е с с о р. Допустим, так. Но как вы будете сравнивать натуральные числа? В частности, как вы устано-вите, что какие-нибудь два числа равны? Ведь Ваши аксиомы не определяют равенство двух чисел.

П р о с т а к. Почему нет? Допустим, с помощью своего определения я получил два натуральных числа. Если они окажутся внешне одинаковыми (например, 5 и 5), то они будут равными. Впрочем, каждое натуральное число определяется своим положением в ряду всех натуральных чисел. То, что натуральные числа однозначно упорядочены, является одним из главных свойств всей совокупности натуральных чисел. Может быть, определяя натуральное число, не следует прибегать к определению операции прибавления 1, а прямо воспользоваться представлением об упорядоченности натуральных чисел? Как Вы считаете, Профессор?

П р о ф е с с о р. Мне нравится ход Ваших мыслей. Я не буду разбирать их по отдельности, а вместо этого предложу рассмотреть аксиомы для натуральных чисел, предложенные итальянским математиком Д. Пеано:

1. 1 есть натуральное число.
2. 1 не следует ни за каким натуральным числом.
3. Следующее за натуральным числом есть натуральное число; всякое натуральное число имеет следующее за ним натуральное число.
4. Два натуральных числа равны, если равны следующие за ними натуральные числа.
5. Аксиома полной индукции.

Джузеппе Пеано (1858—1932)

Как видите, данное аксиоматическое определение базируется на понятии "следует", т. е. на интуитивно ясном понятии упорядоченности. Из первых трех аксиом вытекает, что у каждого натурального числа, кроме 1, есть предшественник и у каждого натурального числа, без исключений, есть последователь. Таким образом, утверждается, что натуральные числа упорядочены. Четвертая аксиома устанавливает, в каком случае два натуральных числа равны. А именно: они равны, если занимают в своем упорядоченном ряду одинаковые места. Заметьте, при сравнении чисел не требуется устанавливать идентичность их символических обозначений.

Особую роль играет пятая аксиома — аксиома полной, или математической, индукции. Эта аксиома постулирует применимость одного из важнейших математических принципов или методов доказательства. Я поясню ее содержание.

Предположим, что имеется бесконечная упорядоченная последовательность высказываний, пронумерованных натуральными числами:

$A_1, A_2, \ldots, A_n, A_{n+1}, \ldots$

Допустим, что:

1. Установлена верность высказывания A_1 (это — база индукции).
2. Для произвольного $n > 1$ доказано, что если верно высказывание A_n, то верно и высказывание A_{n+1} (это — индуктивный переход).

Тогда верны все высказывания бесконечной последовательности A_1, A_2, …, A_n, A_{n+1}, …

Совокупность всех натуральных чисел бесконечна. Действительно, какое бы натуральное число мы ни взяли, существует по крайней мере еще одно натуральное число, следующее за данным. Поэтому чтобы делать какие либо утверждения о любых натуральных числах, необходим метод доказательства, позволяющий оперировать с бесконечными совокупностями. Такой метод базируется на принципе полной индукции.

П р о с т а к. Я вспоминаю, что о математической индукции нам что-то говорили в школе. Не могли бы Вы привести какой-нибудь пример использования принципа математической индукции?

П р о ф е с с о р. Хорошо, рассмотрим пример, который, если не ошибаюсь, обычно приводится в школах.

Требуется доказать, что для любых натурального числа n и вещественного числа x, не равного 1, выполняется следующее равенство:

$$1 + x + x^2 + ... + x^n = \frac{1 - x^{n+1}}{1 - x}.$$

Это равенство определяет сумму членов геометрической прогрессии, т. е. последовательности чисел, в которой частное от деления двух соседних чисел одинаково для любой пары соседних членов последовательности. Нетрудно заметить, что при $x = 1$ правая часть рассматриваемого равенства становится неопределенной, поэтому данный частный случай предусмотрительно и был исключен из общей формулировки задачи.

Докажем это равенство для $n = 1$ и произвольного x, не равного 1:

$$1 + x = \frac{(1 + x)(1 - x)}{1 - x} = \frac{1 - x^2}{1 - x} = \frac{1 - x^{1+1}}{1 - x}.$$

Допустим, что исходное равенство, которое требуется доказать, верно для некоторого $n > 1$, и докажем его для $n + 1$. С этой целью прибавим к обеим частям данного равенства x^{n+1}:

$$1 + x + x^2 + ... + x^n + x^{n+1} = \frac{1 - x^{n+1}}{1 - x} + x^{n+1} =$$

$$\frac{1 - x^{n+1} + x^{n+1} - x^{(n+1)+1}}{1 - x} = \frac{1 - x^{(n+1)+1}}{1 - x},$$

откуда следует, что исходное равенство выполняется для любого натурального n.

Теперь вернемся к разговору об аксиомах. Как я уже говорил, от аксиоматического определения требуется, чтобы оно было непротиворечивым. Для этого достаточно подобрать или построить хотя бы один пример объекта, удовлетворяющего всем аксиомам. В нашем случае нам надо проверить выполнимость аксиом хотя бы для двух натуральных чисел, поскольку четвертая аксиома есть утверждение о двух числах. Но возможно ли это?

Возьмем числа, например, 1 и 2. Эта совокупность, как легко видеть, удовлетворяет аксиомам 1, 2, 4 и 5. Однако для того, чтобы она удовлетворяла аксиоме 3, необходимо еще одно натуральное число, а именно число 3. При этом необходимо, чтобы совокупность чисел 1, 2, 3 удовлетворяла всем пяти аксиомам. Но третья аксиома требует, чтобы число 4 было целым числом и чтобы совокупность 1, 2, 3, 4 удовлетворяла всем пяти аксиомам. Так можно рассуждать сколь угодно долго. Следовательно, нет возможности доказать аксиомы для нескольких чисел, не доказывая их для всех, а значит приходится отказаться от доказательства путем предъявления примера. Остается собрать все логические следствия из данных аксиом и проверить, не заключают ли они в себе противоречия. Это было бы возможно, если бы количество следствий было конечным, но это, к сожалению, не так: они охватывают если не всю математику, то по крайней мере всю арифметику.

Выяснение непротиворечивости арифметики — один из фундаментальных вопросов обоснования математики в целом. Мы не будем на нем останавливаться. Я привел один из вариантов определения натуральных чисел лишь для иллюстрации идеи аксиоматического определения вообще.

Любая аксиоматическая теория использует неопределяемые понятия. Так, например, аксиомы геометрии Евклида используют в своих формулировках неопределяемые понятия точек, прямых и плоскостей. При этом не важно, что именно вы имеете в виду, используя термины "точка", "прямая" и "плоскость", хотя мы и знаем, откуда они взялись. Математическое понятие точки происходит от зрительно воспринимаемого пятна, оставленного краской на бумаге. Разум же образует понятие точки, лишая пятно таких качеств, как протяженность, цвет и др. Аналогично, понятие линии (в том числе, прямой) имеет в основе своего появления след, оставленный карандашом на бумаге или палкой на песке; этот след имеет ширину, как и любая полоса.

Математическая линия это — полоса без ширины. Аксиомы определяют отношения между неопределяемыми объектами. Например, "Между двумя точками можно провести только одну прямую", "Через данную точку можно провести лишь одну прямую, параллельную данной прямой" и т. д. Аксиомы Пеано, определяющие натуральные числа, опираются на неопределяемое понятие натурального числа и отношение следования одного числа за другим.

Смысл неопределяемых понятий уточняется, становится более определенным и, следовательно, понятным по мере развертывания теории путем логического вывода теорем из аксиом и ранее доказанных теорем. Вместе с тем, наряду с аксиомами, в дополнение к ним в качестве комментариев дают и псевдоопределения неопределяемым понятиям. Например, мы могли бы сказать, что точка — это пятно без протяженности, а линия — полоса нулевой ширины. Такого рода определения предназначены для стимулирования интуиции обучаемых, для создания чувственного образа того, с чем придется иметь дело. Это лишь подспорье для изобретения теорем, которые потом придется строго доказать или опровергнуть, подспорье для поиска путей применения полученных результатов на практике.

Теперь рассмотрим аксиоматические непредикативные определения. О непредикативных определениях мы уже говорили ранее, отмечая, что они могут привести к противоречию. В рассматриваемом сейчас случае аксиомы выражают некоторое отношение между определяемым объектом и всеми членами класса (рода), которому принадлежит определяемый объект. Данная ситуация возникает, когда принимаются следующие две аксиомы:

- определяемый объект X таким-то образом связан со всеми членами класса K;
- объект X входит в состав класса K.

Аналогичная ситуация возникает и при установлении следующих трех аксиом:

- определяемый объект X таким-то образом связан со всеми членами класса K;
- объект Y таким-то образом связан с объектом X;
- объект Y входит в состав класса K.

Данные определения объекта X можно рассмотреть с двух точек зрения. Согласно первой точке зрения (назовем ее конструктивистской), подобные определения вовлекают нас в порочный круг. Так, нельзя определить объект X, не зная всех членов класса K и, значит, не зная самого X, который является одним из таких членов. Согласно второй точке зрения, класс K дан нам целиком и, следовательно, мы уже знаем все его члены, а цель определения объекта X — лишь выделить тот объект, который находится с остальными в указанной связи. Приверженцы конструктивистской точки зрения не соглашаются с этим. Они говорят, что знание класса еще не есть знание всех его членов. Знание класса дает только возможность построить столько его членов, сколько мы пожелаем. Эти члены начинают существовать после того, как будут построены, но не ранее. Таким образом, объект X существует только после его определения, имеющего смысл только в том случае, когда мы заранее знаем все члены класса K, в том числе, и объект X. Иначе говоря, определять X через его связь с X означает бродить по порочному кругу. Тем не менее, подобные определения широко применяются. В подобных ситуациях необходимо предварительно доказать, что аксиома, устанавливающая связь X с членами класса K, не содержит в себе противоречия. Прежде всего доказывают, что каким бы ни был класс K, все члены которого предполагаются известными, существует некий объект X, связанный с членами этого класса данной зависимостью. Тем самым показывают, что существование X не влечет за собой противоречия. Затем доказывают, что нет противоречий между существованием X и предположением, что этот X сам входит в состав класса K. Однако исследования такого рода проводят не всегда.

2.4. О "женской" логике

Профессор. Вы, должно быть, слышали о противопоставлении мужской и женской логик. На мой взгляд, стоит проанализировать данную тему — не для смеха, а чтобы понять (не столько женщин, сколько самих себя, мужчин).

Простак. Действительно, женщины порой ставят нас в тупик, просто обескураживают своими "логическими рассуждениями".

Зануда. Если что-то кажется нелогичным или непонятным, то проще всего отнести это на счет говорящего, а не самого себя. Парадигма нашего обыденного мышления — проблемы и глупости создают другие, а не понимают нас недоумки. Все мы — и мальчики, и девочки, — несем в себе "Янь" и "Инь", но в разных пропорциях.

Профессор. Давайте не будем искать объяснение различиям между мальчиками и девочками в физиологии и половом поведении.

Простак. Но почему нет? Ведь, скажем, проводят отдельные спортивные состязания для мужчин и для женщин — даже в шахматах, где конструктивные особенности тел не играют никакой роли. Стало быть, мужской интеллект отличается от женского.

Профессор. Даже у народов и рас есть различия в миропонимании и логике. Наблюдаемые различия не подлежат оценке по шкале "плохо/хорошо". Так устроен мир, и если вы заметили какую-то разницу, то просто учитывайте ее, общаясь с людьми разных национальностей и полов. Насмешливо-пренебрежительное отношение к женской логике — обычная реакция обывателя при встрече с чем-то чуждым и недоступным. Замечу, что так называемой женской логике посвятил свои вполне серьезные и интересные исследования профессор математики Д. В. Беклемишев. Он обратил внимание на то, что женская логика пока еще ждет своего Аристотеля, который формализовал бы ее подобно обычной (мужской) логике, а также на то, что она наиболее эффективно применяется в спорах, причем конкретных, а не отвлеченных (философских). Логика спора, как и логика вопросов, — особая логика, в формализации которой пока мало фундаментальных результатов. Замечу также, что и мужчины часто спорят, переходя от логических доводов к банальной драке. Я хотел сказать, что порой и мужчины не могут разрешить спор с помощью своей же хваленой логики.

Мужская логика применяется в спорах так же, как и в отвлеченных рассуждениях. Как рассматривается и разрешается спор согласно мужской логике? Сам спор обычно возникает из-за того, что двое, исходя из общих предпосылок, приходят к различным выводам. При этом оба спорщика одинаково и однозначно понимают и применяют правила вывода следствий. Таким образом, обоим заранее известно, что только один из них прав, а другой — нет, поскольку сделал какую-то логическую ошибку. Спор как раз и затевается ради того, чтобы обнаружить эту ошибку, и уже не важно, кто именно из спорщиков ее допустил, — истина важнее. Мужчины очень уважают то, что из чего-нибудь следует, иногда им даже не важно, из чего именно.

С незапамятных времен умение вести спор не столько ради выяснения истины, сколько с целью просто победить противника, настоять на своем, является искусством. И чтобы овладеть им, следует, как заметил Д. В. Беклемишев, усвоить женскую логику. Я попытаюсь изложить несколько выявленных им правил этой логики.

- *Утверждение, оставшееся без возражения, является доказанным.* В споре одна из сторон нередко приводит подряд несколько доводов в пользу своей точки зрения. Обычно представляется, что чем больше частных аргументов, тем убедительнее доказательство в целом. Если противнику нечем возразить, то он проиграл спор. Однако ему совсем не обязательно опровергать все ваши аргументы в каскаде, достаточно опровергнуть лишь последний, чтобы разбить всю вашу систему доказательств. Если в этот момент вы замолчите, то последнее слово, а значит, и поле боя, останется за оппонентом. Отсюда, в частности, следует, что не нужно приводить в одной фразе слишком много доводов, иными словами: не используйте в споре длинные фразы. В споре с женщиной лучше не применять формулу "как я сказал ранее", поскольку ее внимание сосредоточено лишь на последнем вашем доводе. Исключением из данного правила является случай, когда вы можете перед своей тирадой заставить оппонента замолчать ("Помолчи, когда я говорю!"). Интересно также, что если закончить свою последовательность доводов оскорблением, то противник начнет отвечать именно на оскорбление, оставив ваши доводы без ответа, а значит, признав их доказанными. Спор продолжится, но в другой плоскости (ругань или драка). Остыв, спорщики вспомнят, чей довод (не оскорбление) остался без ответа и кто поэтому оказался самым убедительным.

- *Каждое высказывание может быть не только опровергнуто, но и отвергнуто.* Если в обычной логике любое утверждение может быть либо истинным, либо ложным, то в женской оно может быть еще и бессмысленным, не заслуживающим внимания. Высказывания, оцененные как бессмысленные, отвергаются, т. е. не рассматриваются как довод. Если вы отвергли последнее высказывание оппонентки, то предпоследнее в вашей цепочке утверждение оказывается без ответа и тем самым считается доказанным. Чтобы отвергнуть высказывание, достаточно сказать "Ну и что?" или "Ну ты даешь!".
- *Опровержение примером.* В обычной логике общие суждения нельзя доказать никаким числом подтверждающих примеров, но можно опровергнуть одним противоречащим примером. В женской логике это правило не действует. Так, если один пример еще не доказывает общее суждение, то двух примеров вполне достаточно. С другой стороны, один противоречащий пример не способен ничего опровергнуть, поскольку он только один.
- *Отсутствие силлогизмов.* Силлогизмы, как нам уже известно, это формулы, выражающие правила логического вывода следствий из посылок, т. е. истинных утверждений. Основные силлогизмы были разработаны еще Аристотелем. В женской логике силлогизмы почти не используются, потому что в споре они малоэффективны. Женщины не хуже мужчин могут их применять, но в споре каждая из них знает, что из двух посылок может сделать тот же вывод и ее оппонентка. Вопрос в том, устраивает ли оппонентку сделанный вывод сам по себе. Если нет, то оппонентка отвергнет одну или даже все посылки силлогизма. Поскольку такой подход известен обеим спорщицам, они и не применяют силлогизмы. Если же силлогизм применяет мужчина, то женщина чаще всего лишь условно допускает истинность посылок, говоря "ладно, допустим, что это так". Если же окажется, что следствие данного силлогизма возмутительно, то она тут же отбросит временно принятое допущение. Таким образом, применять силлогизмы в разговоре с дамами бесполезно. Отношение женщин к посылкам силлогизмов математик А. Н. Колмогоров сформулировал в виде следующего "закона":

Если из высказывания *A* следует высказывание *B*, и *B* приятно, то *A* истинно.

Глава 3

Немного о множествах

> Ясность в математике приобретается за счет обсуждения классов, а не отдельных свойств …, поскольку, когда мы говорим о классах, у нас имеется четкое представление об их тождестве и различии.
>
> *У. Куайн*

3.1. Что такое множество?

Профессор. Понятие множества является базовым понятием современной математики и имеет столь общий характер, что его невозможно выразить через другие более простые понятия. Вместе с тем, в его основе лежит некий образ, который выдающийся немецкий математик Георг Кантор, основоположник теории множеств, описал как *собрание определенных и различных между собой объектов интуиции или интеллекта, мыслимое как единое целое. Эти объекты называются элементами множества*.

Георг Кантор (1845—1918)

В канторовском определении понятие множества выражается через свой синоним "собрание". Таким образом, данное определение непредикативно *(см. разд. 2.3.4)*. Однако оно выражает собой некую неформальную идейную основу понятия множества, которую мы сейчас рассмотрим более подробно.

Прежде всего, предполагается существование некоторой универсальной совокупности объектов, из которой можно хотя бы мысленно выбирать объекты, чтобы затем составлять из них множества или рассматривать, принадлежат ли они данному множеству или нет. На природу объектов не накладывается никаких ограничений (в частности, объекты сами могут быть множествами), за исключением того, что объекты должны быть определенными и различными. Кроме того, множество объектов само может рассматриваться как единый объект и в качестве такового входить в состав какого-нибудь другого множества.

Итак, есть объекты и множества. Для любых двух объектов, рассматриваемых в качестве элементов некоторого множества, должна иметься возможность определить, различны они или нет. Множества определяются своими элементами — объектами, из которых они состоят. Отсюда следует, что два множества равны тогда и только тогда, кода они состоят из одних и тех же элементов. Последнее означает, что любой элемент первого множества найдется и среди элементов второго множества и, наоборот, любой элемент второго множества присутствует в первом множестве. Если выполняется только какая-нибудь одна часть данного предложения (отделенная от другой словами "и, наоборот"), то говорят, что одно множество является подмножеством (частью) другого. К множествам, состоящим из элементов, добавим еще одно особое множество, в котором нет никаких элементов; это множество называется *пустым*.

Я полагаю, что на данном этапе рассмотрения у вас должно сложиться более или менее отчетливое представление о множествах. Обратите внимание на простоту и общность понятия множества.

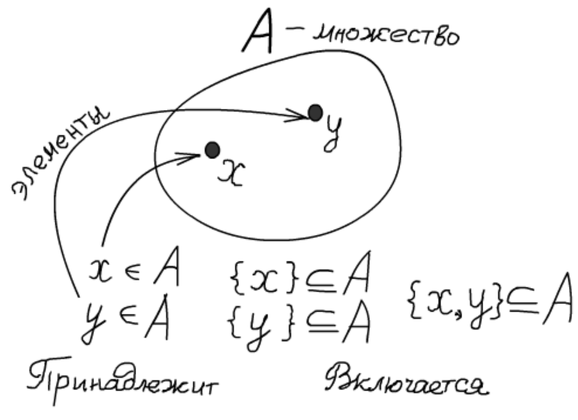
П р о с т а к. Туман поредел, но не рассеялся окончательно. Если я Вас правильно понял, множество подобно мешку с различными предметами. Мешок — это вместилище предметов. Вместо мешка можно использовать ящики, коробки и т. д. Если содержимое двух таких мешков одинаково, то соответствующие этим мешкам множества равны.

П р о ф е с с о р. Ваши образы, Простак, годятся для начальной и простейшей иллюстрации идеи множества. Теперь введем нехитрую символику, чтобы получить возможность рассуждать о множествах более четко и кратко.

Прежде всего нам необходимо особо оговорить отношение принадлежности объекта (элемента) множеству. Чтобы указать, что элемент x принадлежит множеству A, пишут $x \in A$, а если x не принадлежит A, то пишут $x \notin A$. Запись $x_1, x_2, ..., x_n \in A$ будем понимать как сокращение для "$x_1 \in A$ и $x_2 \in A$ и ... $x_n \in A$". Таким образом, символом $\in$ обозначается отношение принадлежности между элементом и множеством. Запись "$x \in A$" можно прочитать различными способами: "x принадлежит A", "x есть член A", "x входит в A". Например, пусть N — множество всех натуральных чисел 1, 2, 3, ..., тогда $5 \in N$, но $0 \notin N$.

Рассмотрим теперь отношения между множествами. Пусть даны два множества, A и B. Если любой (каждый) элемент множества A принадлежит еще и множеству B, то множество A является подмножеством множества B или, другими словами, множество A включается (входит) в множество B. В этом случае пишут $A \subseteq B$, обозначая через $\subseteq$ отношение включения (вхождения) между множествами. Например, множество натуральных чисел является подмножеством множества целых чисел. Очевидно, что каждое множество является подмножеством самого себя: $A \subseteq A$.

Если $A \subseteq B$, но множество B содержит элементы, не принадлежащие множеству A (т. е. неверно, что $B \subseteq A$), то пишут $A \subset B$ и говорят, что A строго включается в B или, иначе, A является собственным подмножеством множества B. Очевидно, что если выполняется строгое включение $A \subset B$, то выполняется и включение (нестрогое) $A \subseteq B$, но обратное утверждение, вообще говоря, неверно.



Принадлежность элемента множеству и включение множества
в другое множество — различные отношения

Внимание

Символы $\in$ и $\subseteq$ обозначают различные отношения: первый из них обозначает отношение между элементом и множеством, а второй — отношение между множествами.

П р о с т а к. Но Вы же говорили, что на природу элементов не накладывается особых ограничений кроме тех, согласно которым элементы должны быть определенными и различными. Эти ограничения не запрещают рассматривать в качестве элемента некоторое множество. Пусть $x \in A$ и x — некоторое множество. Тогда ничто не мешает мне рассматривать одновременно и отношение принадлежности $x \in A$, и отношение включения $x \subseteq A$.

П р о ф е с с о р. Разумеется, Вы можете одновременно рассматривать оба отношения — $x \in A$ и $x \subseteq A$. Я лишь обращаю внимание на то, что это различные отношения. Если для отношения $x \in A$ должно быть задано некоторое правило, согласно которому определяется, принадлежит ли элемент x множеству A, то для отношения $x \subseteq A$ правило выяснения, включается ли множество x в множество A, уже определено. Другими словами, отношение $\in$ первично, поскольку задается при определении множества, а отношение включения $\subseteq$ вторично, поскольку определяется для уже заданных множеств. Далее мы еще не раз вернемся к данному вопросу, но сначала нам потребуются еще несколько важных понятий, связанных с множествами.

Итак, мы уже знаем, как понимать отношение включения одного множества в другое. Теперь определим отношение равенства двух множеств. Два множества равны, если каждое из них является подмножеством другого. Точнее, множества A и B равны (обозначается как $A = B$), если выполняются два включения: $A \subseteq B$ и $B \subseteq A$. В противном случае множества не равны (обозначается как $A \neq B$). Зная, что такое отношение включения, можно сформулировать определение равенства двух множеств следующим образом: $A = B$, если для любого объекта x из того, что $x \in A$ следует, что $x \in B$ и, наоборот, из $x \in B$ следует $x \in A$. Неравенство двух множеств означает, что найдется такой объект x, который принадлежит одному множеству, но не принадлежит другому.

Если требуется доказать, что $A \subseteq B$, нередко поступают следующим образом. Сначала допускают, что некоторый произвольный объект x принадлежит множеству A, а затем показывают каким-нибудь способом, что этот же элемент принадлежит множеству B. Поскольку в допущении речь шла о произвольном (любом) x, значит, все элементы множества A принадлежат и множеству B и, следовательно, выполняется включение $A \subseteq B$. Это так называемое *прямое* доказательство. Бывают и доказательства *от противного*, в которых сначала допускают утверждение, обратное доказываемому: неверно, что $A \subseteq B$. Но это означает, что существует элемент x, принадлежащий одному множеству и не принадлежащий другому. Далее показывается каким-нибудь способом, что возникает противоречие и, следовательно, допущение о невыполнимости включения $A \subseteq B$ неверно.

Приведенное выше определение равенства двух множеств основано на той исходной неформальной идее, что множества полностью задаются своими элементами. Однако нам необходимо еще одно специальное множество, которое не содержит ни одного элемента. Это множество, называемое пустым, обозначается как $\varnothing$. В манипуляциях с множествами оно играет роль, аналогичную нулю в арифметике.

Из изложенного уже можно сделать некоторые выводы.

- Для любого множества A имеет место отношение $A \subseteq A$ — любое множество включает само себя в качестве подмножества. Очевидно, что для любого x, если $x \in A$, то $x \in A$.
- Для любого множества A выполняется $\varnothing \subseteq A$ — пустое множество является подмножеством любого множества. Чтобы доказать это, мы не можем воспользоваться прямым методом, поскольку не можем начать рассуждение с фразы "Пусть $x \in \varnothing$ ...". Действительно, множество $\varnothing$ пусто по определению, и допущение, что $x \in \varnothing$, очевидно противоречит ему. Начав рассуждение с заведомо ложного предположения, мы можем прийти к чему угодно. Поэтому воспользуемся методом доказательства от противного. Предположим, что отношение включения $\varnothing \subseteq A$ не выполняется. Тогда мы должны признать, что существует объект x, такой что $x \in \varnothing$ и $x \notin A$. Но это невозможно, т. к. противоречит определению множества $\varnothing$. Следовательно, предположение о невыполнимости $\varnothing \subseteq A$ неверно, а, значит, отношение $\varnothing \subseteq A$ выполняется.

Обозначим некоторое множество символом A, а все элементы, входящие в это множество, — символами $x_1, x_2, ..., x_n$. Тогда запись $A = \{x_1, x_2, ..., x_n\}$ означает, что множество A содержит только те элементы, которые указаны в фигурных скобках. Напомню, что все элементы в множестве различны, а их порядок перечисления в фигурных скобках не имеет значения. Например, множество $\{x\}$ состоит из единственного элемента x; при этом имеет место отношение $x \in \{x\}$, однако элемент x и множество $\{x\}$ это не одно и то же: можно записать $x \in \{x\}$, но неверно, что $x \subseteq \{x\}$. Пустое множество $\varnothing$ можно обозначить как $\{\}$; заметьте, что множество $\{\varnothing\}$ не пусто, т. к. содержит в качестве своего элемента множество $\varnothing$.

П р о с т а к. Множество $\{\varnothing\}$ не пусто, поскольку не пуст пустой мешок, содержащий другой пустой мешок.

П р о ф е с с о р. Рассмотрим в качестве примера множество букв, из которых составлено слово "множество". Это множество состоит из 8, а не 9 букв: {м, н, о, ж, е, с, т, в}. Напомню, что элементы множества должны быть различны. Слово "множество" представляет собой не просто неупорядоченное множество букв, а конструкцию с более сложной структурой — последовательность вхождений букв в одно и то же слово. Так, буква "о" имеет два вхождения в слово "множество", а остальные буквы — по одному. Вхождение буквы в слово — это пара, состоящая из буквы и номера ее места в слове. Тогда множество всех вхождений букв в слово "множество" можно представить так: {$м_1$, $н_2$, $о_3$, $ж_4$, $е_5$, $с_6$, $т_7$, $в_8$, $о_9$}.

Как уже неоднократно отмечалось, элементы множеств сами могут быть множествами. При этом важно, я не устану это повторять, не путать отношения $\in$ и $\subseteq$. Пусть, например, выполняются отношения $x \in A$ и $A \in B$. Ясно, что A и B — множества. При этом множество A является элементом множества B (но не его подмножеством) и, кроме того, это совсем не означает, что $x \in B$.

Другой пример: если $A = \{B, c, d\}$ и $B = \{a, d\}$, то $B \in A$, но неверно, что $B \subseteq A$. Верно лишь, что $\{B\} \subseteq A$, т. е. множество, состоящее из элемента *B*, является подмножеством множества *A*. Иначе говоря, элементы некоторого множества *A* сами могут быть множествами, но элементы последних не обязательно являются элементами множества *A*. Так, в рассматриваемом примере $a \in A, a \in B, d \in B$, но $d \notin A$.

Множества могут быть конечными и бесконечными. Для конечного множества A количество его элементов обозначается как $|A|$. Если $A = \{x_1, x_2, ..., x_n\}$, то $|A| = n$; $|\varnothing| = 0$. Если конечные множества равны, то равны и количества их элементов, но обратное утверждение, вообще говоря, неверно. Пусть например, A={1, 2, 3}, B={a, b, c}. Тогда |A| = |B| = 3, но $A \neq B$. В случае бесконечных множеств мы не можем указать число, соответствующее количеству элементов в множестве. Вместе с тем, как мы увидим позже, одни бесконечные множества могут быть больше или меньше других бесконечных множеств, а некоторые бесконечные множества могут оказаться равными, хотя это совсем не очевидно на первый взгляд.

Рассмотренный способ задания множества явным перечислением его элементов годится в случае небольших по объему конечных множеств. Это — так называемый *экстенсиональный* (явный) способ определения множества, при котором оно представляется в завершенном виде. Множество можно задать и посредством указания свойства, которым обладают все элементы этого множества и только они. Это — так называемый *интенсиональный* (неявный) способ определения множества, при котором оно представляется не в завершенном, а в потенциальном виде. В этом случае предполагается, что множество можно строить, выбирая элементы из универсума и определяя, обладают они характерным для данного множества свойством или нет. Другими словами, интенсиональный способ основан на использовании предиката — двузначной функции, которая для каждого элемента универсума определяет, принадлежит он данному множеству или нет. Пусть для некоторого множества A определен такой предикат $P_A(x)$. Если вместо переменной x подставить конкретный элемент a, то результатом вычисления $P_A(a)$ будет либо ИСТИНА (И), либо ЛОЖЬ (Л), в зависимости от того, принадлежит или нет элемент a множеству A. С другой стороны, мы можем определить некоторым образом какой-нибудь предикат $Q(x)$, принимающий в качестве аргумента значения из некоторого универсума. Тогда этот предикат задает некоторое множество. А именно множество всех тех элементов универсума, для которых он истинен. С помощью предикатов можно задать как конечные, так и бесконечные множества.

Запись вида $A = \{x : P(x)\}$ определяет множество A через предикат (свойство) P. Она читается так: множество A состоит из всех таких элементов x, для которых предикат $P(x)$ истинен. Например, запись $A = \{x : "x — \textit{красный}"\}$ задает множество всех красных объектов. Другой пример: пусть N — множество натуральных чисел, тогда множество $A = \{x : (x > 5) \,\&\, x \in N\}$ можно рассматривать как множество всех таких натуральных чисел, которые больше 5.

Отношения включения и равенства между множествами можно интерпретировать через предикаты, определяющие эти множества. Пусть $P_A(x)$ и $P_B(x)$ — предикаты, определяющие множества A и B соответственно. Иначе говоря, $P_A(x)$ истинен для всех элементов множества A и ложен для других элементов, а предикат $P_B(x)$ истинен для всех элементов множества B и ложен для других элементов. Тогда $A \subseteq B$, если и только если импликация $P_A(x) => P_B(x)$ истинна для любого объекта универсума. Равенство множеств $A = B$ выполняется тогда и только тогда, когда эквивалентность $P_A(x) <=> P_B(x)$ истинна для любого объекта универсума или, что равносильно, когда истинны две импликации: $P_A(x) => P_B(x)$ и $P_B(x) => P_A(x)$.

Итак, определение множества через предикат как будто дает нам возможность задавать очень большие конечные, а также бесконечные множества. Легко сказать "рассмотрим множество, определяемое свойством P", но подчас трудно сформулировать правила (алгоритм) проверки, обладает ли тот или иной объект этим свойством. Речь идет об эффективном вычислении значения предиката $P(x)$. Например, рассмотрим множество всех погибших в Куликовской битве. Никто не сомневается, что это множество конечно. В нем должно быть несколько тысяч или десятков тысяч элементов, если принять во внимание летописные свидетельства. Но как переформулировать предикат "x погиб в Куликовской битве" так, чтобы действительно можно было определить для каждого из живших в те времена людей, погиб он в Куликовской битве или умер по другим причинам? Да и сам универсум людей, живших во времена Куликовской битвы, также неясно очерчен. Что вы думаете об этом, друзья?

Зануда. Согласно Кантору, множество состоит из определенных и различных между собой элементов. В нашем случае элементы — люди, жившие несколько сотен лет тому назад, — не могут быть достаточно хорошо определены. Поэтому и не может быть определено множество погибших в Куликовской битве. Следовательно, совокупность всех погибших в Куликовской битве не является множеством.

Простак. Занимаясь множествами, я сконцентрировался на способах их задания. Я полагал, что если одно характеристическое свойство множества неудачно, то можно подобрать другое. При этом я забыл, что элементы множества сами должны быть достаточно хорошо определены. Ведь множество определяется своими элементами. Кроме того, я подспудно считал, что множество можно составить из чего угодно и нет ничего такого, чего нельзя было бы представить в виде какого-то множества.

Профессор. Ваши признания, уважаемый Простак, хорошо отражают суть дела. Далее мы еще неоднократно увидим, что не все, что похоже на множество, может быть признано множеством. А вот еще один пример. Рассмотрим множество всевозможных последовательностей нулей и единиц длиной 100. Это конечное, но очень большое по объему множество. Количество его элементов равно 2^{100} и превосходит число всех атомов в видимой части вселенной. Нетрудно написать для компьютера программу, которая генерирует такие последовательности. Код этой программы займет не больше одной страницы, но ее работа даже на самом быстродействующем компьютере потребует времени, многократно превышающего продолжительность жизни многих поколений людей, а для распечатки результатов не хватит земных ресурсов — не то что бумаги, но даже атомов.

Рассмотренные выше трудности определения некоторых множеств можно отнести к категории практических. Однако в математике практическая и теоретическая невозможности чего-либо — различные вещи. Практические трудности игнорируются математиками, а выявление принципиальной недостижимости (фундаментальных пределов) чего-либо является для них чрезвычайно ценным результатом.

Что первично, а что вторично — множество или свойство? Например, свойство "быть красным" может рассматриваться само по себе или, другими словами, как первичное по отношению к множеству всех красных тел. С другой стороны, если применительно к какому-нибудь мешку орехов вводится свойство ореха "быть в мешке", то данное множество оказывается первичным по отношению к данному свойству. Каково бы ни было значение этого различия, оно игнорируется математиками. При любом свойстве P мы можем говорить о множестве P всех элементов, обладающих этим свойством, и относительно любого данного множества P мы можем говорить о свойстве быть элементом множества P. Как следствие такого уравнивания в правах понятий множества и свойства, мы должны принять и правомерность такой записи: $P = \{x : x \in P\}$. Например, "Множество красных тел состоит из красных элементов". Это — непредикативное определение множества через предикат. К чему это может привести, мы рассмотрим позже, а сейчас перейдем к алгебре множеств, т. е. к операциям над множествами.

Я должен отметить соглашение о словоупотреблении, важное для дальнейшего. Когда нам потребуется сказать "Множество A, такое что выполняется включение $A \subseteq B$", мы будем говорить короче: $A \subseteq B$. Например, фразу "Пусть дано множество $A \subseteq B$" следует читать как "Пусть дано множество A, которое включается в множество B". Аналогично, вместо "Элемент x из множества A" мы будем говорить короче: $x \in A$.

3.2. Операции над множествами

П р о ф е с с о р. Над множествами можно выполнять различные операции: объединение, пересечение и вычитание. В результате применения этих операций образуется, вообще говоря, другое множество.

Пусть A и B — некоторые множества, а $P_A(x)$ и $P_B(x)$ — соответствующие им предикаты. Другими словами, $A = \{x : P_A(x)\}$ и $B = \{x : P_B(x)\}$.

- *Объединением* множеств A и B (обозначается как $A \bigcup B$) называется множество, состоящее из всех элементов универсума, которые принадлежат множеству A или множеству B.

  Символически: $A \bigcup B = \{x : P_A(x) \vee P_B(x)\}$ или $A \bigcup B = \{x : (x \in A) \vee (x \in B)\}$.

  Пример: $A = \{a, b, c, x, y\}, B = \{x, y, z\}, A \bigcup B = \{a, b, c, x, y, z\}$.

  Обратите внимание, что в данном примере $|A| + |B| \neq |A \bigcup B|$, т. е. сумма количеств элементов исходных множеств не равна количеству элементов их объединения. Очевидно, что для любого конечного множества A выполняется равенство $A = A \bigcup A$ и, как следствие этого, $|A| = |A \bigcup A|$.

- *Пересечением* множеств A и B (обозначается как $A \bigcap B$) называется множество, состоящее из всех элементов универсума, принадлежащих и множеству A, и множеству B.

  Символически: $A \bigcap B = \{x : P_A(x) \,\&\, P_B(x)\}$ или $A \bigcap B = \{x : (x \in A) \,\&\, (x \in B)\}$.

  Пример: $A = \{a, b, c, x, y\}, B = \{x, y, z\}, A \bigcap B = \{x, y\}$.

  Очевидно, что для любого конечного множества A выполняется равенство $A = A \bigcap A$ и, как следствие этого, $|A| = |A \bigcap A|$.

  Говорят, что множества не пересекаются, если их пересечение есть пустое множество. Обратите внимание: операция пересечения определена и для непересекающихся множеств.

- *Разностью* множеств A и B (обозначается как $A \setminus B$) называется множество, состоящее из всех элементов универсума, которые принадлежат множеству A и не принадлежат множеству B.

  Символически: $A \setminus B = \{x : P_A(x) \,\&\, \overline{P_B(x)}\}$ или $A \setminus B = \{x : (x \in A) \,\&\, (x \notin B)\}$.

  Разность $A \setminus B$ еще называют *относительным дополнением* множества B до множества A, а операцию взятия разности (относительного дополнения) называют также *вычитанием*.

  Пример: $A = \{a, b, c, x, y\}, B = \{x, y, z\}, A \setminus B = \{a, b, c\}$.

- *Дополнением* множества A (обозначается $\overline{A}$) называется множество, состоящие из всех элементов универсума, которые не принадлежат множеству A.

  Символически: $\overline{A} = \{x : \overline{P_A(x)}\}$ или $\overline{A} = \{x : x \notin A\}$.

  Операцию взятия дополнения множества A можно понимать как операцию вычитания множества A из универсума или как операцию взятия относительного дополнения множества A до универсума.

Для наглядного изображения операций над множествами часто используют так называемые *диаграммы Венна*. Универсум объектов изображают в виде квадрата, представляя элементы универсума внутренними точками. Множество элементов универсума изображают в виде некоторой области квадрата, ограниченной замкнутой линией.

Ниже приведены основные равенства, используемые при преобразовании выражений с несколькими операциями над множествами. В этих равенствах универсум обозначен через *U*.

$\overline{\overline{A}} = A$ — закон двойного дополнения (отрицания)

$A \bigcup \overline{A} = U$ — закон исключения третьего

$A \cap \overline{A} = Ш$ — закон противоречия

$A \cup Ш = U$

$A \cap U = A$

Законы Моргана:

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

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

Законы коммутативности (перестановки):

$A \cup B = B \cup A$

$A \cap B = B \cap A$

Законы ассоциативности (группировки):

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

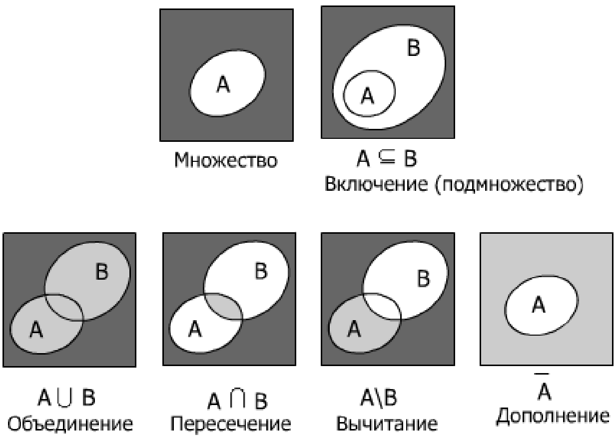
$A \cap (B \cap C) = (A \cap B) \cap C$

Законы дистрибутивности (сочетания операций):

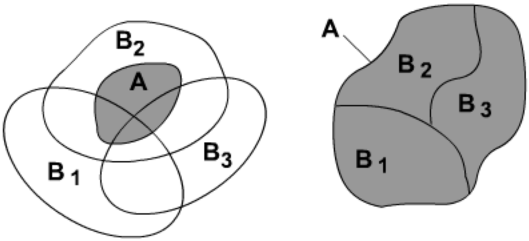
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

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

Говорят, что множества $B_1, B_2, ...$ образуют *покрытие множества* A, если $A \subseteq B_1 \cup B_2 \cup ...$. Множества $B_1, B_2, ...$ образуют *разбиение множества* A, если $A = B_1 \cup B_2 \cup ...$ и, кроме того, попарно не пересекаются. Напомню, что множества B_1 и B_2 не пересекаются, если $B_1 \cap B_2 = Ш$.



Диаграммы Венна, иллюстрирующие действие операций над множествами. Светло-серым цветом выделен результат применения операции



Покрытие (слева) и разбиение (справа) множества A множествами B_1, B_2 и B_3

П р о с т а к. Нетрудно заметить соответствие между операциями НЕ, И и ИЛИ ($\bar{\ }$, $\&, \vee$) алгебры логики с одной стороны и операциями дополнения, пересечения и объединения ($\bar{\ }, \bigcap, \bigcup$) алгебры множеств с другой стороны.

З а н у д а. Это и не удивительно, поскольку операции над множествами определялись через операции над предикатами.

П р о ф е с с о р. Вы оба совершенно правы, заметив и объяснив двойственность операций над предикатами и множествами.

Как уже отмечалось, чтобы доказать равенство двух множеств, обычно показывают, что первое множество включается во второе, а второе — в первое. Кроме того, в самом доказательстве относительно множеств используют предложения наподобие "Пусть некоторый элемент x принадлежит такому-то множеству". Поскольку рассмотрение ситуации начинается с достаточно произвольного элемента, в итоге мы получаем утверждение относительно не только этого элемента, но и всего множества, которому этот "произвольный" элемент принадлежит. Это — общие правила доказательства равенства и включения множеств. В качестве упражнения докажем одно из приведенных выше равенств, а именно один из законов дистрибутивности: $A \bigcup (B \bigcap C) = (A \bigcup B) \bigcap (A \bigcup C)$. Теперь вы знаете, что для этого достаточно доказать два включения:

$$A \bigcup (B \bigcap C) \subseteq (A \bigcup B) \bigcap (A \bigcup C)$$

$$(A \bigcup B) \bigcap (A \bigcup C) \subseteq A \bigcup (B \bigcap C).$$

1. Доказательство включения $A \bigcup (B \bigcap C) \subseteq (A \bigcup B) \bigcap (A \bigcup C)$. Выберем произвольный элемент $x \in A \bigcup (B \bigcap C)$. Тогда $x \in A$ или $x \in B \bigcap C$. Если $x \in A$, то верно и то, что $x \in A \bigcup B$ и $x \in A \bigcup C$, а, следовательно, x есть элемент пересечения этих множеств. Если же $x \in B \bigcap C$, то $x \in B$ и $x \in C$. Следовательно, и в этом случае $x \in A \bigcup B$ и $x \in A \bigcup C$. Таким образом, $x \in (A \bigcup B) \bigcap (A \bigcup C)$. Поскольку мы рассматривали произ-

вольный элемент $x \in A \bigcup (B \bigcap C)$, требуемое включение множеств доказано.

2. Доказательство включения $(A \bigcup B) \bigcap (A \bigcup C) \subseteq A \bigcup (B \bigcap C)$. Пусть $x \in (A \bigcup B) \bigcap (A \bigcup C)$. Тогда $x \in A \bigcup B$ и $x \in A \bigcup C$. Следовательно, $x \in A$ или же $x \in B$ и $x \in C$, т. е. $x \in A \bigcup (B \bigcap C)$.

На этом доказательство равенства двух множеств закончено.

Предлагаю вам в качестве упражнения доказать другие равенства алгебры множеств.

3.3. Первые парадоксы теории множеств

Профессор. Уже на данном этапе рассмотрения теории множеств обнаруживаются некоторые парадоксы, связанные с определением понятия множества. Ранее мы уже рассматривали так называемый "Парадокс брадобрея", обнаруженный Б. Расселом. Напомню его содержание: брадобрей это человек, который бреет только тех людей, которые не бреются сами; если при этом допустить, что брадобрей должен брить себя, то сразу же следует, что он не должен брить себя, и, наоборот, из допущения, что брадобрей не должен брить себя, вытекает, что он должен это делать. Рассмотрим теоретико-множественную трактовку этого парадокса, для чего определим множество K клиентов брадобрея:

$K = \{x : x\ \textit{не бреет}\ x\}$ — множество всех таких x, для которых предикат "x *не бре*ет x" истинен.

Обозначим брадобрея буквой b. Предположим, что $b \in K$, т. е. брадобрей принадлежит множеству своих клиентов. Тогда верно высказывание "b не бреет b", т. е. брадобрей не является собственным клиентом и, следовательно, $b \notin K$. С другой стороны, допустим $b \notin K$, тогда высказывание "b не бреет b" ложно, т. е. верно высказывание "b бреет b" и, следовательно, $b \in K$. Таким образом, при любом допущении о том, является брадобрей собственным клиентом или нет, мы приходим к противоречию.

Далее, в различного рода рассуждениях может возникнуть конструкция, которая включает самое себя в качестве элемента. Рассмотрим множество всех объектов за исключением какого-нибудь одного, например, арбуза. Очевидно, это множество не является арбузом. Но тогда оно должно содержать себя в качестве элемента. Таким образом, возникает уродливая конструкция вида $A = \{A, ...\}$.

Простак. А в чем все-таки уродливость множества, которое содержит в качестве элемента само себя?

З а н у д а. Позвольте мне попытаться объяснить это. У меня есть два аргумента.

Во-первых, пусть для ясности $A = \{A\}$, т. е. множество A состоит из одного элемента A. Тогда, подставляя $\{ A \}$ в $\{ A \}$ вместо A, последовательно будем получать: $\{\{ A \}\}$, $\{\{\{ A \}\}\}$, $\{\{\{\{ A \}\}\}\}$ …, $\{\ldots\{ A \}\ldots\}$, ... и так далее сколь угодно долго. Таким образом, описание даже одноэлементного множества оказывается бесконечным, что, согласитесь, выглядит по меньшей мере странным.

Во-вторых, пусть даны два множества: A и $B = \{A\}$. Рассмотрим обстоятельства, при которых могло бы выполняться равенство $A = B$, а следовательно, и $A = \{A\}$. Чтобы $A = B$, необходимо выполнение включения $A \subseteq B$.

- ❒ Если множество A пусто, то $A \subseteq B$ выполняется, т. к. Ш⊆ В для любого множества B.
- ❒ Пусть множество A не пусто. Возьмем произвольный элемент $x \in A$, отличный от A (разумеется, мы не исключаем, что A может содержать, кроме x, и элемент A). Тогда для выполнения включения $A \subseteq B$ необходимо, чтобы $x \in B$. Но этого быть не может, поскольку множество B по определению содержит единственный элемент A, причем отличный от x. Следовательно отношение включения $A \subseteq B$ не выполняется, а значит, $A \neq B$ и $A \neq \{A\}$. Таким образом, если A не пусто и содержит элемент отличный от A, то нельзя доказать, что $A = \{A\}$.

П р о ф е с с о р. Да, множества, содержащие себя в качестве элемента, плохи в том смысле, что они связаны с логическими несуразностями. Мы можем попытаться исключить их из категории "правильных" множеств. Однако и в этом случае мы не застрахованы от парадоксов. Так, каталог книг содержит наименования других книг кроме своего собственного, а мы хотим составить каталог всех таких каталогов. Спрашивается, должен ли указанный каталог содержать сам себя?

П р о с т а к. На первый взгляд кажется, что ответ должен быть отрицательным, поскольку каталог не должен содержать самого себя по определению. Но я теперь знаю, как разбирать подобные задачи.

Определение каталога каталогов, не содержащих самих себя в качестве элемента, можно задать в виде множества

$$K = \{x : x \notin x\}.$$

Вопрос заключается в том, имеет ли место отношение $K \in K$?

Допустим сначала, что отношение $K \in K$ выполняется, но тогда истинно высказывание "$K \notin K$". С другой стороны, предположим, что $K \notin K$, но тогда высказывание " $K \notin K$ " ложно, т. е. истинно противоположное высказывание " $K \in K$ ". Таким образом, обнаруженное противоречие не позволяет нам ни включить в каталог каталогов самого себя, ни исключить его из своего содержимого.

З а н у д а. Выходит, что плохо в обоих случаях — когда множество содержит себя в качестве элемента и когда не содержит.

П р о ф е с с о р. Вы неточно выразились. Плохо, когда множество содержит себя в качестве элемента и когда оно состоит из всех множеств, не содержащих себя в качестве элемента.

Все рассмотренные до сих пор парадоксы имеют общую природу и допускают множество сюжетных вариантов. В каждом из них некое множество определяется как содержащее те и только те элементы, которые не находятся в определенном отношении к себе. Парадокс обнаруживается при попытке ответить на вопрос: принадлежит ли данное множество самому себе?

К рассмотренным ранее парадоксам я добавлю еще один, который был придуман Д. Бери и получил название "Парадокса слов". Это наглядная перефразировка парадокса, обнаруженного Ж. Ришаром в 1905 г. Каждое целое число допускает множество различных словесных описаний. Например, число 3 можно задать одним словом "Три" или фразой "Целое число, непосредственно следующее за числом два". Рассмотрим все возможные описания целых чисел, состоящие, например, не больше чем из 100 русских букв, знаков препинания и пробела. Как известно, в русском алфавите 33 буквы. Добавим к этому множеству букв пробел, запятую и точку. В итоге получим 36 символов. Тогда описаний чисел может быть не больше чем 36^{100}. Это хоть и очень огромное, но все же конечное число. Поскольку целых чисел бесконечно много, существуют какие-то числа, не задаваемые описаниями длиной не более 100 символов. Рассмотрим *наименьшее целое число, которое не может быть описано не более чем ста символами*. Но описание такого числа только что было дано фразой "Наименьшее целое число, которое не может быть описано не более чем ста символами", состоящей из 80 символов.

Зная причину возникновения парадоксов, вы без труда сможете придумать собственные варианты. Замечу, что однажды в популярной телевизионной игре "Что? Где? Когда?" знатокам был задан некий парадоксальный вопрос, хотя, как мне кажется, это противоречит правилам игры. Знатоки не разглядели подвоха и остановились на одном из ответов, не проверив другой. Ответ, разумеется, не был признан правильным. Не установив противоречивость вопроса, знатоки были просто обречены на проигрыш.

П р о с т а к. Итак, уважаемый профессор, я не сомневаюсь в том, что следует избегать "неправильных" множеств, манипуляция которыми чревата парадоксами. Мы уже выяснили, как получаются некоторые из них, и готовы их отбросить. Но можем ли мы быть уверенными, что нет других типов "неправильных" множеств? Не лучше ли как-то скорректировать саму теорию множеств, чтобы гарантированно избежать неприятностей в дальнейшем?

Зануда. Лично я за пересмотр теории множеств. Нам следует лишь уточнить определение множества, т. е. добавить что-то к свойствам элементов и множеств, а также наложить некоторые ограничения на отношение принадлежности между ними. Определение множества, данное Кантором, слишком расплывчато, поэтому и стали возможными противоречивые конструкции.

Профессор. Превосходные мысли! Приблизительно так же подумали и величайшие математические умы в начале XX века, когда Б. Рассел обнаружил первый парадокс теории множеств. Для наглядности он облек его в форму притчи о брадобрее. К этому времени выдающийся немецкий математик Г. Фреге завершил второй том своего фундаментального труда "Основания арифметики", в котором логическое обоснование арифметики, а значит и всей математики, покоилось на теории множеств. Но теория множеств, развитая Фреге, допускала образование множества всех множеств, не содержащих себя в качестве элемента. Мы теперь уже знаем, что такое множество противоречиво. Об этом и сообщил Рассел в 1902 г. в письме Фреге. Таким образом получалось, что арифметика Фреге опиралась на противоречивую теорию множеств, а значит не могло быть и речи о непротиворечивости самой арифметики и математики в целом. Фреге пришлось дописать к своей книге, уже лежавшей в типографии, небольшое приложение, в котором он признался: "Вряд ли с ученым может приключиться что-нибудь худшее, чем если у него из-под ног выбьют почву в тот самый момент, когда он завершит свой труд. Именно в таком положении оказался я, получив письмо от Бертрана Рассела, когда моя работа уже была почти закончена". С этого момента некоторые из ведущих математиков бросились спасать теорию множеств, стремясь сделать ее непротиворечивой. Среди них был и Б. Рассел:

> Я предпринял решительную попытку найти решение парадоксов. Их существование я рассматривал почти как личный вызов и, если бы потребовалось, посвятил бы всю оставшуюся жизнь попыткам разрешить их. Однако по двум причинам такая приверженность идее избавления от парадоксов казалась мне нежелательной. Во-первых, вся проблема представлялась мне тривиальной. Во-вторых, сколько я ни пытался, мне не удавалось ни на шаг продвинуться в ее решении. 1903 и 1904 годы почти целиком ушли на борьбу с парадоксами, но без сколько-нибудь ощутимых признаков успеха.

Выход из создавшегося положения был найден Б. Расселом и А. Уайтхедом. с помощью созданной ими теории типов.

Альфред Норт Уайтхед (1861—1947)

Теория типов довольно сложна, но основана на простой идее, которая напрашивается чуть ли не сама собой. Первым делом Рассел и Уайтхед ввели требование, согласно которому то, что содержит все элементы множества, не должно быть элементом того же множества. Чтобы удовлетворить этому требованию, они ввели иерархию типов объектов: индивидуумы (например, упомянутый Саша или какое-то конкретное яблоко) имеют тип 0; множество индивидуальных объектов принадлежит типу 1; множество множеств индивидуальных объектов — типу 2 и т. д. Так, если x принадлежит y, то y должен принадлежать более высокому типу, чем x. Кроме того, нельзя говорить о множестве, которое принадлежит самому себе.

Аналогичная иерархия типов была предложена и для логики: индивидуумы — объекты типа 0; любое утверждение о свойстве индивидуума имеет тип 1; любое утверждение о свойстве свойства индивидуума имеет тип 2 и т. д. При переходе к предикатам теория типов несколько усложняется. Ни один из аргументов предиката не должен определяться через сам предикат. При данном условии предикат имеет более высокий тип, чем входящие в него переменные.

В рамках теории типов можно избежать всех рассмотренных парадоксов и их аналогов. Например, рассмотрим еще раз определение гетерологических прилагательных: прилагательное называется гетерологическим, если оно неприменимо к самому себе *(см. разд. 2.3.4)*. Это определение множества всех гетерологических прилагательных и поэтому оно принадлежит более высокому типу, чем каждое конкретное гетерологическое прилагательное (например, "русский", "многосложный" и т. д.). Таким образом, вопрос о том, является ли гетерологическим слово "гетерологический", просто неправомерен, и, следовательно, парадокс не возникает.

Аналогично, правило "Все правила имеют исключения" относится к более высокому типу (т. е. к метаправилам), чем конкретное правило, например, "Все компьютерные программы содержат ошибки". Все, что утверждается в метаправиле, к нему самому не относится, и, таким образом, парадокс не возникает.

В теории типов разрешается и парадокс лжеца, проанализированный еще Аристотелем. Он возникает при попытке оценить истинность высказывания "Это утверждение ложно". При обычной интерпретации получается следующее. Если предложение, указанное в кавычках, истинно, то истинно и то, что оно утверждает. Следовательно, предложение ложно. Если же предложение ложно, то ложно и то, что оно утверждает. Следовательно, предложение истинно. Другой вариант этого парадокса: является ли высказывание "То, что я говорю, — ложь" истинным или ложным? Если сказавший это действительно лжет, то он говорит правду. Если же он говорит правду, то он лжет. Таким образом, возникает противоречие.

Парадокс лжеца разрешается следующим образом. Высказывание "То, что я говорю, — ложь" означает "Я высказываю некоторое утверждение X, и X ложно". Последнее есть утверждение относительно высказывания X, причем более высокого типа, чем тип высказывания X. Следовательно, если утверждение относительно X истинно, то само X ложно, а если утверждение относительно X ложно, то само X истинно. Таким образом, противоречие не возникает.

Простак. Если я правильно понял, то в обычной речи нужно следить за уровнями (типами) высказываний, являющихся составными частями более сложного утверждения.

Профессор. По крайней мере, это необходимо делать при ощущении, что вас могут неправильно понять. На самом нижнем уровне находятся высказывания об объектах. Это — объектный язык. Слова "истина" и "ложь" не входят в него. Чтобы говорить об истинности или ложности высказываний, выраженных на объектном языке самого нижнего уровня, необходимо воспользоваться языком следующего, второго уровня, который выполняет роль метаязыка. Метаязык содержит в себе объектный язык, но не только его. Как говорил польский математик А. Тарский, высказывание "Снег белый" принадлежит объектному языку, а высказывание "Утверждение

'Снег белый' истинно" — метаязыку. В обычной речи понимания и использования двух уровней языка, как правило, оказывается достаточно. Однако наш разум может позволить себе заняться значениями истинности высказываний самого метаязыка. В этом случае мы должны осознавать, что поднимаемся на следующий уровень, на котором уже используется метаметаязык. Язык более высокого уровня позволяет формулировать утверждения об истинности или ложности высказываний всех языков более низких уровней.

3.4. Бесконечные множества

Профессор. Теория множеств создавалась как инструмент для выяснения устройства бесконечных совокупностей объектов. Бесконечность всегда привлекала внимание людей. Термином "бесконечность" сначала обозначали все, что было невозможно сосчитать или перечислить. Бесконечное — это что-то запредельное, невообразимо большое или, напротив, чрезвычайно малое, к чему можно стремиться сколь угодно долго, но достичь чего невозможно. Приведите, пожалуйста, пример бесконечного объекта.

Простак. Первое, что приходит мне на ум, — Вселенная, состоящая из бесчисленного количества звезд, атомов и других частиц.

Зануда. Множество целых чисел бесконечно: какое бы число мы ни взяли, всегда можно перейти к следующему, прибавив к предыдущему 1. Таким образом, мы никогда не сможем сказать: "Вот, все целые числа перечислены и других целых чисел больше нет".

Профессор. На каком основании Вы, Простак, считаете Вселенную бесконечной? В действительности и Вы, и никто другой не знаете, какая она, — бесконечная или конечная. Муравью килограммовый камень кажется бесконечно тяжелым. Число 2^{100} настолько огромно, что простой перебор всех целых чисел, не превышающих его, с помощью самого современного компьютера потребует невообразимо много сроков жизни самого компьютера. Тем не менее, количество таких чисел конечно. Это означает, что процесс их пересчета когда-нибудь да закончится. Мы называем Вселенную бесконечной, имея в виду лишь ее очень большие размеры, а также то, что никто из людей никогда не достигал ее границ. А вот множество всех целых чисел в самом деле бесконечно. Иначе говоря, мы доподлинно знаем, что оно бесконечно, поскольку таковым по определению оно создано нашим разумом. Для обозначения бесконечно большого количества в математике используют специальный символ ∞, а для бесконечно малого — $1/\infty$. Однако

∞ и $1/\infty$ это не числа, а просто символы, поскольку они не подчиняются обычным правилам арифметики. Так, $2\times\infty=\infty$, $100+\infty=\infty$, $\infty/1000=\infty$, т. е. арифметические операции над бесконечно большим количеством оставляют это количество бесконечно большим.

С античных времен Аристотеля бесконечное множество рассматривали как совокупность объектов, данную не в завершенном виде, а как постоянно формирующуюся. По заданному правилу эту совокупность можно было дополнить новыми объектами когда угодно. Так, множество целых чисел считалось существующим не *актуально*, в готовом или завершенном виде, а существующим только *потенциально*. Его можно было постоянно пополнять новыми членами, но этот процесс никогда не мог быть завершен. Такой взгляд на бесконечные множества существовал вплоть до начала 1870-х годов, когда Георг Кантор стал рассматривать бесконечные множества как данные актуально, в завершенном виде. Тогда оказалось, что одни бесконечные множества могут включать в себя другие. Например, множество всех целых чисел включает в себя множества четных и нечетных чисел, причем все эти три множества бесконечны. Разве это не удивительно? Получается, что части столь же велики, что и целое. Естественно возникает вопрос: насколько велико то или иное бесконечное множество?

Если быть более точным, то почти за 25 лет до Кантора чешский математик Б. Больцано уже рассматривал актуально бесконечные множества и считал, что одно бесконечное множество может содержать в себе бесконечное подмножество.

Бернард Больцано (1781—1848)

3.4.1. Сравнение бесконечных множеств

Профессор. Конечные множества можно сравнить (хотя бы принципиально, если не практически) в количественном отношении так: сначала подсчитать количество элементов в каждом из них, а затем сравнить полученные числа. Для бесконечных множеств данный способ не годится, поскольку подсчет количества их элементов никогда не закончится или, лучше сказать, не завершится какими-то числами, которые можно сравнивать. Тем не менее, интуиция подсказывает нам, что одни из бесконечных множеств могут быть больше других. Попробуйте сравнить бесконечные множества всех натуральных и четных натуральных чисел.

Простак. Натуральные числа это 1, 2, 3, ... Поскольку только каждое второе из них четное, несомненно, четных чисел меньше в два раза, хотя и тех, и других бесконечно много.

Зануда. Странно все это. С одной стороны, оба множества бесконечны и не имеют количественной характеристики, а с другой — одно из них больше другого, и даже видно, во сколько раз.

Профессор. Похоже, вы оба согласны с тем, что бесконечные множества можно сравнивать. Однако результат сравнения, полученный Простаком, оказался неверным. В действительности четных натуральных чисел столько же, сколько и всех натуральных чисел. Прежде чем разобраться с этой задачей, попробуйте решить более простую. Пусть имеются конечные множества болтов и гаек. Как определить без подсчета, равны ли их количества?

Простак. Я бы стал навинчивать гайки на болты или просто выкладывать пары болт-гайка. Если для каждого болта найдется своя гайка, то количества всех болтов и гаек равны, а в противном случае — нет. Таким образом, я решил задачу, не выясняя, сколько именно было гаек и болтов.

Профессор. Прекрасно, хотя сформулировал эту идею еще Б. Больцано, а затем Г. Кантор использовал ее в качестве главного принципа сравнения объемов множеств. В своей книге "Парадоксы бесконечного" Больцано в частности писал: "Я утверждаю: два множества, оба бесконечные, могут быть в таком отношении друг к другу, что, с одной стороны, можно связать в пару каждый предмет, принадлежащий одному из них, с каким-то предметом, принадлежащим другому из них, таким образом, чтобы ни один предмет в обоих множествах не остался без пары, а также чтобы ни один из них не встретился в двух и более парах..."

Теперь воспользуйтесь методом Больцано применительно к бесконечным множествам чисел.

П р о с т а к. Идея метода состоит в сопоставлении друг другу элементов различных множеств. Выпишем ряд натуральных чисел, а под ним — ряд натуральных четных чисел:

1, 2 , 3, 4, 5, 6, … — натуральные числа;

2, 4, 6, 8, 10, 12, … — четные натуральные числа.

Выписать все члены рядов, очевидно, невозможно, но уже по нескольким первым из них нетрудно заметить, что каждому натуральному числу однозначно соответствует некоторое четное натуральное число. Так, натуральному числу n соответствует четное натуральное число $2n$. И наоборот, каждому четному натуральному числу k однозначно соответствует натуральное число $k/2$. Выходит, можно образовать пары из элементов рассматриваемых множеств так, что ни один элемент не окажется без пары. Следовательно, множества натуральных и четных натуральных чисел равночисленны, если позволительно так сказать.

П р о ф е с с о р. Вы прекрасно справились с данной задачей. Поскольку количество элементов бесконечного множества не может быть выражено каким-либо числом, говорят не о количестве, а о *мощности* множества. Для конечных множеств понятия количества элементов и мощности совпадают. Итак, только что установлено, что множества всех натуральных чисел и четных натуральных чисел равномощны. Очевидно, что аналогичным способом можно доказать, что равномощны множества натуральных и нечетных натуральных чисел.

З а н у д а. Я вынужден согласиться с приведенным доказательством равномощности рассматриваемых множеств, хотя результат не согласуется с интуицией, которая говорит нам, что множество натуральных чисел состоит из четных и нечетных натуральных чисел, взятых вместе. Таким образом, мы видим, что нарушается принцип "целое больше своей части".

П р о ф е с с о р. Это один из примеров того, что иногда математика позволяет выяснить нечто неподвластное нашей интуиции. Соотношения, выполняющиеся для конечных множеств, могут не выполняться применительно к бесконечным множествам. Бесконечность имеет особые свойства, которых нет в конечном. Так в чем же состоит метод, с помощью которого мы устанавливаем равномощность или неравномощность множеств?

З а н у д а. Как мы видели, суть метода состоит в попытке установить соответствие между элементами сравниваемых множеств. Это соответствие должно быть взаимно однозначным, чтобы множества были равномощными. Иначе говоря, каждому элементу одного множества должен соответствовать некоторый единственный элемент другого множества, и наоборот.

П р о ф е с с о р. Вы верно выразили суть метода сравнения множеств по их мощности. Теперь вы сможете сравнить множества натуральных и целых чисел. Попробуйте сделать это.

П р о с т а к. Целые числа это натуральные числа плюс те же натуральные числа, но со знаком "минус", и еще 0. Иначе говоря, целые числа можно представить рядом: …, –3, –2, –1, 0, 1, 2, 3, …. Очевидно, что множество целых чисел содержит в качестве своего подмножества все натуральные числа. Однако это еще не основание говорить о том, что данные множества неравномощны. Попробуем установить взаимно однозначное соответствие между ними, для чего выпишем ряды чисел один под другим:

…, –3, –2, –1, 0, 1, 2, 3, … — целые числа;

1, 2, 3, … — натуральные числа.

Нетрудно заметить, что целых чисел больше, чем натуральных.

З а н у д а. Совершенно неверно! Простак выбрал способ расположения числовых рядов, который сам собой напрашивается, и сразу же сформулировал категорический вывод. А может быть существует другой способ расположения рядов, при котором вывод будет совершенно противоположным? Давайте упорядочим целые числа иначе: 0, 1, –1, 2, –2, 3, –3, …, а натуральные числа оставим в их естественном порядке. Расположим ряды этих чисел один под другим:

0, 1, –1, 2, –2, 3, –3, … — целые числа;

1, 2, 3, 4, 5, 6, 7, … — натуральные числа.

Нетрудно заметить, что n-е по порядку целое число (обозначим его через z_n) можно вычислить по следующей формуле:

- $z_n = n / 2$ — если n четное;
- $z_n = -(n - 1) / 2$, если n нечетное.

С другой стороны, для каждого целого числа z можно однозначно указать его номер (натуральное число):

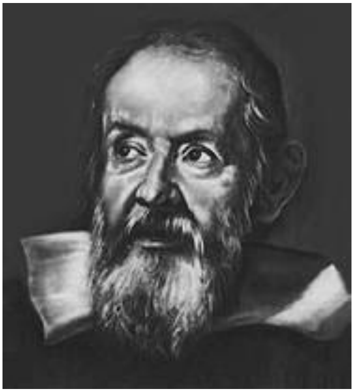
- $n = 2z$, если $z > 0$;
- $n = -2z + 1$, если $z \leq 0$.

Например, четвертое целое число равно $4/2 = 2$, а пятое равно $-(5 - 1) / 2 = -2$; с другой стороны, числу 3 сопоставляется натуральное число $2 \times 3 = 6$, а числу –3 сопоставляется номер (натуральное число) $-2 \times (-3) + 1 = 7$.

Таким образом, между целыми и натуральными числами установлено взаимно однозначное соответствие: каждому натуральному числу однозначно сопоставляется некоторое единственное целое число и, наоборот, каждому целому числу однозначно соответствует некоторое единственное натуральное число. Следовательно, натуральных чисел не меньше, но и не больше, чем целых, а значит, их множества равномощны.

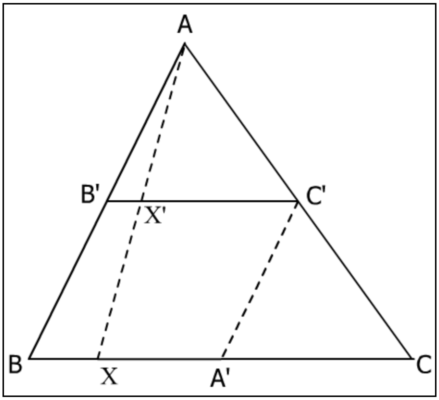
П р о ф е с с о р. Хорошее решение, приводящее к правильному результату.

Итак, мы только что убедились, что множества всех натуральных, только четных натуральных, только нечетных натуральных и целых чисел равномощны. При этом равномощность множеств определяется как существование взаимно однозначного соответствия между элементами данных множеств. Замечу, что кажущаяся парадоксальность полученных результатов имеет своей причиной рассмотрение бесконечных множеств как актуально данных в своем завершенном виде. Такой подход к бесконечным множествам был выполнен Кантором преднамеренно, хотя еще Галилео Галией указывал, что он приведет к необходимости признать: четных чисел столько же, сколько четных и нечетных вместе, а это, по мнению подавляющего большинства его современников и предшественников, считалось абсурдным.



Галилео Галилей (1564—1642)

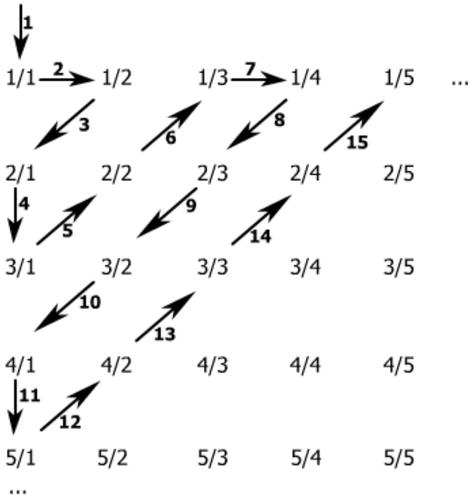
Для бесконечных множеств такое утверждение, как "часть меньше целого" перестает быть верным. Это заметил еще древнегреческий математик Зенон, рассматривая следующее обстоятельство. Пусть дан треугольник ABC, в котором отрезок B'C' параллелен основанию BC и соединяет середины сторон, выходящих из вершины A. Спроектируем из этой вершины на BC все точки отрезка B'C'. Тогда каждой точке отрезка B'C' будет соответствовать некоторая точка основания BC и, наоборот, каждой точке основания треугольника будет соответствовать какая-то точка отрезка B'C'. Например, точке X' будет соответствовать точка X, а точке X — точка X'. Следовательно, на отрезке B'C', параллельном и равном отрезку BA' (A' — середина основания BC), размещается столько же точек, сколько и на вдвое большем отрезке BC. При этом Зенон не знал, что аналогичный парадокс возникает при сравнении бесконечных подмножеств множества целых чисел.



На отрезке B'C' размещается столько же точек,
сколько на вдвое большем отрезке BC

А теперь рассмотрим множество рациональных чисел, т. е. целых и дробей. Рациональное число, как известно, можно представить как частное от деления двух целых чисел или в виде пары двух целых чисел m/n, где m — числитель, а n — знаменатель. Особенность множества рациональных чисел состоит в том, что между любыми соседними целыми числами находится бесконечно много рациональных чисел. Например, между 0 и 1 находятся дроби 1/2, 1/3, 1/4, 1/5, … . Возникает подозрение, что рациональных чисел больше, чем целых: целых бесконечно много, а между любыми соседними целыми находится тоже бесконечно много дробных чисел. Получается бесконечное множество, составленное из бесконечных множеств, что, согласитесь, представляется более сложным образованием, чем множество натуральных чисел. Однако, как показал Кантор в 1874 г., рациональных чисел столько же, сколько и натуральных. Трудность, с которой он столкнулся, заключалась в поиске способа нумерации рациональных чисел. Кантор расположил рациональные числа не в один ряд, а в бесконечной квадратной таблице, т. е. в бесконечно много рядов. Далее оставалось только найти зигзагообразный путь обхода всех чисел, позволяющий последовательно пронумеровать каждое из них.

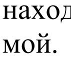
Всякое множество, элементы которого можно взаимно однозначно сопоставить натуральным числам, Кантор назвал *счетным*. Другими словами, элементы счетного множества можно пересчитывать.



Метод доказательства счетности рациональных чисел. Стрелками указан маршрут обхода рациональных чисел. Около стрелки указан номер, взаимно однозначно сопоставляемый рациональному числу, на которое эта стрелка указывает. Например, числу 2/3 соответствует натуральное число 8

З а н у д а. А существуют ли несчетные множества, элементы которых пересчитать нельзя?

П р о ф е с с о р. Да, существуют. Например, множество всех действительных чисел, содержащее кроме рациональных еще и иррациональные числа (например, $\sqrt{2}$, π), несчетно. Множество действительных чисел находится во взаимно однозначном соответствии с множеством точек прямой. Обратите внимание, что множество рациональных чисел значительно "плотнее" множества целых чисел (между любыми соседними целыми числами располагается бесконечное множество рациональных чисел), тем не менее, оба эти множества равномощны. Плотность же множества действительных чисел существенно больше, и пересчитывать их нельзя. Последнее означает, что не существует взаимно однозначного соответствия между множествами действительных и натуральных чисел.



П р о с т а к. Мне трудно поверить в это. Что же может помешать нам последовательно, без пропусков, перебирать действительные числа, приписывая им натуральные номера?

Зануда. Сначала всем нам было трудно поверить, что натуральных чисел столько же, сколько целых и даже рациональных, а теперь трудно поверить, что есть множество, для нумерации элементов которого натуральных чисел просто не хватит. Но ведь именно последнее лучше всего согласуется с нашей интуицией. Воистину, наш разум способен сначала из сложного сделать простое, чтобы потом еще больше все усложнить! Вы хотите сказать, Профессор, что нельзя придумать способ нумерации действительных чисел?

Профессор. Именно так! Рассмотрим идею доказательства этого утверждения, придуманного Кантором. Сначала он предположил, что взаимно однозначное соответствие между натуральными и действительными числами существует. Затем он показал, что данное предположение приводит к противоречию и, следовательно, взаимно однозначное соответствие между натуральными и действительными числами невозможно. Таким образом, Кантор использовал метод доказательства от противного.

Доказательство можно заметно упростить, рассматривая только подмножество всех действительных чисел, заключенных между 0 и 1. Если уж это бесконечное подмножество окажется больше множества всех натуральных чисел, то все множество действительных чисел и подавно больше него.

Итак, предположим, что действительные числа в промежутке между 0 и 1 могут быть пронумерованы без повторов и пропусков натуральными числами. Другими словами, мы допускаем гипотезу, что все действительные числа, расположенные между 0 и 1, могут быть поставлены во взаимно однозначное соответствие с множеством натуральных чисел. Это означает, что мы можем составить некий бесконечный перечень действительных чисел, каждое из которых можно представить в виде бесконечной десятичной дроби.

Всякое действительное число можно представить бесконечной десятичной дробью. Такие бесконечные десятичные дроби, как 0,5000…, представим в виде эквивалентной бесконечной дроби 0,4999…. Перечень всех десятичных дробей, представляющих все действительные числа в промежутке от 0 до 1, может быть любым с точки зрения их порядка. Просто представим мысленно этот список: бесконечное количество различных бесконечно длинных десятичных представлений действительных чисел от 0 до 1. Каким бы ни был этот список, пронумеруем его элементы натуральными числами без пропусков и повторений. Перечень или, другими словами, список это — последовательность элементов, которым можно сопоставить натуральные числа: первому в списке элементу сопоставим 1, второму — 2 и т. д.

Теперь задача состоит в том, чтобы построить такую десятичную дробь, которой нет в указанном списке. Если нам это удастся, то тем самым мы докажем, что наш первоначальный перечень всех действительных чисел от 0 до 1 не полон и, следовательно, его нумерация не является нумерацией всех десятичных дробей от 0 до 1. Разумеется, мы можем пополнить начальный перечень вновь сконструированным числом. Однако ничто не мешает нам повторить аналогичные рассуждения, с помощью которых мы создадим еще одно новое действительное число, которого раньше не было в списке, и т. д. Итог будет один — список нумерованных действительных чисел всегда не полон. А это означает, что наша нумерация относится не к тому объекту, для которого она предназначалась изначально. Иначе говоря, наша нумерация нумерует не все действительные числа от 0 до 1 и, следовательно, она не является нумерацией этого подмножества действительных чисел. Так как это положение вещей будет сохраняться при сколь угодно долгом пополнении первоначального списка вновь созданными действительными числами, мы должны признать, согласно Кантору, что нумерации действительных чисел просто не существует. Такова идея доказательства.

Зануда. Уважаемый Профессор, если я Вас правильно понял, получается следующее. Сначала мы предполагаем, что множество действительных чисел можно представить в виде перечня всех его элементов, пусть даже бесконечного. Само понятие перечня предполагает некоторую, хотя бы произвольную, упорядоченность его элементов. Так, перечень создается из элементов любого множества следующим образом: сначала както выбирается первый элемент, затем второй и, далее, все остальные. Коль скоро мы можем сделать это, то можем и пронумеровать элементы этого списка натуральными числами 1, 2, 3, … . Таким образом, гипотеза о возможности нумерации действительных чисел уже провозглашена. Далее Вы, Профессор, соглашаетесь с Кантором, что любая нумерация этого множества таковой не является по той простой причине, что само множество не соответствует своему определению, т. е. не является множеством *всех* действительных чисел в промежутке от 0 до 1 — ведь, по-вашему, можно построить число, не входящее в исходный перечень. На этом основании Вы заключаете, что нумерация действительных чисел невозможна.

Профессор. Вы верно поняли мысль Кантора, которую я намерен лишь только растолковать, не претендуя ни на что большее. Давайте теперь рассмотрим ее техническую сторону. Это интересно как феномен изобретательской деятельности человека.

Итак, мы должны построить новое число, которого не было в первоначальном списке. Это число должно отличаться по крайней мере одним десятичным знаком (цифрой в одном из разрядов) от каждого из действительных чисел в списке. Вместе с тем, оно должно быть действительным числом, расположенным между 0 и 1.

Пусть имеется бесконечный перечень бесконечных десятичных представлений действительных чисел от 0 до 1. В этом перечне такие бесконечные десятичные дроби, как 0,5000…, представим в виде эквивалентной бесконечной дроби 0,4999… . Для построения нового числа, не входящего в указанный перечень, выполним следующие действия:

1. Берем первое число в исходном перечне. Первую десятичную цифру в новом числе определяем так:
   - Если первая цифра десятичного представления рассматриваемого числа из перечня — 1, то пишем 9 на первом месте после разделительной точки (запятой);
   - Если первая цифра десятичного представления рассматриваемого числа — не 1, то пишем 1 на первом месте после разделительной точки (запятой).
2. Берем второе число в исходном перечне и определяем вторую цифру в новом числе:
   - Если вторая цифра десятичного представления рассматриваемого числа — 1, то пишем 9 на втором месте после разделительной точки (запятой);
   - Если вторая цифра десятичного представления рассматриваемого числа — не 1, то пишем 1 на втором месте после разделительной точки (запятой).
3. Построение нового числа продолжается путем изменения *n*-ой цифры *n*-го числа в исходном списке аналогичным образом:
   - Если *n*-я цифра десятичного представления рассматриваемого числа — 1, то пишем 9 на *n*-м месте после разделительной точки (запятой);
   - Если *n*-я цифра десятичного представления рассматриваемого числа — не 1, то пишем 1 на *n*-м месте после разделительной точки (запятой).

$$\begin{array}{lcl}
1 & \longleftrightarrow & 0.\;1\;\;1\;\;1\;\;1\;\;1\ \ldots \\
2 & \longleftrightarrow & 0.\;4\;\;0\;\;1\;\;2\;\;5\ \ldots \\
3 & \longleftrightarrow & 0.\;5\;\;6\;\;7\;\;1\;\;2\ \ldots \\
4 & \longleftrightarrow & 0.\;6\;\;0\;\;3\;\;0\;\;4\ \ldots \\
5 & \longleftrightarrow & 0.\;7\;\;9\;\;8\;\;9\;\;6\ \ldots \\
\ldots & & \\
\hline
 & & \mathbf{0.\;9\;\;1\;\;1\;\;1\;\;1\ \ldots}
\end{array}$$

Диагональный метод доказательства того,
что множество действительных чисел несчетно

Новое число будет отличаться по крайней мере одним десятичным знаком от каждого действительного числа в исходном перечне, и, вместе с тем, оно будет действительным числом, расположенным в промежутке от 0 до 1. Таким образом, предположение, что действительные числа можно взаимно однозначно сопоставить с натуральными числами, приводит к противоречию, а потому должно быть отброшено. Обратите внимание, что при доказательстве мы просматривали цифры чисел в перечне по диагонали. Поэтому метод, которым был получен данный результат, называют *диагональным*.

3.4.2. Множество всех подмножеств данного множества

Профессор. Ранее мы рассматривали операции (объединение, пересечение, вычитание), с помощью которых можно было из одних множеств получать другие множества *(см. разд. 3.2)*. Новые множества могут быть построены и другими способами. Так, мы можем рассмотреть множество, составленное из всех его подмножеств. Пусть, например, дано множество $A=\{a,b,c\}$ из трех элементов. Тогда множество всех его подмножеств (обозначим его как 2^A) состоит из восьми элементов:

$2^A=\{$

$\{a,b,c\}$ — само множество A,

$\{a,b\},\{a,c\},\{b,c\}$,

$\{a\},\{b\},\{c\}$,

$\varnothing$ — пустое множество

$\}$

Обратите внимание, что элементами множества 2^A являются множества. Если множество A состоит из n элементов, то множество 2^A всех его подмножеств состоит из 2^n элементов. Если обозначить количество элементов множества A как $|A|$, то $|2^A|=2^{|A|}$. В частности, пустое множество не имеет элементов ($|\varnothing| = 0$), поэтому $2^{Ш}=\{Щ\}$, $|2^{Ш}|=2^{|Щ|}=2^0=1$. Я надеюсь, что вы понимаете, в чем состоит различие между 2^A и $2^{|A|}$. Множество 2^A еще называют *булеаном* множества A.

Очевидно, что множество всех подмножеств *конечного* множества всегда больше данного множества: $|2^A|>|A|$. А выполняется ли это неравенство в случае бесконечных множеств?

З а н у д а. Сразу и не скажешь. Надо проверить, возможно ли взаимно однозначное соответствие между множествами A и 2^A. Но как его построить?

П р о с т а к. Быть может, следует попытаться доказать неравенство $|2^A|>|A|$ методом от противного, чтобы не заниматься взаимно однозначным соответствием в явной форме?

П р о ф е с с о р. Это хорошая идея. Попробуйте реализовать ее.

П р о с т а к. Пусть это сделает Зануда со всей присущей ему тщательностью.

З а н у д а. Хорошо, я попытаюсь. Итак, теорему сформулируем следующим образом: любое множество A неравномощно множеству 2^A всех его подмножеств.

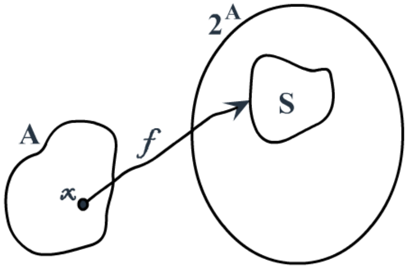
Доказывать будем методом от противного, т. е. предположим, что A равномощно множеству 2^A. Но это означает, что существует взаимно однозначное соответствие между элементами множества A и элементами множества 2^A. Здесь я позволю себе ввести несложные символические обозначения. Пусть $f: A \longrightarrow 2^A$ — указанное взаимно однозначное соответствие; $f(x)$ — подмножество множества A, которое соответствие f сопоставляет элементу x. Надеюсь понятно, что $x \in A$, $f(x) \in 2^A$.

Очевидно, что каким бы ни было взаимно однозначное соответствие f, для любого элемента $x \in A$ возможны два варианта:

- $x \in f(x)$ — элемент принадлежит сопоставляемому множеству;
- $x \notin f(x)$ — элемент не принадлежит сопоставляемому множеству.

Вот здесь начинается самое интересное. Рассмотрим подмножество S множества A всех тех элементов x, для которых $x \notin f(x)$. Не исключено, что множество S пусто, но это не имеет значения. Так как соответствие f взаимно однозначно, существует элемент $x \in A$, для которого $f(x) = S$. Спрашивается, что имеет место: $x \in S$ или $x \notin S$?

Допустим сначала, что $x \in S$. Но тогда $x \notin f(x)$, а $f(x) = S$ и, следовательно, $x \notin S$. Получаем противоречие. Предположим обратное: $x \notin S$. Но тогда $x \in f(x) = S$ и опять получаем противоречие. Другими словами, с одной стороны $x \in S$, а с другой — $x \notin S$. Данное противоречие мы получили в предположении, что множества A и 2^A равномощны. Отсюда следует, что предположение было неверным. Значит, множества A и 2^A не равномощны. На этом доказательство заканчивается.



Соответствие *f* сопоставляет элементу *x* множества *A*
некоторое его подмножество *S*

П р о с т а к. Как только Зануда ввел в рассмотрение подмножество S из всех таких элементов, что $x \notin f(x)$, я сразу понял, что следует ожидать противоречия. Это подмножество определяется таким же способом, как и множество гетерологических прилагательных, или множество тех, кого должен брить брадобрей. Только в данном случае противоречивость такого множества сыграла нам на руку, а не просто обескуражила.

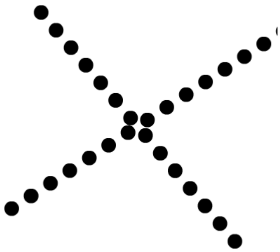
П р о ф е с с о р. Как мы уже знаем, множества натуральных, целых и рациональных чисел счетны, а множество действительных чисел — нет. Теперь мы узнали, что множество всех подмножеств данного множества больше последнего. Например, множество всех подмножеств натуральных чисел больше множества всех рациональных чисел.

3.4.3. Прямая и плоскость

П р о ф е с с о р. Мы знаем со времен французского мыслителя и математика Рене Декарта (лат. Cartesius — Картезий), привнесшего числа и алгебру в геометрию, что каждой точке прямой можно сопоставить число — координату этой точки. Пусть дан отрезок прямой, левому концу которого соответствует число 0, а правому — 1. Внутренним точкам данного отрезка взаимно однозначно сопоставлены числа в промежутке от 0 до 1 в их естественном порядке. Но какие числа? Хватит ли для этой цели только рациональных чисел (т. е. дробей)?

П р о с т а к. Любой отрезок прямой состоит из бесконечного количества точек. Точка не имеет протяженности. Если бы мы взяли лишь конечное их количество на отрезке прямой, то между ними образовались бы промежутки. С увеличением количества точек ширина промежутков между ними уменьшалась бы, оставаясь все же больше нуля. Но тогда возможен случай, когда два отрезка пересекаются не в точке, а в промежутке, что противоречит аксиоме евклидовой геометрии, согласно которой две пересекающиеся прямые имеют одну и только одну общую точку.

Рене Декарт (1596—1650)



При недостаточной плотности точек пересекающиеся отрезки могут не иметь общей точки

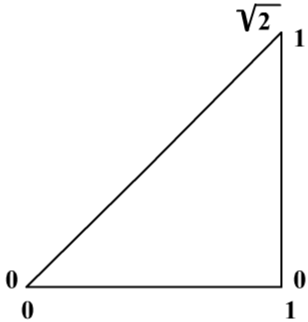
З а н у д а. То, что точек в отрезке конечной длины бесконечно много, понятно и так.

П р о с т а к. А я и не собирался это доказывать. Просто я хотел наглядно показать, что может быть, если множество точек не достаточно плотно.

З а н у д а. Физическая линия, например, след карандаша на бумаге, под лупой с достаточным увеличением предстанет как набор дискретных пятен. Математическая же линия и под микроскопом любой силы будет выглядеть так же, как и без него. Впрочем, математическую линию мы видим не глазами, а умом. Плотность точек на ней столь велика, что между ними нет никаких промежутков. Не пойму, куда ты клонишь?

П р о с т а к. Минуточку терпения, Зануда. Допустим теперь, что каждой точке отрезка прямой можно взаимно однозначно сопоставить рациональные числа из подмножества, например, от 0 до 1. Возьмем два таких отрезка и используем их в качестве катетов прямоугольного треугольника. Заметьте, что концевые точки гипотенузы являются одновременно и концевыми точками катетов, на которые гипотенуза опирается. Другими словами, вершины треугольника это общие точки гипотенузы и смежных с ней катетов. Мы не можем удалить концевую точку ни у гипотенузы, ни у катета, ибо вершина треугольника это, по определению, точка пересечения сторон треугольника. Длины катетов равны 1, а гипотенуза, согласно теореме

Пифагора, имеет длину $\sqrt{2}$. Это число, сопоставленное концевой точке гипотенузы, не является рациональным, т. е. оно не может быть получено делением одного целого числа на другое целое число. Следовательно, рациональных чисел недостаточно, чтобы можно было установить взаимно однозначное соответствие между ними и точками отрезка прямой.



Рациональных чисел недостаточно, чтобы можно было установить взаимно однозначное соответствие между ними и точками отрезка прямой

П р о ф е с с о р. Будем считать Ваши рассуждения, Простак, не строгим доказательством, а лишь проясняющими существо дела. Однако я должен сделать одну важную оговорку. В начале своих рассуждений, уважаемый Простак, Вы заявили, что любой отрезок прямой состоит из бесконечного количества точек. Отсюда можно было бы заключить, что вся прямая тоже состоит из бесконечного числа точек. Мне не нравится здесь употребление слова "состоит". Прямая не состоит из точек. Дело в том, что прямую, как и любой ее отрезок, можно бесконечно делить на отрезки все меньшей и меньшей длины. Длина таких отрезков стремится к нулю, но никогда не достигает его. А точка неделима и тем она отличается от любого, сколь угодно короткого отрезка. Не может быть половины точки. Так считали еще древние греки. Хотя точка не имеет протяженности, она все же не является отрезком нулевой длины. С другой стороны, отрезок нулевой длины это и не точка, и не отрезок, а некий образ, возникающий в нашем воображении как предельный результат бесконечного стягивания концов отрезка (концевых точек) друг к другу. Точки, прямые и плоскости являются исходными сущностями геометрии. Так, Д. Гильберт в своей книге "Основания геометрии" начинает с того, что вводит в рассмотрение три системы вещей, которые предлагает называть точками, прямыми и плоскостями. Мы не знаем, что это такое, и можем претендовать лишь на то, чтобы усвоить относящиеся к ним аксиомы, констатирующие отношения между введенными вещами, как, например, "Две различные точки определяют прямую". Кстати, можно договориться формулировать эту аксиому и иначе: "Прямая проходит через две точки" или "Точки расположены на прямой".

Давид Гильберт (1862—1943)

Итак, прямая не состоит из точек, а содержит точки, через которые она проходит. Точек на любом отрезке прямой бесконечно много и больше, чем рациональных чисел. В действительности их столько же, сколько действительных чисел — рациональных и иррациональных вместе взятых. Это означает, что между этими двумя множествами существует взаимно однозначное соответствие.

> Говорят, что множества точек прямой (вообще любой линии) и действительных чисел образуют *континуум* — бесконечную и непрерывную совокупность. Между любыми двумя сколь угодно близкими точками линии находится бесконечно много точек, а между любыми двумя действительными числами находится бесконечно много действительных чисел.

А что вы скажете о возможности взаимно однозначного соответствия между множествами всех точек прямой и плоскости, например, между точками отрезка единичной длины и квадрата со стороной, равной 1?

П р о с т а к. Мне кажется, что в данном случае взаимно однозначного соответствия быть не может, поскольку точек в квадрате явно больше, чем на прямой. Можно представить себе квадрат как бесконечно много отрезков параллельных прямых, расположенных один подле другого и покрывающих всю площадь квадрата. При этом мощность множества таких отрезков должна быть такой же, как и мощность множества точек любого из них.

А эта мощность, как нам уже известно, равна мощности множества всех действительных чисел.

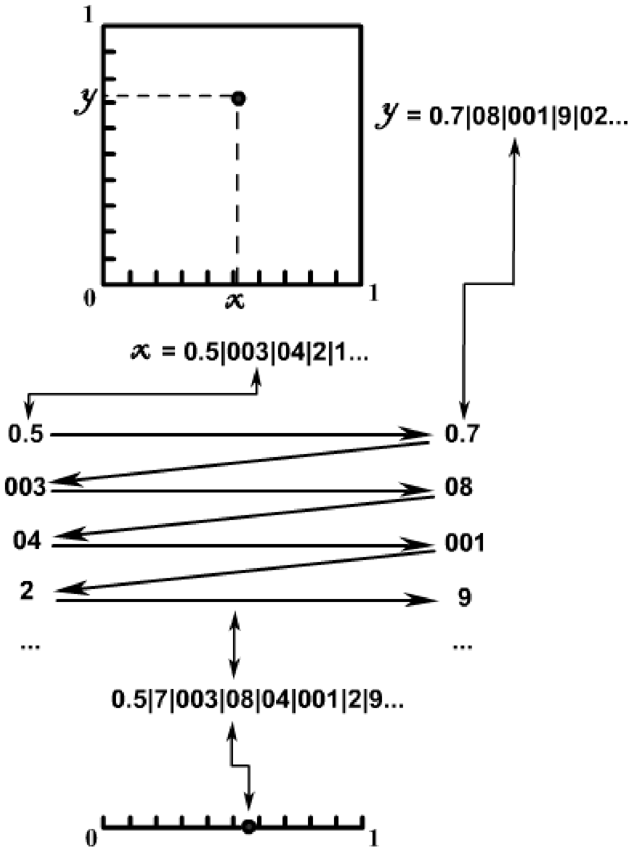
Зануда. Аналогия с покрытием квадрата очень узкими полосами лишь делает наглядными трудности построения или хотя бы выяснения возможности взаимно однозначного соответствия между точками квадрата и отрезка прямой. Интуиция подсказывает мне, что Простак скорее прав, чем не прав. Однако, имея дело с бесконечными множествами, мы должны быть готовы ко всему.

Профессор. Оказывается, точки квадрата *можно* отобразить на точки отрезка прямой взаимно однозначным образом. Доказав это в 1877 г., Кантор сам был чрезвычайно удивлен полученным результатом: "Я вижу это, но никак не могу этому поверить!" Для большинства математиков это было настоящей сенсацией, а немецкий математик Л. Кронекер вообще не принял доказательства. Кронекер был известен очень высокими требованиями к строгости определений математических объектов. Считая, что "Бог создал целые числа, а все остальное — дело рук человеческих", он, в частности, не признавал существования иррациональных чисел. Так, число π он представлял бесконечной суммой рациональных чисел. Известно несколько вариантов представления π таким образом — рядом. Например Лейбниц в XVII веке предложил такую формулу: $\pi/4 = 1 - 1/3 + 1/5 - 1/7 + ...$.

Леопольд Кронекер (1823—1891)

Я лишь кратко поясню идею доказательства Кантора возможности взаимно однозначного соответствия между точками плоскости и прямой. Каждая точка плоскости представляется парой чисел — координатами *x*, *y*, которые можно представить бесконечными десятичными дробями. Эти две дроби комбинируются по определенному правилу, чтобы получить одну дробь, которая сопоставляется с точкой на прямой. Данная процедура обратима и, следовательно, она устанавливает взаимно однозначное соответствие между точками плоскости и прямой.

Комбинация двух десятичных дробей, представляющих точку на плоскости, производится следующим образом. Цифры дроби последовательно разбиваются на группы. Каждая цифра, кроме 0, начинает новую группу. Бесконечная дробь, соответствующая точке на прямой, составляется из полученных групп цифр: первая группа из числа *x*, первая группа из числа *y*, вторая группа из числа *x*, вторая группа из числа *y* и т. д.



Установление взаимно однозначного соответствия
между точками плоскости и прямой

3.5. Парадоксы бесконечности

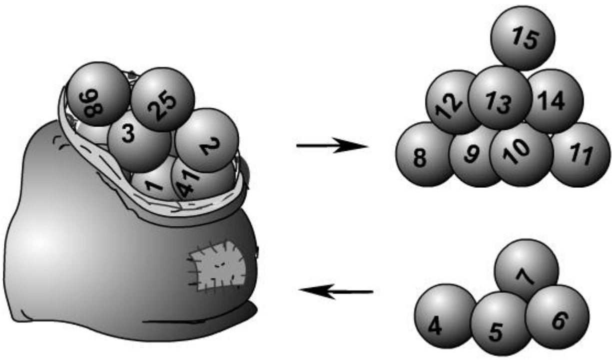
3.5.1. Не оскудеет рука дающего

П р о ф е с с о р. Представим себе мешок с бесконечным количеством шаров, пронумерованных натуральными числами. В течение одной минуты шары вынимаются из мешка и возвращаются обратно по следующему алгоритму:

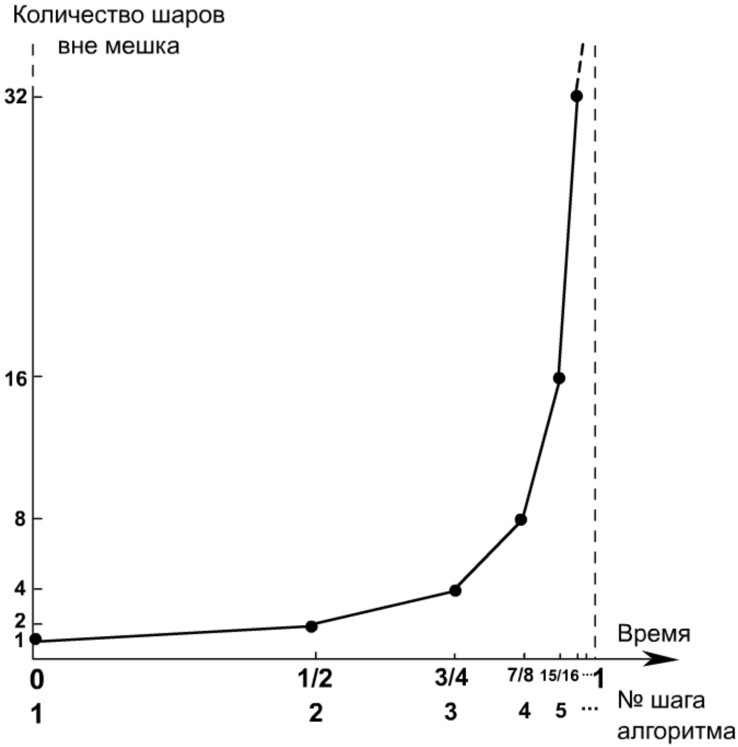
1. За одну минуту до окончания данного алгоритма из мешка вынимается шар с номером 1.
2. Через 1/2 минуты после шага 1 из мешка вынимаются шары с номерами 2 и 3, а шар 1 возвращается в мешок.
3. Через 1/4 минуты после шага 2 из мешка вынимаются шары с номерами 4, 5, 6, 7, а шары 2 и 3 возвращаются.
4. Через 1/8 минуты после шага 3 из мешка вынимаются шары с номерами 8, 9, …, 15, а шары 4, 5, 6, 7 возвращаются. Следующие шаги данного алгоритма выполняются аналогичным образом.

Нетрудно заметить, что в каждый момент из мешка извлекается в два раза больше шаров, чем возвращается назад. Сколько шаров окажется вне мешка ровно через минуту после начала работы описанного алгоритма?

З а н у д а. На первом шаге, в самом начале минуты, вне мешка был один шар. На втором шаге, через 1/2 минуты, — 2 шара, на третьем, еще через 1/4 минуты, — 4 шара. На n-м шаге, через $1/2^{n-1}$ минуты после предыдущего шага или за это же время до окончания минутного срока, вне мешка будет 2^{n-1} шара.



В каждый момент из мешка извлекается вдвое больше шаров, чем возвращается обратно



Зависимость количества шаров, находящихся вне мешка,
от времени и номера шага алгоритма

Очевидно, что необходимо сначала определить, сколько шагов будет пройдено в течение всей минуты. Первый шаг был сделан в начале минуты, второй — через 1/2 (т. е. $1/2^1$) минуты, третий — через $1/2^1 + 1/2^2$ минуты, четвертый — через $1/2^1 + 1/2^2 + 1/2^3$, n-й шаг был сделан через $1/2^1 + 1/2^2 + 1/2^3 + \ldots + 1/2^{n-1}$ минуты после начала. Очевидно, что количество шагов n, выполненное за одну минуту, можно найти из уравнения:

$$1/2^1 + 1/2^2 + 1/2^3 + \ldots + 1/2^{n-1} = 1.$$

Правда, я не знаю, как его решить, но это, я думаю, трудность технического характера, которую можно преодолеть с помощью, например, математического справочника.

Простак. Чтобы не искать справочник, давайте попробуем составить другое уравнение. А именно попытаемся определить, сколько времени еще осталось до истечения минуты на n-м шаге алгоритма. На первом шаге остается еще целая минута, на втором — $1/2 = 1/2^1$ минуты, т. к.

этот шаг был сделан через полминуты после начала работы. На третьем шаге оставалось $1/4 = 1/2^2$ минуты, а на n-м шаге — $1/2^{n-1}$ минуты. Чтобы определить номер шага, на котором минута будет полностью исчерпана, достаточно решить очень простое уравнение:

$$1/2^{n-1} = 0.$$

Теперь мы хорошо видим, что ни при каком конечном числе n это равенство не выполняется точно, но в то же время замечаем, что с ростом n левая его часть очень быстро приближается к 0. Можно сказать, что при $n \to \infty$ величина $1/2^{n-1}$ равна нулю со сколь угодно большой точностью. Например, на 10-м шаге работы алгоритма до окончания минуты останется меньше 0,12 секунды, а на 25-м шаге — меньше 0,000 004 секунды. Это я подсчитал с помощью калькулятора.

З а н у д а. Но, насколько я помню, в задаче спрашивалось, сколько шаров окажется вне мешка ровно через минуту. На n-м шаге работы алгоритма это количество, как мы уже видели, равно 2^{n-1}. На любом шаге количество шаров вне мешка конечное, хотя и увеличивается с огромной скоростью с каждым новым шагом. Вместе с тем, минута будет исчерпана при $n \to \infty$ и, следовательно, вне мешка окажется бесконечно много шаров.

П р о с т а к. Не кажется ли вам, что здесь намечается противоречие: на каждом шаге количество шаров вне мешка конечно, но когда минута будет исчерпана, это количество вдруг станет бесконечным. При этом, однако, и внутри мешка будет также бесконечно много шаров.

З а н у д а. А почему это тебя удивляет? Нас же теперь не шокирует тот факт, что если из множества всех целых чисел удалить все только четные числа, коих бесконечно много, то оставшихся чисел по-прежнему будет бесконечно много.

П р о с т а к. Я не это имел ввиду, просто неудачно выразился. Меня удивляет, что по истечении минуты любой конкретный шар окажется в мешке, несмотря на то, что вне мешка будет бесконечно много шаров. Например, где окажется k-й шар в момент истечения минуты?

В момент, сколь угодно близкий к завершению минуты, шары 1, 2, 3 и так далее до, возможно, какого-то k-го шара уже будут в мешке. Особым является момент извлечения/возвращения шаров, когда и решается судьба дальнейшего пребывания k-го шара. Давайте рассмотрим его подробнее.

Можно заметить, что на n-м шаге работы алгоритма из мешка извлекаются шары с номерами:

$$2^n, 2^n + 1, \ldots, 2^{n+1} - 1,$$

а назад возвращаются (при $n > 1$) шары с номерами:

$$2^{n-1}, 2^{n-1}+1, \ldots, 2^n-1.$$

Для любого номера k найдется шаг n алгоритма такой, что

$$2^{n-1} \leq k \leq 2^n - 1$$

и, следовательно, на этом шаге k-й шар окажется в мешке. Это произойдет до истечения минуты или в момент, сколь угодно близкий к ее окончанию. Таким образом, получается, что любой конкретный шар (с любым заданным номером) к моменту окончания минуты будет в мешке.

З а н у д а. Однако к этому же моменту, как мы убедились чуть ранее, вне мешка будет находиться бесконечно много шаров. Выходит, это множество содержит неконкретные шары? Парадокс!

П р о ф е с с о р. Вы оба рассуждали довольно разумно, но упустили из виду важное обстоятельство, что и привело к противоречию. Обратите внимание, что алгоритм, состоящий из бесконечного количества извлечений/возвращений шаров, привязанных к моментам времени внутри минуты, не определен для последнего момента этой минуты. Данный алгоритм состоит из бесконечного количества шагов, но продолжительность его работы меньше одной минуты, хотя и сколь угодно близка к этой величине. Действительно, время работы алгоритма определяется суммой бесконечного количества временны́х интервалов между его шагами — $1/2^1 + 1/2^2 + 1/2^3 + \ldots + 1/2^{n-1} + \ldots$. Для любого, сколь угодно большого числа шагов эта сумма меньше 1. Эту единицу можно рассматривать лишь как тот предел, к которому сумма постоянно приближается с добавлением каждого нового члена, но ни при каком конечном числе всех членов не достигает его. Последний момент минуты не принадлежит временно́му интервалу работы алгоритма, на котором он определен, а, значит, мы не можем сказать, что он сотворит в этот предельный для него момент. В любой момент времени, не равный окончанию минуты, вне мешка будет находиться конечное количество шаров, а внутри мешка — бесконечное число шаров, что ведет к противоречию, которое исчезает, если выбросить из рассмотрения этот момент.

П р о с т а к. Тем не менее, мы можем предсказать, что будет в момент, сколь угодно близкий к концу минуты.

П р о ф е с с о р. Разумеется, поскольку для такого момента всегда найдется интервал, концы которого привязаны к двум смежным шагам алгоритма, и который содержит этот момент.

П р о с т а к. Кажется, я начинаю понимать. Но как Вы объясните следующее обстоятельство. С одной стороны, алгоритм выполняет бесконечное количество шагов, т. е. никогда не завершает своей работы. С другой стороны, длительность его работы не может превысить одной минуты. Что же мы будем реально наблюдать в момент окончания минуты и далее? С одной стороны, спустя минуту и более алгоритм должен продолжать работу, т. к. в противном случае он не содержал бы бесконечное количество шагов. Другими словами, мы не можем сказать, что спустя минуту и более алгоритм перестал работать. Но, с другой стороны, он не определен на интервале времени, начинающемся с конца рассматриваемой минуты. Так что же он делает в это время, работает или нет? Словом, я не могу представить себе эту противоречивую ситуацию. Моя голова просто "пухнет".

Профессор. Противоречивую ситуацию всегда трудно себе представить наглядно, на то она и противоречивая. Поэтому математики и считают противоречивые объекты несуществующими.

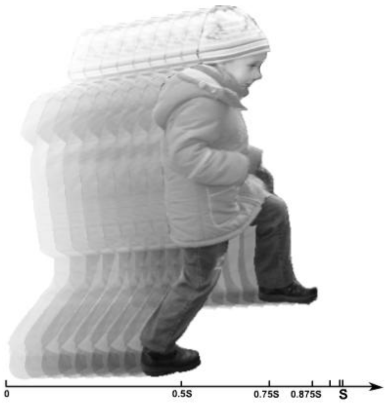
В данном примере сталкиваются две точки зрения на бесконечность. С одной стороны, бесконечность актуальна: в мешке находится бесконечно много шаров, мы их созерцаем умом как нечто данное в законченном виде. С другой стороны, бесконечность потенциальна: алгоритм выполняет шаги, один за другим, никогда не останавливаясь (т. е. сколь угодно долго); мы можем точно сказать, что происходит на каждом его конкретном шаге, но не можем отчетливо предсказать ситуацию на предельном шаге, потому что его не существует для алгоритма; предельный шаг запределен для алгоритма (потенциальной бесконечности шагов); последняя фраза и выражает противоречие между двумя пониманиями бесконечности почти буквально.

В качестве сеанса успокоительной психотерапии предложу вам следующее объяснение. Если наблюдатель пристально следит за действиями алгоритма извлечения/возвращения шаров, фиксируя своими чувствами каждый его шаг, то он начинает жить как бы в собственном времени этого алгоритма. Для такого наблюдателя алгоритм работает вечно, никогда не останавливаясь. Теперь представим себе наблюдателя, который следит за первым наблюдателем, а вместе с ним и за алгоритмом. Этот наблюдатель живет в своем, если угодно, обычном времени, и для него первый наблюдатель, как и алгоритм, исчезают из поля зрения, спустя ту самую минуту, в течение которой они существуют. В дальнейшем второй наблюдатель просто их не воспринимает. Впрочем, и первый наблюдатель теряет из виду второго, как только тот пересекает запредельную для него метку окончания той самой замечательной минуты. Оба наблюдателя могут заметить друг друга только в течение этой минуты, но не в момент ее окончания и не во все последующие моменты.

3.5.2. Дорогу осилит идущий

П р о ф е с с о р. Парадоксы, аналогичные только что рассмотренному, были известны древним грекам 25 веков тому назад. Зенон Элейский предложил так называемый парадокс "Стадии", в котором бегун якобы не может преодолеть конечную дистанцию, поскольку она состоит из бесконечного количества отрезков пути. Так, сначала ему следует пробежать 1/2 всей дистанции, затем 1/2 оставшейся половины и т. д. Таким образом, вся дистанция S разбивается на бесконечное количество стадий уменьшающейся длины: $S/2$, $S/4$, $S/8$, …, $S/2^{n-1}$, … . Сумма длин всех отрезков сколь угодно близко приближается к S при их количестве n, стремящемся к бесконечности, но не достигает величины S в точности ни при каком сколь угодно большом n.

Другой аналогичный парадокс Зенона называется "Ахилл и черепаха". В начальный момент времени между Ахиллом и черепахой некоторое расстояние. Быстроногий Ахилл якобы не может догнать медлительную черепаху, поскольку когда он добежит до места, с которого стартовала черепаха, она успеет отползти на некоторое расстояние. Данная ситуация повторяется сколько угодно раз.



Конечную дистанцию можно разбить на бесконечное количество конечных отрезков

П р о с т а к. Оба парадокса разрешаются очень просто. Дистанция, которую должен преодолеть бегун, конечна. Бегун движется с постоянной скоростью. Следовательно, время, через которое он достигнет цели, тоже конечно. Ахилл, перемещаясь равномерно с большей скоростью, чем черепаха, в конце концов догонит и перегонит ее. А представление Зенона о том, что конечные пространственные и временны́е интервалы можно разбить на бесконечное количество конечных интервалов, вовсе ни при чем и только затемняет существо дела.

П р о ф е с с о р. Зенон не был идиотом и не имел целью дурачить своих сограждан. Его парадоксы были придуманы, чтобы вскрыть противоречия некоторых давних интуитивных представлений о бесконечно малых и бесконечно больших величинах. Ранее считали, что сумму бесконечно многих крайне малых величин ε можно сделать сколь угодно большой ($\infty \times \varepsilon = \infty$), а сумма бесконечного или конечного количества величин нулевого размера равна нулю: ($\infty \times 0 = 0, n \times 0 = 0$). Зенон с помощью своих парадоксов выступил с критикой таких представлений. Они сформулированы так, чтобы подчеркнуть противоречия в понятиях движения и времени. Проблемы, приведшие к парадоксам Зенона, постоянно возникают в философских дискуссиях, и они связаны с соотношением потенциальной и актуальной бесконечностей.

До появления канторовской теории актуально бесконечных множеств бесконечность воспринималась только как нечто потенциальное, как *возможность* ничем не ограниченного наращивания какой-либо совокупности за счет новых членов. Эта совокупность всегда остается незавершенной, а потому не может рассматриваться как нечто целое. Множество по Кантору, напротив, завершенное и потому является целостностным образованием. С точки зрения теории актуально бесконечных множеств, ни пространственные области, ни временны́е интервалы недопустимо рассматривать как состоящие из бесконечно большого количества дискретных, изолированных друг от друга элементов. Рассуждения Зенона как раз и были направлены против представления учениками Пифагора пространства как суммы дискретных и неделимых точек.

3.6. Аксиомы теории множеств

П р о ф е с с о р. Теория множеств, построенная Кантором, является наивной. Это означает лишь то, что главный предмет этой теории, множество, не имеет формального определения. Фраза *"Множество — это собрание определенных и различных между собой объектов интуиции или интеллекта, мыслимое как единое целое"* не является формальным определением множества. Она служит только для того, чтобы наше подсознание приняло в свой обиход некий образ, а сознание — соответствующие термины и способы их употребления в речи. Проще говоря, неформальные определения нужны для начальной ориентации нашей интуиции в предлагаемом пространстве идей, чтобы мы могли понять, хотя бы смутно, о чем идет речь, попросту говоря, "въехать в тему". При этом, однако, предполагается, что изучаемый предмет, нечетко указанный неформальным определением, будет впоследствии, по мере развития теории о нем, уточнен и в конце концов станет достаточно определенным.

Парадоксы, обнаруженные в канторовской теории множеств, заставили ряд выдающихся математиков уточнить понятие множества на основе его аксиоматического определения. Аксиомы должны были быть сформулированы таким образом, чтобы, с одной стороны, они были интуитивно истинными, а с другой стороны, позволили бы избежать всех известных парадоксов. Отправной идеей было то, что не любая совокупность отличающихся друг от друга элементов может образовывать множество. Предполагалось, что явно сформулированные аксиомы позволили бы прояснить, что следует понимать под множеством и какими свойствами оно должно обладать.

Обычно считается, что аксиоматическое определение чего-либо не использует никаких интуитивных образов определяемого объекта. Этот образ должен появиться у нас постепенно, приобретая более или менее ясные очертания по мере рассмотрения всех аксиом, а также в результате логического вывода из них всевозможных следствий (теорем). Так например, мы можем дать аксиоматическое определение блям-блямчиков, не навязывая при этом никаких интуитивных образов, которые якобы должны возникать, как только мы встретимся со словом "блям-блямчик". Что же это в действительности, определяется исключительно с помощью свойств блям-блямчиков и их отношений между собой и с другими объектами, сформулированных явным образом в виде аксиом. Разумеется, в действительности создатель аксиоматической системы, начиная с чистого листа, не может не иметь интуитивного представления об определяемом объекте, но, тем не менее, он формулирует свои аксиомы так, будто этого представления у него нет.

Сначала мы рассмотрим один из первых вариантов аксиоматического определения множества, предложенный немецким математиком Э. Цермело в 1908 г. Чтоб хотя бы частично подавить интуитивный образ канторовского понятия множества, вместо слова "множество" будем использовать в формулировках аксиом английское слово Set, а вместо "подмножество" — Subset.

В своей теории множеств Э. Цермело начинает с того, что рассматривает область каких-то объектов. Может статься, что между двумя из этих объектов x и y существует отношение $x \in y$. В таком случае мы говорим, что x есть элемент y, а y есть Set. Понятно, что это является не определением, а соглашением об употреблении слов и символов. В самом деле, если мы не знаем, что такое Set, то мы не узнаем ничего нового, пока нам не скажут, что такое $\in$. Таким образом, основными неопределяемыми понятиями здесь являются объекты и отношение $\in$ между ними. В любой аксиоматической системе без неопределяемых понятий не обойтись: ведь аксиомы необходимо через что-то выразить. Далее перечислим семь аксиом Цермело, определяющих Set:

1. Два Set, имеющие одни и те же элементы, тождественны (равны).

2. Существует одно Set, не содержащее ни одного элемента, — пустое Set Ø. Если существует один объект a, то существует Set $\{a\}$, единственным элементом которого является этот объект. Если существуют два объекта a и b, то существует Set $\{a, b\}$ как неупорядоченная пара этих двух объектов.
3. Совокупность всех элементов какого-либо Set A, удовлетворяющих условию P, образует Subset A (образование подмножеств).
4. Каждому Set A соответствует другое Set UA, образованное из всех Subset A (образование множества всех подмножеств данного множества).
5. Пусть T есть Set, все элементы которого сами являются Set. Тогда существует Set ST, элементы которого являются элементами элементов T. Например, если T имеет три элемента A, B и C, которые сами являются Set, и если $A = \{a_1, a_2\}$, $B = \{b_1, b_2\}$, $C = \{c_1, c_2\}$, то ST будет иметь шесть элементов: $\{a_1, a_2, b_1, b_2, c_1, c_2\}$.
6. Пусть T есть Set, все элементы которого сами являются Set. Тогда можно выбрать в каждом из этих элементарных Set по одному элементу, и совокупность всех таких элементов образует Subset ST.
7. Существует Set A, которое не может содержать элемент a, не включая в себя одновременно в качестве элемента и Set $\{a\}$.

Первые шесть аксиом можно рассматривать как очевидные, если слову *Set* придать интуитивный смысл слова "множество" и, кроме того, рассматривать только конечное количество элементов. Однако седьмая аксиома утверждает существование бесконечных множеств. Поэтому ее называют *аксиомой бесконечности*. Цермело сформулировал ее столь изощренно, чтобы избежать слова "бесконечный" и, таким образом, рассматривать свои аксиомы как предшествующие различению конечного и бесконечного. Согласно седьмой аксиоме, если A содержит элемент a, то оно будет содержать еще и другие элементы, а именно: Set, единственным элементом которого является a, Set, единственным элементом которого является Set, единственным элементом которого является a, и т. д. Иначе говоря, существует Set A такое, что если a принадлежит A, то $\{a, \{a\}\}$ принадлежит A, $\{a, \{a\}, \{a, \{a\}\}\}$ принадлежит A и т. д. Очевидно, что количество таких элементов бесконечно.

Простак. Аксиомы Цермело это правила, указывающие, когда определенные совокупности объектов образуют Set. Кроме того, эти аксиомы можно рассматривать просто как правила расширения смысла слова Set. Аксиомы определяют, как из некоторых элементов и их совокупностей образовывать то, что называется Set и под чем подразумевается множество. Возникает подозрение, что не любая совокупность элементов может быть Set.


Профессор. Обратите внимание, что среди аксиом нет ни одной, которая бы явно утверждала, что какие-то совокупности не являются Set. Если бы такая аксиома была сформулирована, то появилась бы опасность получить противоречие, т. е. невозможность существования объекта, удовлетворяющего одновременно всем аксиомам. В то же время Цермело отбросил восьмую аксиому, допускающую любые множества: *какие бы то ни было объекты образуют Set*. Сделал он это во избежание ранее обнаруженных противоречий (парадоксов). Цермело запретил рассматривать множества всех объектов, обладающих некоторым свойством, но позволил рассматривать подмножества существующих множеств, элементы которых обладают определенным свойством.

Шестая аксиома Цермело называется *аксиомой выбора*. Она достаточно очевидна для конечных множеств, но не является таковой в случае бесконечных множеств. Применение этой аксиомы оправдывается тем, что доказательства огромного числа математических теорем оказываются весьма изящными и, наоборот, будут громоздкими и сложными при отказе от нее. Некоторые важные результаты вообще нельзя получить без аксиомы выбора. Интересно, что ранее математики использовали аксиому выбора подсознательно, но стоило ее сформулировать в явном виде, как сразу же возникли споры о правомерности ее применения. Великий немецкий математик Д. Гильберт в 1923 г. назвал аксиому выбора общим принципом, который необходим и неоценим в теории математического вывода. Первым на аксиому выбора обратил внимание Д. Пеано в 1890 г., но то, что она действительно является аксиомой, впервые понял Б. Леви в 1902 г., а Э. Цермело узнал об этом в 1904 г. Вместе с тем, явное применение Цермело аксиомы выбора вызвало негодование многих ведущих математиков.


Простак. А почему аксиома выбора неочевидна в случае бесконечных множеств? И почему она вызвала отрицательную реакцию, если она, пусть неявно, многократно использовалась ранее и была очень полезной? Ведь мы же иногда допускаем некоторые постулаты, исходя не столько из их самоочевидности, сколько из соображений технологического удобства.


Профессор. Противники аксиомы выбора считали, что если не указано правило, по которому из каждого множества выбирается по одному элементу, то в действительности выбор не может быть произведен и, следовательно, множество из таких элементов не может быть построено. Сказанное пояснил с помощью примера Б. Рассел в 1905 г. Предположим, что имеется 100 пар обуви и из каждой пары выбирается левый ботинок. Очевидно, что в данном случае правило выбора вполне определено. А теперь предположим, что имеется 100 пар носков и из каждой пары выбирается по

одному носку. В этом случае невозможно указать, какой носок, правый или левый, был выбран из каждой пары, поскольку они неразличимы. Таким образом, нельзя сформулировать правило выбора и, следовательно, построить с помощью этого правила множество представителей носков из каждой пары. Впрочем, для иллюстрации данной ситуации можно привести и более древнюю притчу о буридановом осле, который умер, не сумев решить проблему выбора одной из двух одинаковых охапок сена.

Теперь рассмотрим бесконечное множество. Допустим, что имеется столько пар ботинок, сколько есть целых чисел. Понятно, что здесь речь идет не о числе, а о том, что мы можем пронумеровать пары ботинок целыми числами, начиная с 1 и далее до бесконечности. Будет ли количество ботинок равно количеству пар?

Простак. Да, поскольку достаточно в n-ой паре левый ботинок обозначить числом $2n - 1$, а правый — числом $2n$.

Зануда. Это верно, если ботинки в паре отличаются друг от друга, что вполне естественно предположить. Однако если речь пойдет о парах носков или шнурков, в которых левый неотличим от правого, то мы не сможем выполнить их нумерацию.

Простак. Но мы можем просто взять любой из неразличимых предметов наудачу.

Профессор. Вот видите, Зануда испытал некое затруднение при выборе элемента, а Простак нашел некий способ его разрешения. Чтобы в подобных случаях гарантировать себе хотя бы принципиальную возможность выбора, не вдаваясь в детали его реализации, и нужна аксиома выбора.

Сторонники аксиомы выбора признавали, что правила выбора может и не быть, но не считали его наличие необходимым. По их мнению акты выбора определены просто тем, что их считают определенными. Наиболее стойким защитником аксиомы выбора, как и концепции актуально бесконечных множеств, был французский математик Ж. Адамар. По его мнению, для того чтобы утверждать существование объектов, совсем не требуется их описывать. Если одного лишь утверждения о существовании объекта достаточно для прогресса математики, то это утверждение приемлемо. Аксиому выбора, хотя и с некоторыми оговорками, принимал и А. Пуанкаре.

В середине 1930-х годов К. Гедель показал, что аксиому выбора можно рассматривать как истинную в ряде формальных систем теории множеств, т. е. она либо доказуема в таких системах, либо независима от остальных аксиом системы и при желании может быть добавлена к этой системе. В 1960 г. П. Коэн показал, что в обычных формальных системах теории множеств аксиому выбора можно считать не выполняющейся, не впадая при этом в противоречие. В 1966 г. он опубликовал доказательство того, что аксиома выбора независима от остальных аксиом системы Цермело—Френкеля. Это означает, что если отбросить аксиому выбора, то на основе оставшихся аксиом ее нельзя доказать как теорему. Это же эквивалентно тому, что существует теорема, которую нельзя доказать без аксиомы выбора. Таким образом, возможны формальные теории множеств как с аксиомой выбора, так и без нее. Эти теории не эквивалентны, но обе достаточно интересны. В настоящее время существует много различных формальных теорий множеств.

Жак Адамар (1865—1963)

Курт Гедель (1906—1978)

Система аксиом Цермело была усовершенствована в 1922 г. А. Френкелем. Эта новая система аксиом Цермело—Френкеля и в настоящее время широко используется специалистами по теории множеств, хотя существуют и другие (например, система Гильберта—Бернайса). Я приведу в словесной форме лишь некоторые аксиомы из этой системы:

1. Два множества тождественны (равны), если они состоят из одних и тех же элементов.
2. Существует пустое множество.
3. Если A и B — множества, то неупорядоченная пара $\{A, B\}$ тоже множество.
4. Объединение любого числа множеств есть множество.
5. Существует бесконечное множество (аксиома бесконечности).
6. Любое свойство, формализуемое с помощью языка теории, может быть использовано для определения множества.
7. Совокупность всех подмножеств данного множества есть множество. Процесс образования множества подмножеств можно повторять сколько

угодно раз, т. е. рассматривать совокупность всех подмножеств любого данного множества в качестве нового множества; совокупность подмножеств этого множества также является множеством и т. д.

8. Аксиома выбора. Пусть дано множество, все элементы которого сами являются множествами. Тогда можно выбрать в каждом из этих элементарных множеств по одному элементу, и совокупность всех таких элементов образует множество.
9. x не принадлежит x (т. е. $x \notin x$).

Замечательной особенностью аксиоматики Цермело—Френкеля является то, что она не допускает к рассмотрению множество, содержащее все множества, а также множества, содержащие самих себя в качестве элемента. Поэтому она, возможно, позволяет избежать парадоксов. Все знаменитые парадоксы, большинство из которых мы рассмотрели ранее, исключены. Другие же пока не обнаружены. Однако нет полной гарантии того, что никакие другие парадоксы не могут возникнуть. По этому поводу А. Пуанкаре метко заметил: "Мы возвели ограду вокруг стада, чтобы защитить его от волков, но неизвестно, нет ли волков внутри ограды".

Интересно, что аксиома бесконечности по существу определяет множество целых положительных чисел Z. В литературе для нее есть несколько различающихся формулировок. Одна из них имеет такой вид:

1. $\varnothing \in \mathrm{Z}$ (пустое множество $\varnothing$ существует согласно второй аксиоме).
2. Если $x \in \mathrm{Z}$, то также $\{x\} \in \mathrm{Z}$ — если элемент x принадлежит множеству Z, то и одноэлементное множество $\{x\}$ принадлежит Z в качестве элемента.

Для такого множества Z имеет место следующее:

$\{\varnothing\} \in Z$,

$\{\varnothing, \{\varnothing\}\} \in Z$,

$\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\} \in Z$

и так далее.

Положим $0 = \varnothing$, $1 = \{\varnothing\}$, $2 = \{\varnothing, \{\varnothing\}\} = \{0, 1\}$, …, $n = \{0, 1, 2, \ldots, n-1\}$. Тогда для любого $n \geq 0$ будет выполнено $n \in Z$ и при этом $0 \neq 1$, $0 \neq 2$, $1 \neq 2$, $0 \neq 3$, $1 \neq 3$, $2 \neq 3$, … . Таким образом, целое положительное число n есть множество $\{0, 1, 2, \ldots, n-1\}$, элементами которого являются числа 0, 1, 2, …, $n-1$; при этом число 0 является обозначением пустого множества $\varnothing$ (т. е. $0 = \varnothing$), число 1 — обозначение множества $\{\varnothing\}$, число 2 — обозначение множества $\{\varnothing, \{\varnothing\}\}$ и т. д.

3.7. От перечислимости и разрешимости множеств к алгоритмам

Профессор. В настоящее время решение многих насущных задач, в том числе неоспоримо интеллектуальных, поручается компьютерам. Казалось бы, то, что не подвластно интеллекту человека, можно поручить вычислительной машине. Оптимизация перевозок, стратегии банковских операций, вычисление какого бы то ни было значения по каким-то правилам, поиск всяческих минимумов или максимумов — вот примеры типичных вычислительных задач.

Поручить компьютеру решение какой-то задачи означает *автоматизировать* процесс ее решения. Чтобы сделать это, необходимо составить *алгоритм* решения данной задачи, а затем написать программу на одном из компьютерных языков (например, C++, Java, Pascal, JavaScript, Basic и т. п.), которую компьютер может понять (интерпретировать) и затем выполнить. Алгоритм, более или менее ясный для человека, должен быть достаточно определенным, чтобы его можно было однозначно перевести на другой язык — язык программирования, который может быть правильно воспринят компьютером. Известно, что не все практические задачи более или менее легко поддаются автоматизации. Может показаться, что трудности автоматизации обусловлены то ли несовершенством языков программирования, то ли неумением конкретных программистов. И то, и другое случается в повседневной жизни. Однако во многих сложных практических задачах есть нечто общее, что и получило общее название — *вычислимость*. Легко сказать: "Пусть имеется функция, которая …" Иногда мы не сомневаемся, что такая функция может существовать, даже если не знаем, как ее вычислить, но бывает и так, что вполне понятная по своему описанию функция не может быть вычислена никаким компьютером потому, что алгоритм ее вычисления не существует, т. е. не может быть изобретен в принципе. Иначе говоря, мы не знаем, что именно требуется сделать, чтобы получить конечный результат, называемый *значением* этой функции. В математике часто ограничиваются просто знанием того, что нечто существует, без того, чтобы выяснить подробно, как бы это могло осуществлено. *Существование* в классической математике означает непротиворечивость данной системе аксиом, формально описывающей область исследования. В математическом направлении, называемом *конструктивным* (Л. Брауэр, Г. Вейль и др.) или *конструктивизмом*, это не так. Для конструктивистов, чтобы объект существовал, недостаточно только согласованности его описания с аксиомами мира, объемлющего этот объект. Необходимо также указать способ его построения. Например, иррациональное число $\sqrt{2}$ для конструктивистов не существует в множестве действительных чисел, пока мы не укажем способ (алгоритм) его построения.

Перестройка математики в соответствии с философией конструктивизма обедняет математику в том смысле, что многие ее ценные результаты не вписываются (не выводимы) в систему аксиом, принимаемую в качестве истинной (допустимой) конструктивистами. Тем не менее, именно конструктивистам мы обязаны прояснением того, что такое существование в математике, вычислимость функций и, в конце концов, алгоритм.

В данном разделе мы рассмотрим понятия перечислимости и разрешимости множеств, на которых основаны понятия алгоритма и вычислимости.

Итак, решение какой-либо задачи выполняется по некоторому алгоритму — упорядоченной конечной совокупности достаточно определенных правил, применение которых приводит к некоторому результату. Алгоритм должен быть конструктивным в том смысле, что он может быть выполнен некоторым вычислительным устройством (компьютером). Результатом работы алгоритма может быть одно из двух:

- ❒ данная задача решена за конечное время — собственно полезный результат;
- ❒ алгоритм не останавливается никогда (зависает) либо выдает в качестве результата сообщение о невозможности своего завершения (сообщение об ошибке).

Конструктивность алгоритма в математике рассматривается безотносительно к конкретному типу вычислительного устройства путем его выражения через понятие *вычислимой функции*. Что бы ни делал алгоритм, он, в конце концов, вычисляет значение некоторой функции. При этом говорят проще: алгоритм вычисляет функцию. Не всякая функция вычислима, но если она вычислима, то существует алгоритм, вычисляющий ее за конечное время (конечное количество шагов), и, наоборот, для невычислимой функции не существует алгоритма, который мог бы вычислить ее за конечное время. С понятием вычислимости функции тесно связаны понятия перечислимости и разрешимости множеств, которые мы сейчас рассмотрим.

Зануда. Если я Вас правильно понял, Профессор, понятие алгоритма остается интуитивным, т. е. формально неопределенным.

Профессор. Совершенно верно, мы отправляемся от интуитивного понятия алгоритма как некоторого точного предписания, руководствуясь которым с каждым исходным данным из области возможных данных можно за конечное количество шагов сопоставить некоторый результат применения этого алгоритма.

Зануда. Таким образом, для любого алгоритма следует указать область возможных данных. Значит, существует область данных, к которым рассматриваемый алгоритм неприменим.


П р о ф е с с о р. Область возможных исходных данных — множество исходных данных, к которым алгоритм применим. Например, алгоритм деления чисел "столбиком" применим к числам, но не применим к буквам или звукам.


З а н у д а. А может ли быть так, что алгоритм воспринимает возможные исходные данные, но все равно зависает или выдает сообщение об ошибке?


П р о ф е с с о р. Совершенно не требуется, чтобы алгоритм давал результат, будучи примененным к любым данным из области возможных данных. Так, алгоритм деления числа x на число y применим к числам, но не дает результата, если $y = 0$ (в этом случае компьютерная программа обычно выводит на экран сообщение об ошибке). Множество тех возможных исходных данных, которые перерабатываются алгоритмом в некоторый результат, отличный от сообщения об ошибке или зависания, называется *областью применимости алгоритма*. Очевидно, что область применимости алгоритма содержится в области его возможных данных.

Итак, я начинаю подходить к определению перечислимых множеств. Пусть даны множества X и Y, а также частично определенная функция f, которая однозначно преобразует элементы множества X в элементы множества Y. Такую функцию в общем виде обычно обозначают как $f : X \to Y$. Частичная определенность функции f означает, что, возможно, не для каждого элемента $x \in X$ найдется значение $f(x) \in Y$ этой функции. Однозначность преобразования означает, что для любого элемента x, для которого функция f определена, значение $f(x)$ является некоторым единственным элементом из множества Y. Подмножество $X_0 \subseteq X$, такое что для всех его элементов функция f определена, называется *областью определения функции* f. Понятно, что частично определенная функция $f : X \to Y$ полностью определена на множестве $X_0 \subseteq X$.

Функция f называется *вычислимой*, если существует алгоритм с областью применимости X_0, такой что для всякого $x \in X_0$ значение $f(x)$ совпадает с результатом применения этого алгоритма к x.


П р о с т а к. Как я понял, вычислимая функция — это функция, для которой имеется алгоритм вычисления ее значений по исходным данным, к которым данный алгоритм применим. Но если функция существует, значит, существует и алгоритм ее вычисления. Так что я не вижу никакой разницы между алгоритмом и вычислимой функцией.


П р о ф е с с о р. Определение функции можно дать различными способами, в том числе и словесным описанием. Однако при этом может оказаться, что нельзя дать конечную упорядоченную совокупность правил, выполнение которых приведет к вычислению значения функции. Пусть, например, дано такое определение: функция $f : \Re \to \mathrm{N}$ однозначно отображает

действительные числа $x \in \Re$ в множество натуральных чисел N; другими словами, функция f нумерует действительные числа. Однако для вычисления этой функции алгоритма не существует. Кстати, поскольку алгоритма нет, говорят, что и функция $f : \Re \to \mathrm{N}$ *не существует*. В общем случае для любой функции требуется доказать, существует ли для ее вычисления алгоритм. При этом не обязательно приводить алгоритм в явном виде, хотя для практики это важно. Так, например, аксиома выбора декларирует, что существует функция выбора элементов из множеств некоторой совокупности. Противники этой аксиомы, конструктивисты, приводят примеры случаев, когда не ясно, как именно осуществить выбор. Они подвергают сомнению существование функции выбора из-за сомнения в возможности построения алгоритма вычисления этой функции.

Итак, я продолжу. Пусть дана частично определенная функция $f : X \to Y$, причем подмножество $X_0 \subseteq X$ — область применимости алгоритма, вычисляющего эту функцию. Если $x \in X_0$, то алгоритм возвращает, как говорят, значение $f(x)$ данной функции. Если $x \in X \setminus X_0$, то относительно результатов работы алгоритма не выдвигается никаких требований, за исключением того, что он не должен давать заведомо неверный результат; поэтому и принимается, что в этом случае алгоритм либо "зависает", либо выдает специальное сообщение об ошибке. Впрочем, сообщение об ошибке может быть и пустым: алгоритм останавливается, ничего не выдавая.

Обозначим через N множество всех натуральных чисел, включая 0 (0, 1, 2, 3, …). Множество X называется *перечислимым*, если существует всюду определенная функция $f : N \to X$. Говорят, что функция f перечисляет множество X: каждому натуральному числу она сопоставляет некоторый элемент множества X.

Пустое множество Ø будем считать вырожденным случаем перечислимого множества. Оно перечисляется алгоритмом, который при любом $n \in N$ не дает никакого результата.

Так как перечислимое множество есть множество значений вычислимой функции f, определенной на множестве натуральных чисел с нулем, то, вычисляя последовательно значения $f(0), f(1), f(2),...$ и т. д., будем порождать элементы множества X (возможно, с повторением). Таким образом, каждое перечислимое множество является эффективно порождаемым в том смысле, что для него существует вычислительный процесс, последовательно генерирующий элементы этого множества. Существуют различные виды эффективных порождающих процессов, однако все они сводятся к процессу перечисления.

Рассмотрим в качестве примера множество A^* всех слов, составленных из символов конечного алфавита A. Это множество бесконечно. Перенумеруем символы алфавита A натуральными числами: $a_1, a_2, ..., a_n$. Числу 0 сопоставим пустое слово, которое обозначим через e, числу 1 сопоставим слово, состоящее из единственного символа a_1, числу 2 сопоставим слово, состоящее из единственного символа a_2 и так далее, числу n сопоставим слово a_n. Далее, числу $n+1$ сопоставим слово a_1a_1, числу $n+2$ — слово a_1a_2 и так далее. Из двух слов разной длины предшествующим считается то, которое короче. Если же слова имеют одинаковою длину, то, сравнивая слова посимвольно слева направо, необходимо найти первую слева пару различающихся символов. Тогда предшествующим будет то слово, у которого символ из этой пары имеет меньший индекс (символы пронумерованы). Таким образом, мы описали алгоритм перечисления слов из множества A^*. Поэтому множество A^* перечислимо. Описанный алгоритм устанавливает взаимно однозначное соответствие f между множествами N и A^*. Поэтому существует и обратный алгоритм определения по слову $x \in A^*$ его номера $n \in N$. В самом деле, чтобы определить номер слова, необходимо просто последовательно вычислять $f(0), f(1), f(2),...$, пока для некоторого n не окажется $f(n) = x$. Тогда найденное n и есть результат работы алгоритма. Таким образом, множество A^* счетно.

Теперь покажем, что конечное множество также перечислимо. Пусть дано конечное непустое множество X. Выберем из него какой-нибудь элемент и обозначим его через x_0. С числом 0 сопоставим x_0. Выберем из множества $X \setminus \{x_0\}$ новый элемент, обозначим его через x_1 и сопоставим его числу 1. Будем поступать подобным образом до тех пор, пока не окажется, что $X \setminus \{x_0, x_1,..., x_n\} = Ш$. С остальными числами $n+1, n+2,...$ сопоставим один и тот же элемент x_n. Таким образом, посредством только что описанного алгоритма определена вычислимая функция из N на X. Следовательно, X — перечислимое множество.

Относительно перечислимых множеств можно доказать следующие факты.

- Объединение, пересечение и декартово произведение конечного числа перечислимых множеств перечислимо.

> **Примечание**
>
> Декартово произведение множеств мы рассмотрим в *разд. 4.2*.

- Множество всех перечислимых множеств счетно. Поскольку множество всех подмножеств множества натуральных чисел не счетно, то существуют неперечислимые множества.

Пусть $X_0 \subseteq X$ — перечислимое подмножество множества X. Пусть также $x \in X$. Требуется установить, принадлежит ли x подмножеству X_0. Для этого можно поступить следующим образом. Поскольку X_0 перечислимо, то существует вычислимая функция f, перечисляющая X_0. Будем последовательно вычислять $f(0), f(1),...$ до тех пор, пока не выполнится равенство $f(n) = x$. Если $x \in X_0$, то данная процедура приведет к успеху. Однако если $x \notin X_0$, то алгоритм, вычисляющий функцию f, либо остановится на каком-то шаге, не выдав результата, и тем самым мы не узнаем, что $x \notin X_0$, либо будет работать бесконечно долго, и мы никогда не определим, то ли подходящее число n еще не достигнуто, то ли $x \notin X_0$. Таким образом, установить принадлежность элемента перечислимому множеству удается не всегда. Это связано с тем, что дополнение $X \setminus X_0$ перечислимого множества может быть неперечислимым.

Подмножество X_0 множества X называется *разрешимым* относительно X, если как само X_0, так и его дополнение $X \setminus X_0$ перечислимы.

Очевидно, что понятие разрешимого множества является частным случаем понятия перечислимого множества.

Для разрешимого множества X_0 существует разрешающий алгоритм, который распознает принадлежность элементов этому множеству или его дополнению $X \setminus X_0$. У функции f, соответствующей разрешающему алгоритму, область возможных значений состоит из двух элементов. Если $x \in X_0$, то положим $f(x)=1$, в противном случае $f(x)=0$.

Построим алгоритм вычисления разрешающей функции f. Пусть f_1, f_2 — вычислимые функции, перечисляющие множества X_0 и $X \setminus X_0$ соответственно. Последовательно вычисляем значения этих функций: $< f_1(0), f_2(0) >$, $< f_1(1), f_2(1) >, < f_1(2), f_2(2) >, \ldots$ Если для некоторого числа n окажется, что $f_1(n)=x$, то положим $f(x)=1$. Если для некоторого n встретится $f_2(n)=x$, то положим, что $f(x)=0$. Ясно, что f — вычислимая функция и для любого x значение $f(x)$ вычисляется за конечное число шагов.

Поскольку множество всех перечислимых множеств счетно, то и множество вычислимых функций счетно. Более того, оказывается, что множество вычислимых функций, отображающих множество X в множество Y, перечислимо. Поэтому существует вычислимая функция g, которая каждому натуральному числу i сопоставляет некоторую вычислимую функцию $f_i : X \to Y$, причем так, что если i пробегает все множество натуральных чисел, то $g(i)$ перечисляет все вычислимые функции из X в Y. Отсюда следует, что существует универсальная функция $\Phi(i,x)=(g(i))(x)$, которая как бы содержит в себе все вычислимые функции из X в Y.

На этом мы завершаем рассмотрение элементов теории множеств, оставляя без внимания многие ее важные и интересные разделы, в том числе один из главных — трансфинитные числа. Трансфинитные числа характеризуют "размеры" бесконечных множеств. Счетные и несчетные бесконечные множества, например, имеют различные мощности, которым сопоставлены различные трансфинитные числа. Последние могут быть упорядочены подобно обычным числам и для них существует своя арифметика. Однако данная тема выходит за рамки этой книги.

ЧАСТЬ II

Шутки в сторону

Профессор. Теперь мы переходим к анализу понятия отношения между объектами, которое играет очень важную роль не только в науках, но и в обычной жизни. Мы привыкли рассматривать объекты различной природы в их отношениях друг к другу, поскольку в отношениях раскрывается суть объектов — их внутреннее устройство, внешние взаимосвязи и поведение. Все познается в сравнении. Наша задача состоит в том, чтобы первоначальные, довольно простые, но расплывчатые категории перевести на язык математики и далее исследовать математическими средствами уже уточненные понятия.

Начиная с данной темы, я перехожу от диалога к монологу, чтобы немного увеличить темп изложения. Если хотите, я начинаю читать лекции. В *главе 4* мы рассмотрим общие сведения об отношениях: что это такое, как они определяются и в каких формах представляются. Нам потребуется освоить некоторую несложную символику, чтобы понять материал последующих глав. Главное, что предстоит усвоить из *главы 4*, заключается в том, что отношение это множество, элементы которого мы теперь рассматриваем как имеющие структуру определенного сорта. Это — последовательности объектов, находящихся в рассматриваемом отношении. Последовательность объектов тоже является объектом с некоторой структурой: элементы последовательности упорядочены, они следуют друг за другом. Стоит только изменить этот порядок, и мы получаем другую последовательность. Таким образом, теория отношений является по существу теорией множеств, состоящих из последовательностей объектов.

Глава 4

Отношения

> Истинные отношения между реальными предметами представляют собой единственную реальность, которую мы можем постигнуть… Раз отношения нам известны, то уже не существенно, какое образное выражение мы считаем удобным применить.
>
> *А. Пуанкаре*

4.1. Что такое отношение?

В обычной практике различные объекты сравниваются между собой с некоторой точки зрения или рассматриваются в некоторой связи. Вот несколько примеров отношений:

1. Людей можно рассматривать в отношении старшинства по возрасту или занимаемой должности: "Антон старше Петра", "Иван — начальник Сергея".
2. Товары могут сравниваться в отношении предпочтения с точки зрения некоторого потребителя: "Автомобиль лучше булочки с маком".
3. Число 2 меньше числа 5.
4. Некоторые точки плоскости могут быть связаны тем условием, что лежат на одной и той же линии; числа x и y связаны между собой соотношением $y = \sin x$.
5. Слова "меч" и "мяч" похожи, поскольку различаются только одной буквой; эти же слова имеют одинаковое количество букв.

Некоторые три числа могут образовывать дату в формате *день.месяц.год*: тройка чисел 06.11.1952 образует дату 6 ноября 1952 года, а тройка чисел 31.06.2007 не соответствует никакой дате, т. к. в июне только 30 дней.

Таблица со столбцами "Наименование товара", "Поставщик", "Количество", "Цена" содержит данные, представляющие отношение, которое можно назвать, например, "Товары на складе".

Мы видим, что отношения между объектами могут быть весьма разнообразными. В то же время, между различными по смыслу отношениями можно заметить нечто общее. Так, между вариантами отношения старшинства из примера 1 и отношениями из примеров 2 и 3 общим является то, что все объекты, находящиеся в этих отношениях, можно расположить друг относительно друга в некотором порядке. Например, можно образовать последовательность, в которой старший (более предпочтительный) объект будет находиться правее младшего (менее предпочтительного). Поэтому подобные отношения называют *отношениями порядка*, или просто *порядками*.

В примере 4 объекты связаны между собой зависимостью, которая выражена некоторой математической формулой.

В примере 5 приведен частный случай отношения сходства (похожести) между словами. Очевидно, что оно отличается от отношения порядка хотя бы тем, что из сходства двух объектов не вытекает какое-либо их упорядочение. Так, если объект x похож на объект y, то имеет место и обратное: объект y похож на объект x. В силу этой симметричности отношения сходства мы и не можем упорядочить объекты. Допустим, мы определили, что два слова похожи, если они отличаются не более чем одной буквой с учетом ее места в слове. Тогда слова "меч" и "мяч" похожи. Разумеется, мы могли бы определить сходство между словами и иначе. Например, два слова похожи, если они различаются не более, чем n буквами, без учета их места в слове. Пусть $n = 1$, т. е. два слова похожи, если они отличаются не более чем одной буквой. Тогда слово "меч" похоже на слово "мяч", а слово "мяч" похоже на слово "мяу", но слово "мяч" не похоже на слово "мяу". Как нередко бывает, в цепочке родителей и детей сходство постепенно утрачивается, даже исчезает совсем, но потом вдруг возникает снова. Для отношения сходства, независимо от конкретного способа определения, характерно, что если объект x похож на y, а y похож на z, то не обязательно x похож на z. В отличие от отношения сходства, для отношения эквивалентности объектов, как бы оно ни было определено, обязательно должно выполняться условие: если объект x эквивалентен y, а y эквивалентен z, то x эквивалентен z. Отношение эквивалентности является обобщением отношения равенства. Можно сказать, что эквивалентность — это равенство с определенной точки зрения или неразличимость при заданном наборе отличительных признаков. А для равенства всегда выполняется условие: если $x = y$, а $y = z$, то $x = z$. Задав, так или иначе, отношение эквивалентности, мы можем разбить множество всех объектов, находящихся в данном отношении, на непересекающиеся классы так, что эквивалентные друг другу объекты окажутся в одном и том же классе, а неэквивалентные — в разных классах. Например, между словами можно определить отношение "состоять из одинакового количества букв". Тогда множество всех слов разобьется на непересекающиеся классы слов одинаковой длины: класс однобуквенных слов, класс двухбуквенных слов, класс трехбуквенных слов, …, класс n-буквенных слов (если, разумеется, имеются слова из n букв).

В примерах 1—5 рассматривались отношения между двумя объектами. Такие отношения называются *бинарными*, или *двухместными*. Однако отношения могут быть и многоместными. В примерах 6 и 7 приведены трехместное и четырехместное отношения. Возможны также одноместные отношения, например, "помидор — красный", "2 — четное число", "числа x, удовлетворяющие уравнению $x^2 - 1 = 0$". Одноместные отношения еще называют *унарными*, или *свойствами* объектов, а трехместные отношения — *тернарными*.

Обратите внимание: в предложениях, выражающих отношения, фигурируют *имя отношения* и *имена объектов*. Например, в предложении "Антон старше Петра" слово "старше" является именем отношения, а Антон и Петр — именами объектов, находящихся в данном отношении. При рассмотрении подобных предложений сама собой напрашивается гипотеза о том, что имена объектов можно заменить другими. Что при этом произойдет? Возможны следующие три варианта:

- отношение будет выполняться и для других объектов;
- отношение перестанет выполняться;
- отношение потеряет смысл.

Например, если в предложение "2 меньше 5" вместо числа 2 подставить число 3, то отношение все равно будет выполняться; если же вместо 2 подставить 10, то отношение перестанет выполняться. Если, заменив объекты, мы получим предложение "вода меньше звезды", то отношение "меньше" потеряет смысл, хотя в художественной литературе, особенно в поэзии, часто используются предложения с переносным смыслом. Понимая предложение "вода меньше звезды" в каком-то переносном смысле, мы должны отказаться от того, что слово "меньше" выражает отношение между количественными характеристиками объектов.

Чтобы избежать бессмысленных предложений об отношениях, достаточно зафиксировать множества объектов, для которых данное отношение имеет смысл. Например, в предложении "2 меньше 5" неявно предполагается, что 2 и 5 — числа. А может быть это не числа, а индексы (цифровые ярлыки), обозначающие, например, слона и мышь? В таком случае отношение "меньше" не будет выполняться. Для большей определенности мы могли бы использовать такую формулировку: "число 2 меньше числа 5". Здесь слово "число" указывает на множество всех чисел, которому принадлежит объект, рассматриваемый в отношении "меньше". Аналогично, переформулировав предложение "вода меньше звезды" в виде "слово "вода" меньше слова "звезда" по количеству букв", мы получаем предложение с прямым (а не переносным) смыслом.

Итак, мы теперь знаем, как избавиться от бессмысленности отношений: достаточно зафиксировать множества (категории) объектов, для которых рассматриваемое отношение имеет смысл. При этом, однако, само отношение может выполняться для одних объектов и не выполняться для других. что не является какой-либо аномалией. Остается только определить само отношение. С одной стороны, отношения могут по-разному называться, но обладать одинаковыми свойствами, а с другой — одноименные отношения могут существенно различаться.

4.2. Как задать отношение?

Чтобы задать отношение, следует указать, между какими объектами оно выполняется. Это можно сделать путем явного перечисления таких объектов или указанием правила (способа), с помощью которого можно определить, находятся данные объекты в задаваемом отношении или нет. Первый способ, называемый еще *экстенсиональным* (*явным*), подходит в случае конечных и небольших по объему множеств объектов, а второй, называемый *интенсиональным* (*неявным*), более универсален и может применяться даже в случае бесконечных множеств объектов.

Рассмотрим, как задать бинарное (двухместное) отношение между двумя объектами. Прежде всего, необходимо указать два множества (обозначим их как A и B), между элементами которых задается отношение. Затем необходимо определить множество всех пар (x, y) элементов x и y таких, что $x \in A$, $y \in B$ и x находится в рассматриваемом отношении с y. Множества A и B могут быть как различными, так и равными.

Таким образом, чтобы определить бинарное отношение, следует указать:

1. Его область задания — множество всех возможных пар объектов.
2. Подмножество тех и только тех пар объектов, которые находятся в рассматриваемом отношении.

Область задания бинарного отношения формируется из двух множеств A и B с помощью операции, называемой *декартовым произведением* этих множеств и обозначаемой как $\times$. В результате применения этой операции получается множество $A \times B$ (также называемое *декартовым произведением* A *и* B). Это множество состоит из упорядоченных пар, в которых первый элемент принадлежит множеству A, а второй — множеству B.

Определение 4.1

Декартовым произведением множеств A и B называется множество $A \times B$, состоящее из тех и только тех пар (a, b), для которых $a \in A$, $b \in B$. Символически это можно записать так: $A \times B = \{(a, b) : a \in A \,\&\, b \in B\}$.

Пусть, например, $A = \{a_1, a_2, a_3\}$ и $B = \{b_1, b_2\}$,

тогда $A \times B = \{(a_1, b_1), (a_1, b_2), (a_2, b_1), (a_2, b_2), (a_3, b_1), (a_3, b_2)\}$.

Внимание

Элементы декартова произведения двух множеств — упорядоченные пары элементов. Поэтому, вообще говоря, $A \times B \neq B \times A$. Так, в приведенном выше примере $(a_1, b_1) \in A \times B$, но $(b_1, a_1) \notin A \times B$.

Декартово произведение $n \geq 2$ множеств $A_1, A_2, ..., A_n$ определяется следующим образом:

$$A_1 \times A_2 \times ... \times A_n = \{(a_1, a_2, ..., a_n) : a_1 \in A_1 \,\&\, a_2 \in A_2 \,\&\, ... \,\&\, a_n \in A_n\}.$$

Если все множества одинаковы ($A_1 = A_2 = ... = A_n = A$), то их декартово произведение обозначают как A^n и называют *n*-ой *степенью* множества A. Например, $A^2 = A \times A$.

Примечание

Операция декартова произведения множеств названа так в честь французского мыслителя и математика Р. Декарта. Не он придумал эту операцию, но он первым стал сопоставлять точкам числа — координаты и, тем самым, стал основоположником аналитической геометрии. Прямоугольную систему координат называют *декартовой*.

На рис 4.1 в виде прямоугольника показано декартово произведение множеств X и Y, являющихся подмножествами множества действительных чисел. Понятно, что в данном случае декартово произведение это не прямоугольник как геометрическая фигура, а множество всех его точек. Множество X — это отрезок прямой, содержащий точки, которым, в свою очередь, взаимно однозначно сопоставлены числа — координаты этих точек на горизонтальной прямой, являющейся одной из осей декартовой (прямоугольной) системы координат. То же самое можно сказать и о множестве Y. Элементам множества Y (точкам) также взаимно однозначно сопоставлены числа — координаты этих точек на вертикальной прямой. Горизонтальная и вертикальная прямые перпендикулярны друг другу (пересекаются под прямым углом, равным, извините за банальность, 90º) и образуют конфигурацию, называемую *прямоугольной*, или *декартовой системой координат*. Итак, декартово произведение множеств X и Y — это множество, обозначаемое как $X \times Y$, которое состоит из всех точек с координатами (x, y), такими что $x \in X$ и $y \in Y$. Таким образом, декартово произведение $X \times Y$ есть множество всех точек плоскости, находящихся во взаимно однозначном соответствии с множеством пар всех (x, y) координат этих точек. Следовательно, декартово произведение $X \times Y$ и есть множество таких пар, а то, что изображено на рис. 4.1, является лишь геометрическим (визуальным) его образом.

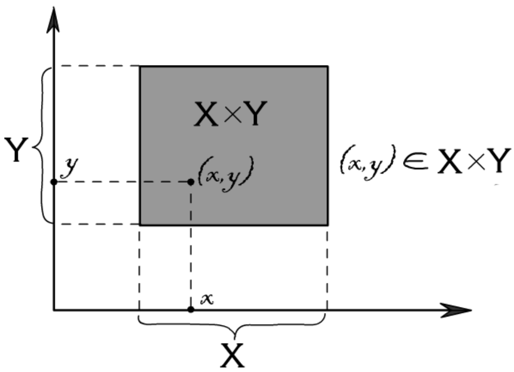


Рис. 4.1. Декартово произведение двух подмножеств множества действительных чисел

Элементы декартова произведения n множеств, как мы видели, представляются упорядоченными последовательностями элементов из этих множеств. Эти последовательности в математике обычно называются *кортежами* (франц. cortege — торжественная процессия). Важно понимать, что кортеж это не множество, а упорядоченная последовательность элементов. Стоит только изменить порядок элементов в кортеже, и мы получим, вообще говоря, другой кортеж. Количество элементов в кортеже называется его *длиной*. Обращаясь к геометрической интерпретации, элементы декартова произведения n множеств можно представить как точки n-мерного пространства. Множества, элементами которых являются кортежи длины n, иногда называют n-мерными множествами. Обычные множества можно рассматривать как одномерные.

Примечание

Если хотя бы одно из множеств, A или B, пусто, то их декартово произведение $A \times B$ также пусто. Это можно доказать следующим образом. Допустим, что $A = \varnothing$ (пустое множество). Чтобы сформировать множество $A \times B$, следует построить все его элементы. Для этого необходимо взять элемент множества A и приписать к нему справа какой-нибудь элемент множества B, но в множестве A нет элементов, и, следовательно, мы не можем построить ни одного элемента множества $A \times B$. Таким образом, $A \times B = \varnothing$. Аналогично рассматривается случай, когда $B = \varnothing$.

Итак, мы выяснили, как определяется область задания отношения. Для n-местного отношения это — декартово произведение n множеств объектов, которые находятся или не находятся в определяемом отношении. Для определения самого отношения этого, очевидно, недостаточно. Необходимо еще определить подмножество тех кортежей из области задания, элементы которых находятся в интересующем нас отношении.

Определение 4.2

Пусть $A_1 \times A_2 \times \ldots \times A_n$ — область задания отношения, а $R \subseteq A_1 \times A_2 \times \ldots \times A_n$ — некоторое подмножество области задания. Тогда пару $\langle A_1 \times A_2 \times \ldots \times A_n, R \rangle$ называют *отношением*, а множество R — *графиком* этого отношения. При этом число n называют *арностью*, или *местностью* отношения, говоря об n-арном или, что то же самое, n-местном отношении. Для $n \leq 3$ употребляются специальные названия:

- унарное ($n = 1$);
- бинарное ($n = 2$);
- тернарное ($n = 3$).
- Если $R = A_1 \times A_2 \times \ldots \times A_n$ (график отношения полностью совпадает с областью задания), то отношение называют *полным*.

Если $R = \varnothing$, то отношение называют *пустым*.

Когда из контекста ясно, о какой области задания идет речь, нередко отношением называют его график. Мы тоже будем так поступать. Отношения (графики отношений) суть множества и, следовательно, к ним применимы все теоретико-множественные операции, такие как объединение, пересечение и др. Но в силу особенности отношений как многомерных множеств к ним применимы еще и специальные операции, которыми мы займемся позже.

Как уже отмечалось в *главе 3*, множества можно задавать с помощью предиката — двузначной функции, определяемой как некоторое условие (правило). Этот предикат принимает значение ИСТИНА (И), если элемент принадлежит данному множеству; в противном случае он принимает значение ЛОЖЬ (Л). Пусть $R \subseteq A_1 \times A_2 \times \ldots \times A_n$ — некоторое отношение. Обозначим предикат, определяющий это множество R, тем же символом, что и само множество: $R(x_1, x_2, \ldots, x_n)$. Тогда отношение $R \subseteq A_1 \times A_2 \times \ldots \times A_n$ можно определить и в следующей интенсиональной форме:

$$R = \{(x_1, x_2, \ldots, x_n) : x_1 \in A_1 \,\&\, x_2 \in A_2 \,\&\ldots\&\, x_n \in A_n \,\&\, R(x_1, x_2, \ldots, x_n)\}.$$

Когда мы говорим, что элементы $a_1, a_2, \ldots, a_n$ из множеств соответственно $A_1, A_2, \ldots, A_n$ находятся в некотором отношении R, то мы имеем в виду, что $(a_1, a_2, \ldots, a_n) \in R$, т. е. кортеж $(a_1, a_2, \ldots, a_n)$ принадлежит отношению R.

В случае бинарных отношений, если $(x, y) \in R$, нередко употребляют и такое обозначение: xRy. Например, обычно отношение "меньше или равно" записывают как $x \leq y$, а не в виде $(x, y) \in \leq$.

Рассмотрим в качестве примера бинарное (двухместное) отношение "родители-дети" между элементами множества родителей Р = {*мама*$_1$, *папа*$_1$, *мама*$_2$, *папа*$_2$} и множества детей Д = {*Саша*, *Маша*, *Антон*, *Иван*}. Область задания этого отношения есть множество $Р \times Д$, состоящее из 4 × 4 = 16 всевозможных пар родитель-ребенок. Разумеется, совсем не обязательно, чтобы элементы каждой такой пары находились в отношении "родители-дети". График этого отношения (обозначим его через R) состоит из тех и только тех пар родитель-ребенок, в которых первый элемент является родителем второго элемента. Допустим, что *Саша*, *Маша* и *Антон* имеют общих родителей *мама*$_1$ и *папа*$_1$, а *Иван* имеет других родителей — *мама*$_2$ и *папа*$_2$. Тогда график отношения "родители-дети" есть следующее множество:

R = {(*мама*$_1$, *Саша*), (*папа*$_1$, *Саша*), (*мама*$_1$, *Маша*), (*папа*$_1$, *Маша*),

(*мама*$_1$, *Антон*), (*папа*$_1$, *Антон*), (*мама*$_2$, *Иван*), (*папа*$_2$, *Иван*)}.

Данное отношение удобно представить графически, проводя стрелки от родителей к их детям.

Теперь, исходя из рассмотренного отношения "родители-дети", определим отношение "быть братом" между элементами множества детей $Д$: x является братом y, если они имеют двух общих родителей и x — мужчина (очевидно, что женщина не может быть кому-либо братом). Из общепринятого смысла данного отношения следует также, что никакой x не может быть братом самому себе. Таким образом, мы сформулировали условия, ограничивающие выбор пар детей из множества $Д \times Д$. Этим условиям удовлетворяет, например, следующее отношение:

$Б = \{(Саша, Маша), (Саша, Антон), (Антон, Саша), (Антон, Маша)\}$. Отношения "родители-дети" и "быть братом" показаны на рис. 4.2.

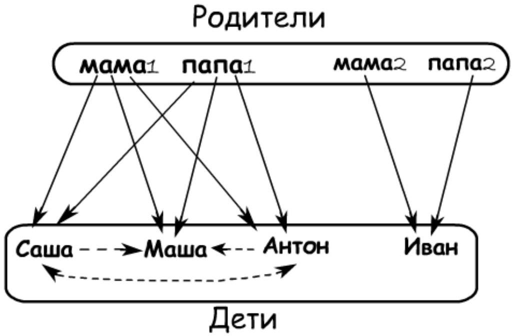


Рис. 4.2. Отношения "родители-дети" (сплошные стрелки) и "быть братом" (штриховые стрелки), определяемое как "иметь двух общих родителей"

Рассмотрим отношение "быть братом" подробнее.

- Если допустимо, что Саша является братом Маши, то обратное утверждение — нет. Поэтому пара (*Маша*, *Саша*) отсутствует в множестве, представляющем отношение "быть братом". То же самое справедливо и для Антона и Маши. Таким образом, бинарное отношение может быть несимметричным.
- Очевидно, что никто не является братом самому себе, т. е. ни одна пара вида (x, x) не принадлежит множеству, представляющему отношение. Таким образом, данное отношение, как говорят, не рефлексивно.
- Иван не является братом никому. Таким образом, не обязательно, чтобы для каждого элемента нашелся хотя бы один элемент, с которым он находится в отношении.
- Если Саша — брат Антона, а Антон — брат Маши, то отсюда необходимо следует, что Саша — брат Маши. Таким образом, отношение может иметь некоторые внутренние зависимости. В данном примере эта зависимость обусловлена тем, что по определению "быть братом" означает "иметь двух общих родителей". Если же считать братьями также и тех, у кого только один общий родитель, то эта зависимость может и не выполняться.

Вообще говоря, дети могут иметь только одного общего родителя, являясь в этом случае сводными братьями или сестрами. Рассмотрим пример такого отношения "родители-дети":

$$R = \{(\text{мама}_1, \text{Саша}), (\text{папа}_1, \text{Саша}), (\text{мама}_3, \text{Маша}), (\text{папа}_1, \text{Маша}),$$
$$(\text{мама}_1, \text{Антон}), (\text{папа}_3, \text{Антон}), (\text{мама}_2, \text{Иван}), (\text{папа}_2, \text{Иван})\}.$$

Определим теперь на основе данного отношения R отношение "быть братом" $Б$: x является братом y, если они имеют хотя бы одного общего родителя, x — мужчина и для всех x $(x,x) \notin Б$. Данному определению удовлетворяет следующее множество пар детей:

$$Б = \{(\text{Саша}, \text{Маша}), (\text{Саша}, \text{Антон}), (\text{Антон}, \text{Саша})\}.$$

Рассмотренные отношения представлены на рис. 4.3. В данном случае Саша является братом Маши и Антона, но последний не является братом Маши. Это обстоятельство обусловлено тем, что Саша, Маша и Антон имеют только одного общего родителя, причем попарно, а не втроем. Таким образом, Саша — сводный брат Маши по папе$_1$ и сводный брат Антона по маме$_1$, но у Антона и Маши нет общих родителей и поэтому он не является ее братом. Иван, как и в предыдущем примере, не имеет общих родителей ни с кем и, следовательно, не имеет братьев (не принадлежит отношению "быть братом").

Рассмотрим еще один пример. Пусть X — множество всех действительных чисел, а Y — множество чисел в интервале от 0 до 1, включая границы. Тогда уравнение $y = \sin x$ задает отношение между числами $x \in X, y \in Y$, которое графически можно представить в виде синусоиды (волнообразной линии с амплитудой колебаний, равной 1). Обозначим график этого отношения через S, тогда $S \subseteq X \times Y$ и $S = \{(x, y) : x \in X \;\&\; y \in Y \;\&\; (y = \sin x)\}$.

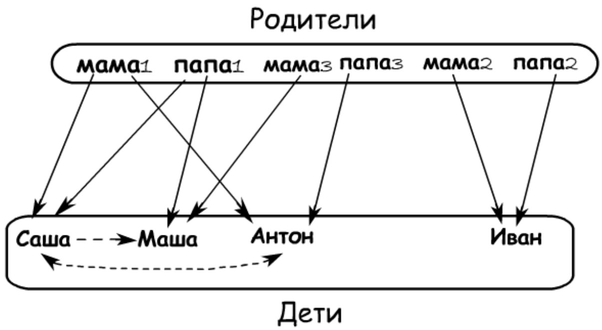


Рис. 4.3. Отношения "родители-дети" (сплошные стрелки) и "быть братом" (штриховые стрелки), определяемое как "иметь хотя бы одного общего родителя" и, следовательно, допускающее сводных братьев

Бинарные и тернарные отношения, если их области задания являются конечными и относительно небольшими множествами, удобно представлять в форме *графа*. В таких графах объекты изображаются точками, помеченными именами объектов из двух множеств или просто этими именами, без точек; от одного объекта к другому проводят стрелку, если первый из них находится в данном отношении со вторым объектом. Стрелка может быть двунаправленной, если отношение между данными двумя объектами симметрично. В этом случае стрелки можно вообще не рисовать. На рис. 4.1 и 4.2 показаны примеры представления бинарных отношений в виде графов.

Иногда бинарное отношение удобно представлять в виде матрицы — прямоугольной таблицы, устроенной следующим образом. Строкам этой таблицы соответствуют элементы одного множества, а столбцам — элементы другого. В клетке на пересечении строки и столбца записывается 1, если соответствующие элементы находятся в рассматриваемом отношении, и записывается 0, если соответствующие элементы не находятся в этом отношении. Иногда вместо 1 указывают какой-нибудь символ, например "+", а вместо 0 ничего не пишут.

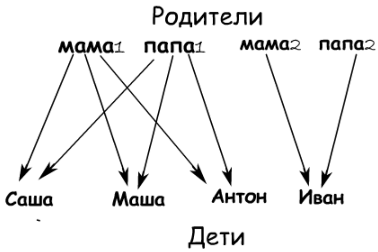
Бинарное отношение можно представлять и в виде таблицы, два столбца которой соответствуют множествам элементов, а в строках записываются имена тех элементов, которые находятся в данном отношении.

На рис. 4.4 показано отношение "родители-дети", изображенное в виде графа, матрицы и таблицы.

В случае тернарных отношений стрелки на графе помечаются элементами из третьего множества. Например, пусть A — множество городов, а S — множество действительных чисел. Тогда отношение "дальность полета" $R \subseteq A \times A \times S$ можно интерпретировать как указатель дальностей полета между городами. Граф такого отношения имеет вид точек с названиями городов, между которыми проведены линии без стрелок или со стрелками на обоих концах, а около каждой линии указано расстояние между городами, которые она соединяет. При матричном представлении такого отношения внутри клеток указываются расстояния между городами. Нетрудно представить себе, как будет выглядеть это отношение в виде таблицы. На рис. 4.5 показано тернарное отношение "дальность полета между городами". В данном примере мы условились, что одноименные города (точнее, кортежи вида (a, a, s)) не принадлежат отношению. Если бы, напротив, мы решили рассматривать в отношении и такие кортежи (точнее, кортежи $(a, a, 0)$), то в графе отношения в каждой вершине следовало бы нарисовать петлю — линию, выходящую из вершины и входящую в эту же вершину; около петли следовало бы написать число 0. В матрице отношения мы указываем нули на диагонали, чтобы показать, что соответствующие одноименные города не принадлежат отношению, хотя их можно интерпретировать и как расстояния, равные нулю.

Наиболее универсальным способом представления конечных отношений является табличный, при котором каждому кортежу отношения взаимно однозначно соответствует строка таблицы. Именно этот способ положен в основу построения реляционных баз данных, в которых данные организованы в виде системы взаимосвязанных отношений (англ. relation — отношение, связь).

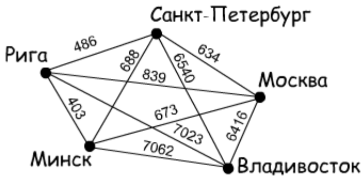
В следующей главе мы рассмотрим важнейший класс отношений между элементами двух множеств — бинарные отношения. В зависимости от ограничений, накладываемых на связи между двумя элементами, эти отношения в целом приобретают интересные свойства, знание которых вооружает нас для исследования разнообразных по своей природе взаимосвязей объектов реального мира. Объекты и связи между ними могут иметь различное происхождение и строение, но между ними есть нечто общее, что и является предметом теории бинарных отношений, т. е. некоторого раздела математики.



$$
\text{Родители}\quad
\begin{array}{c}
 \\ \text{мама}_1 \\ \text{папа}_1 \\ \text{мама}_2 \\ \text{папа}_2
\end{array}
\begin{array}{c}
\begin{array}{cccc}\text{Саша} & \text{Маша} & \text{Антон} & \text{Иван}\end{array} \\
\left\|\begin{array}{cccc}
1 & 1 & 1 & 0 \\
1 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1
\end{array}\right\|
\end{array}
$$

Родители	Дети
мама1	Саша
мама1	Маша
мама1	Антон
папа1	Саша
папа1	Маша
папа1	Антон
мама2	Иван
папа2	Иван

Рис. 4.4. Бинарное отношение "родители-дети", представленное в виде графа, матрицы и таблицы



	Владивосток	Минск	Москва	Рига	Санкт-Петербург
Владивосток	0	7062	6416	7023	6540
Минск	7062	0	673	403	688
Москва	6416	673	0	839	634
Рига	7023	403	839	0	486
Санкт-Петербург	6540	688	634	486	0

Начальный пункт	Конечный пункт	Дальность полета
Владивосток	Минск	7062
Владивосток	Москва	6416
Владивосток	Рига	7023
Владивосток	Санкт-Петербург	6540
Минск	Владивосток	7062
Минск	Москва	673
Минск	Рига	403
Минск	Санкт-Петербург	688
Москва	Владивосток	6416
Москва	Минск	673
Москва	Рига	839
Москва	Санкт-Петербург	634
Рига	Владивосток	7023
Рига	Минск	403
Рига	Москва	839
Рига	Санкт-Петербург	486
Санкт-Петербург	Владивосток	6540
Санкт-Петербург	Минск	688
Санкт-Петербург	Москва	634
Санкт-Петербург	Рига	486

Рис. 4.5. Тернарное отношение "дальность полета", представленное в виде графа, матрицы и таблицы

Глава 5

Бинарные отношения

Арифметические знаки — это записанные фигуры, а геометрические фигуры — это нарисованные формулы.

Д. Гильберт

Среди всевозможных отношений особое место занимают бинарные (двухместные) отношения между объектами из двух множеств. Во-первых, они встречаются повсюду, где какие-нибудь объекты сравниваются между собой или рассматриваются в некоторой связи. Во-вторых, бинарные отношения, как бы хитроумно они ни были определены, достаточно просты, чтобы их можно было изучать математическими средствами вне зависимости от конкретного определения.

В *главе 4* мы рассмотрели определение *n*-арного отношения, частным случаем которого является бинарное отношение.

Определение 5.1

Бинарное отношение есть пара $\langle A \times B, R \rangle$, где $A \times B$ — область задания отношения, а $R \subseteq A \times B$ — график отношения.

В дальнейшем под отношением мы будем понимать его график. Так, говоря о некотором отношении R, заданном между элементами множеств A и B, мы будем иметь в виду некоторое подмножество R декартова произведения множеств A и B, т. е. $R \subseteq A \times B$. С другой стороны, любому подмножеству $F \subseteq A \times B$ соответствует отношение F.

Тот факт, что x и y находятся в отношении R, можно записать любым из следующих трех способов:

❒ xRy — например, "Саша любит Машу";

- $(x, y) \in R$ — кортеж (x, y) принадлежит отношению R; например, (Саша, Маша) $\in$ Любит;
- $R(x, y)$ — предикат, который определяет, находятся ли объекты x и y в отношении R; например, Любит(Саша, Маша).

5.1. Операции над бинарными отношениями

Над бинарными отношениями можно выполнять определенные операции, получая в результате другие отношения. Коль скоро бинарное отношение является множеством, к нему применимы теоретико-множественные операции, такие как объединение, пересечение и вычитание. Но поскольку элементы отношения являются упорядоченными парами объектов, к отношению применимы и специальные операции, определенные с учетом данного обстоятельства.

5.1.1. Теоретико-множественные операции

К теоретико-множественным операциям над отношениями относятся объединение, пересечение и вычитание. Эти операции применимы только к таким отношениям, у которых одинаковы области задания. Точнее говоря, если $R \subseteq A \times B, S \subseteq C \times D$ — два отношения, то теоретико-множественные операции применимы к ним только тогда, когда $A \times B = C \times D$.

Итак, пусть даны два отношения: $R_1 \subseteq A \times B$ и $R_2 \subseteq A \times B$. Определим теоретико-множественные операции над этими отношениями.

- Объединение $R_1 \bigcup R_2$ отношений R_1 и R_2:

  $(x, y) \in R_1 \bigcup R_2$ тогда и только тогда, когда $(x, y) \in R_1$ или $(x, y) \in R_2$.
- Пересечение $R_1 \bigcap R_2$ отношений R_1 и R_2:

  $(x, y) \in R_1 \bigcap R_2$ тогда и только тогда, когда $(x, y) \in R_1$ и $(x, y) \in R_2$.
- Разность (результат вычитания) $R_1 \setminus R_2$ отношений R_1 и R_2:

  $(x, y) \in R_1 \setminus R_2$ тогда и только тогда, когда $(x, y) \in R_1$ и $(x, y) \notin R_2$.

5.1.2. Обратное отношение

Для любого бинарного отношения существует противоположное или, иначе говоря, обратное отношение.

Например, для отношения "x — родитель y" обратным будет отношение "y — ребенок x"; для отношения "Саша — брат Маши" обратным является отношение "Маша — сестра Саши"; для отношения $x \le y$ обратным является отношение $y \le x$ или, что то же самое, $x \ge y$. Теперь дадим точное определение обратному отношению.

Определение 5.2

Пусть $R \subseteq A \times B$ — некоторое отношение. Тогда обратное отношение, обозначаемое как R^{-1}, есть подмножество $R^{-1} \subseteq B \times A$, определяемое следующим образом: $(x, y) \in R^{-1}$ тогда и только тогда, когда $(y, x) \in R$.

Операция, с помощью которой из исходного отношения получается обратное ему, называется *операцией обращения* отношения.

Если в графе, изображающем отношение R, изменить направление стрелок на противоположное (при этом петли, если они есть, остаются на месте), то получится граф обратного отношения R^{-1}. Матрица отношения R^{-1} получается из матрицы отношения R путем замены строк столбцами. Таблица обратного отношения получается из исходной путем изменения порядка столбцов.

У операции обращения отношения есть важное свойство: $(R^{-1})^{-1} = R$.

Доказательство. Действительно, $x(R^{-1})^{-1}y$ равносильно тому, что $yR^{-1}x$, а последнее равносильно тому, что xRy.

Ниже приведены еще несколько свойств операции обращения отношения:

$$(R_1 \bigcup R_2)^{-1} = R_1^{-1} \bigcup R_2^{-1};$$

$$(R_1 \bigcap R_2)^{-1} = R_1^{-1} \bigcap R_2^{-1};$$

$R_1 \subseteq R_2$ тогда и только тогда, когда $R_1^{-1} \subseteq R_2^{-1}$ (изотонность обращения, т. е. сохранение отношения включения при обращении).

5.1.3. Композиция отношений

Если, например, Антон — отец Барбары, а Барбара — мать Василия, то Антон — дед Василия. Отношение "быть дедом" получено из отношений "быть отцом" и "быть матерью" с помощью операции композиции. Теперь дадим точное определение этой операции.

Определение 5.3

Пусть имеются два бинарных отношения $R_1 \subseteq A \times B$ и $R_2 \subseteq B \times C$. Тогда композиция $R_1 \circ R_2$ этих отношений есть подмножество $R_1 \circ R_2 \subseteq A \times C$, определяемое следующим образом:

$(x, y) \in R_1 \circ R_2$ тогда и только тогда, когда существует объект $z \in B$, такой что $(x, z) \in R_1$ и $(z, y) \in R_2$.

Операция композиции обладает следующими свойствами:

$$(R_1 \circ R_2)^{-1} = R_1^{-1} \circ R_2^{-1};$$

$$(R_1 \circ R_2) \circ R_3 = R_1 \circ (R_2 \circ R_3);$$

$$(R_1 \bigcup R_2) \circ R_3 = (R_1 \circ R_3) \bigcup (R_2 \circ R_3);$$

$$(R_1 \bigcap R_2) \circ R_3 \subseteq (R_1 \circ R_3) \bigcap (R_2 \circ R_3);$$

если $R_1 \subseteq R_2$, то $R_1 \circ R_3 \subseteq R_2 \circ R_3$.

5.2. Отображения как бинарные отношения

В математике, и не только, часто используют понятие соответствия, отображения, функции и т. п. Соответствие определяется некоторым правилом сопоставления одних объектов другим. Каким бы способом ни было сформулировано это правило, суть соответствия заключается в том, что оно устанавливается между элементами двух множеств и, стало быть, является бинарным отношением. Поэтому бинарные отношения вида $\langle A \times B, R \rangle$ называют еще *соответствиями* между элементами множеств A и B. При этом используется обозначение $R: A \to B$. Пусть элементу $x \in A$ соответствует некоторое подмножество элементов из множества B. Это подмножество обозначается как $R(x)$ и называется *образом* элемента x. Элементу $y \in B$ соответствует подмножество $R^{-1}(y)$ элементов из множества A. Здесь R^{-1} — соответствие, обратное соответствию R. Множество $R^{-1}(y)$ называется *полным прообразом* элемента y. В общем случае соответствие может быть определено не для каждого элемента $x \in A$ и, с другой стороны, не каждый элемент $y \in B$ имеет непустой прообраз. Если каждый элемент $x \in A$ имеет непустой образ $R(x)$, то говорят, что соответствие R *всюду определено*. В противном случае соответствие называют *частично определенным*. Множество A всюду определенного соответствия $R: A \to B$ называют его *областью (множеством) определения*, а множество B — *областью (множеством) значений*.

Теперь мы рассмотрим некоторые важные виды соответствий, обладающих дополнительными свойствами.

Определение 5.4

Соответствие (бинарное отношение) $R: A \to B$ называется *отображением*, или *функцией*, из множества A в множество B, если

1. Оно всюду определено.

2. Для любого $x \in A$ его образ $R(x)$ есть одноэлементное множество.

Для отображения $R: A \to B$ множество A называется его *областью определения*, а B — *множеством значений*.

Таким образом, отображение (функция) это всюду определенное однозначное соответствие.

Например, формула $y = \sin x$ задает функцию $\sin: X \to Y$, определенную на множестве всех действительных чисел X и принимающую значения из подмножества Y действительных чисел, определенного как интервал от –1 до 1 включительно.

Формула $x^2 + y^2 = 1$ определяет окружность единичного радиуса с центром в начале прямоугольной системы координат. Однако зависимость y от x не является функциональной из-за неоднозначности: подавляющему большинству значений переменной x соответствует не одно, а два значения переменной $y = \pm\sqrt{1 - x^2}$. Исключением являются только $x = 1$ и $x = -1$. Кроме того, данная зависимость определена не на всем множестве действительных чисел, а только на замкнутом интервале от –1 до 1.

Всюду определенные, но неоднозначные соответствия называют еще *многозначными отображениями* (функциями).

Определение 5.5

Отображение $R: A \to B$ называется *инъективным* (инъекцией, однозначным отображением из A в B), если для любых элементов $x, x' \in A$ из равенства $R(x) = R(x')$ их образов следует равенство $x = x'$ самих элементов, т. е. истинно высказывание:

$$\forall x \forall x' ((R(x) = R(x')) => (x = x')) .$$

Таким образом, в инъективном отображении каждый элемент $y \in B$ имеет не более одного прообраза. Пусть, например, A — множество студентов в одной аудитории, а B — множество стульев в этой аудитории; определим отношение xRy как "x сидит на стуле y". Каждый студент сидит на какомнибудь стуле, причем только на одном, поэтому $R: A \to B$ — отображение. Поскольку на каждом из стульев сидит не больше одного студента (некоторые из стульев могут оказаться незанятыми), $R: A \to B$— инъективное отображение.

ОПРЕДЕЛЕНИЕ 5.6

Отображение $R: A \to B$ называется *сюръективным* (сюръекцией, отображением из A на B), если для любого элемента $y \in B$ существует элемент $x \in A$, такой что его образ $R(x)$ совпадает с элементом y, т. е. истинно высказывание:

$$\forall x \exists y (R(x) = y).$$

Таким образом, в сюръективном отображении каждый элемент $y \in B$ имеет непустой прообраз.

В данной книге отображение множества страниц основного текста в множество разделов сюръективно, т. к. каждый раздел содержит хотя бы одну страницу.

ОПРЕДЕЛЕНИЕ 5.7

Отображение $R: A \to B$ называется *биективным* (биекцией, взаимно однозначным соответствием), если оно одновременно и инъективно, и сюръективно.

Таким образом, биекция или, другими словами, взаимно однозначное соответствие есть всюду определенное однозначное соответствие, при котором каждый элемент $y \in B$ имеет одноэлементный прообраз $R^{-1}(y)$.

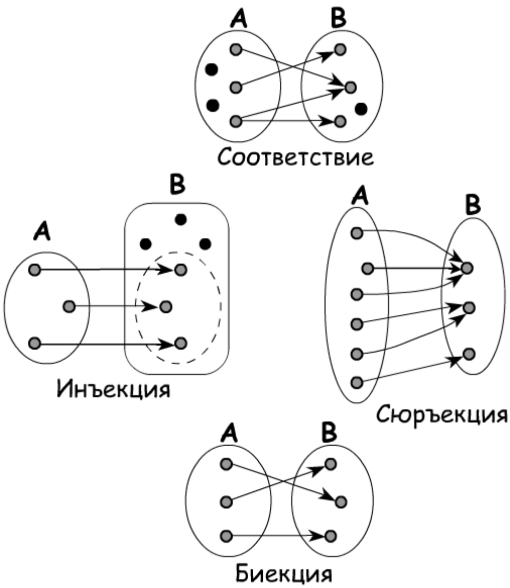


Рис. 5.1. Примеры соответствия, не являющегося отображением, и отображений — инъекции, сюръекции и биекции

В рассмотренном ранее примере отображение множества студентов в множество стульев было инъективным, поскольку некоторые стулья могли быть незанятыми. Если бы таковых не оказалось, то это отображение, оставаясь по-прежнему инъективным, стало бы еще и сюръективным, а в итоге — биективным.

В *главе 3* мы рассматривали взаимно однозначные соответствия между множествами, например, целых и рациональных чисел. Г. Кантор показал, что между множествами натуральных и действительных чисел нельзя установить взаимно однозначное (биективное) соответствие.

На рис. 5.1 показаны примеры соответствия, не являющегося отображением, а также отображений — инъекции, сюръекции и биекции.

К понятиям инъекции, сюръекции и биекции мы еще вернемся в *разд. 5.5*.

5.3. Сечения отношения

Материал данного раздела очень важен для понимания последующего изложения. Поэтому отнеситесь к его изучению наиболее внимательно.

В *разд. 5.2* мы рассматривали бинарные отношения как соответствия вида $R: A \to B$. При этом мы ввели понятия образа и прообраза элемента для данного соответствия. Так, если для элемента $x \in A$ соответствие R определено, то образом этого элемента называется множество всех соответствующих ему элементов множества B; образ элемента x по соответствию R обозначается как $R(x)$. Аналогично определяется прообраз элемента $y \in B$ как множество $R^{-1}(y)$ всех тех элементов $x \in A$, образом которых является y. Другими словами, прообраз элемента $y \in B$ по соответствию R это образ этого элемента по обратному соответствию R^{-1}. При рассмотрении бинарного отношения как множества упорядоченных пар вида $(x, y) \in R$ принято говорить не об образах и прообразах, а о сечении отношения через элемент (иногда сечение называют срезом).

5.3.1. Определения

Определение 5.8

Сечением бинарного отношения через элемент называется множество всех тех элементов, которые находятся в этом отношении с данным элементом. Точнее, сечением отношения $R \subseteq A \times B$ через элемент $x \in A$ называется множество $R(x) = \{y : xRy\}$; сечением этого же отношения через элемент $y \in B$ называется множество $R^{-1}(y) = \{x : xRy\}$.

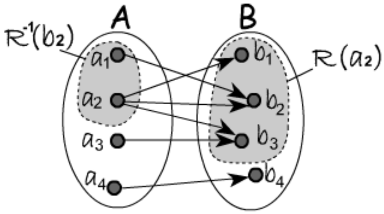


Рис. 5.2. Сечения отношения $R \subseteq A \times B$ через элементы a_2 и b_2

На рис. 5.2 показан пример отношения и его сечений через элементы.

Зная, что такое сечение отношения через единственный элемент, определим сечение через множество элементов. Сечение через множество X естественно определить двумя способами — через операции объединения и пересечения множеств:

1. Как множество всех таких элементов, каждый из которых находится в данном отношении хотя бы с одним элементом множества X ; очевидно, что в этом случае сечение через множество определяется как объединение сечений через все элементы множества X ;
2. Как множество всех таких элементов, каждый из которых находится в данном отношении одновременно со всеми элементами множества X ; очевидно, что в этом случае сечение через множество определяется как пересечение сечений через все элементы множества X .

Теперь дадим точные определения :

Определение 5.9

1. *Сечением 1-го рода* отношения $R \subseteq A \times B$ через непустое множество $X \subseteq A$ называется множество $R(X) = \bigcup_{x \in X} R(x)$; сечением через множество $Y \subseteq B$ называется множество $R^{-1}(Y) = \bigcup_{y \in Y} R^{-1}(y)$. Здесь знак $\bigcup_{x \in X}$ означает операцию объединения по всем элементам $x \in X$.

Для пустого множества положим $R(\text{Ш}) = B, R^{-1}(\text{Ш}) = A$.

2. *Сечением 2-го рода* отношения $R \subseteq A \times B$ через непустое множество $X \subseteq A$ называется множество $R[X] = \bigcap_{x \in X} R(x)$; сечением через множество $Y \subseteq B$ называется множество $R^{-1}[Y] = \bigcap_{y \in Y} R^{-1}(y)$. Здесь знак $\bigcup_{x \in X}$ означает операцию объединения по всем элементам $x \in X$.

Для пустого множества положим $R[\text{Ш}] = B, R^{-1}[\text{Ш}] = A$.

На рис. 5.3 показано некоторое отношение $R \subseteq A \times B$ и его сечения через множество $X = \{a_1, a_2\}$. Сечение 1-го рода $R(X) = R(a_1) \bigcup R(a_2) = \{b_2\} \bigcup \{b_1, b_2, b_3\} = \{b_1, b_2, b_3\}$.

Сечение 2-го рода $R[X] = R(a_1) \bigcap R(a_2) = \{b_2\} \bigcap \{b_1, b_2, b_3\} = \{b_2\}$.

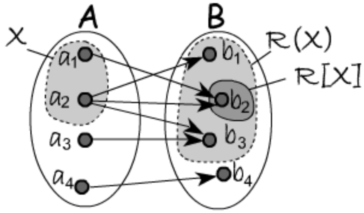


Рис. 5.3. Сечения 1-го рода $R(X)$ и 2-го рода $R[X]$ через множество X

В дальнейшем мы будем широко использовать сечения 2-го рода, поэтому введем для них более удобные, на мой взгляд, обозначения:

$$R[X] = X^{\Delta}, R^{-1}[X] = X^{\nabla}.$$

В частном случае одноэлементного множества $\{x\}$ будем считать для краткости, что $\{x\}^{\Delta} = x^{\Delta}$ и $\{x\}^{\nabla} = x^{\nabla}$. Тогда можно записать:

$$X^{\Delta} = \bigcap_{x \in X} x^{\Delta};$$

$$Y^{\nabla} = \bigcap_{y \in Y} y^{\nabla}.$$

Здесь греческие символы Δ и ∇ называются соответственно *дельта* и *набла*. Эти символы можно рассматривать как обозначения операций сечения (многозначных отображений):

- $\Delta : X \to X^{\Delta}$, где $X \subset A, X^{\Delta} \subseteq B$;
- $\nabla : Y \to Y^{\nabla}$, где $Y \subset B, Y^{\nabla} \subseteq A$.

Примечание

Читатель, знакомый с теорией структур (решеток), может установить, что операции вида $x \to x^{\Delta}$ и $y \to y^{\nabla}$ являются соответствиями Галуа, названными так в честь французского математика Эвариста Галуа.

Пусть, например, $R \subseteq A \times B$ — отношение между множеством A каких-то объектов и множеством B их признаков. Тогда для некоторого непустого подмножества X объектов множество X^{Δ} есть множество всех признаков, которыми обладают все объекты из X. Другими словами, это — общие признаки объектов множества X. Аналогично для непустого подмножества Y признаков множество Y^{∇} есть множество всех таких объектов, которые обладают всеми признаками из Y.

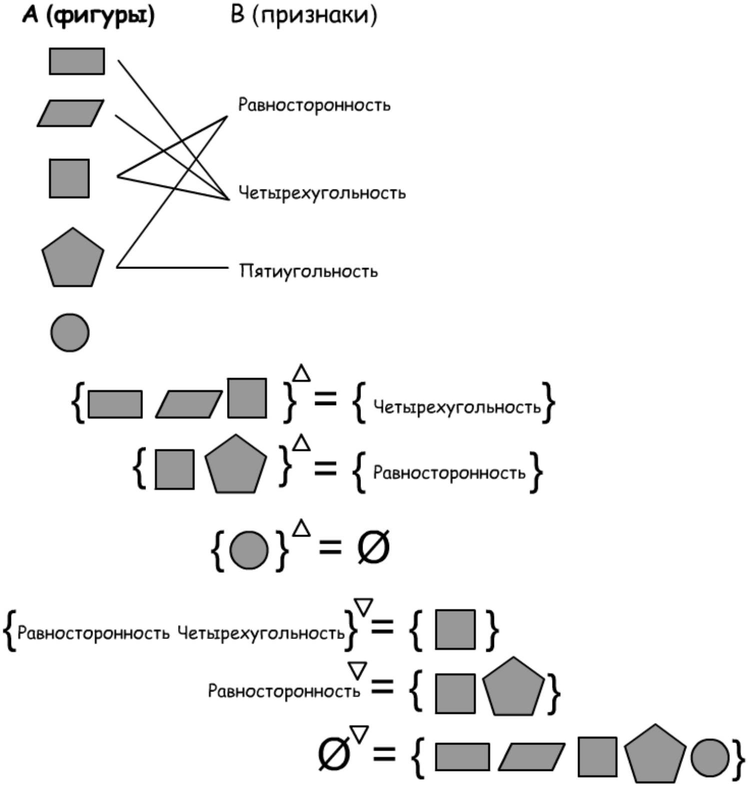


Рис. 5.4. Отношение между геометрическими фигурами и признаками, а также некоторые множества вида X^{Δ} и Y^{∇}

На рис. 5.4 показан пример отношения между геометрическими фигурами и признаками, которыми они обладают или нет, а также некоторые множества вида X^{Δ} и Y^{∇}. Так, некоторые фигуры обладают общими признаками, а другие — нет; некоторые наборы признаков являются общими одновременно для нескольких объектов, другие — только для одного объекта, а некоторым наборам признаков не соответствует ни один объект. В частности, круг вообще не имеет ни одного признака из данного множества всех признаков, поэтому операция Δ, примененная к нему, возвращает пустое множество признаков. Применение операции ∇ к пустому множеству признаков дает по определению множество всех объектов.

Для любого бинарного отношения $R \subseteq A \times B$ можно задать отношение эквивалентности $x \sim x'$ между элементами множества A: $x \sim x'$ тогда и только тогда, когда $x^{\Delta} = x'^{\Delta}$. Для элементов множества B отношение эквивалентности определяется аналогично: $y \sim y'$ тогда и только тогда, когда $y^{\nabla} = y'^{\nabla}$.

Другими словами, два объекта из множества A (B) эквивалентны тогда и только тогда, когда они связаны исходным отношением R (R^{-1}) с одними и теми же объектами из множества B (A). Определенное таким способом отношение эквивалентности называют *индуцированным исходным отношением* R (или R^{-1}), или *ядром отношения* R.

Для бинарных отношений $R \subseteq A \times A$, т. е. заданных, как говорят, на одном множестве A, индуцируются два ядра. Одно ядро индуцируется отношением R, а другое — обратным ему отношением R^{-1}. Однако в таком случае удобно рассматривать одно отношение эквивалентности. Поэтому ядро отношения $R \subseteq A \times A$ определяют так: $x \sim y$ тогда и только тогда, когда $x^{\Delta} = y^{\Delta}$ и $x^{\nabla} = y^{\nabla}$.

Если, например, рассматривается отношение объекты—признаки, то мы считаем два объекта эквивалентными тогда, когда их наборы признаков совпадают. Аналогично, два набора признаков эквивалентны, если ими всеми обладают одни и те же объекты. Так, для отношения между геометрическими фигурами и признаками, показанного на рис. 5.4, эквивалентными являются прямоугольник и параллелограмм: у них один и тот же единственный признак — четырехугольность. Хотя квадрат также имеет этот признак, у него есть еще и равносторонность, т. е. набор признаков квадрата отличается от набора признаков, общих для прямоугольника и параллелограмма.

Мы видели, что операции сечения отношения $\Delta : X \to X^{\Delta}$ и $\nabla : Y \to Y^{\nabla}$ противоположны (обратны) друг другу. Интересными свойствами обладают операции, заключающиеся в последовательном применении операций Δ и ∇:

- $\Delta\nabla : X \to (X^{\Delta})^{\nabla}$ — сначала к множеству X применяется операция Δ, а затем к полученному результату применяется операция ∇;
- $\nabla\Delta : Y \to (Y^{\nabla})^{\Delta}$ — сначала к множеству Y применяется операция ∇, а затем к полученному результату применяется операция Δ.

Для отношения, например, объекты—признаки возможны следующие интерпретации данных операций:

- X^{Δ} — множество признаков, общих для всех объектов множества X; $X^{\Delta\nabla}$ — множество всех объектов, обладающих всеми признаками X^{Δ}.
- Y^{∇} — множество объектов, обладающих всеми признаками множества Y; $Y^{\nabla\Delta}$ — множество всех общих признаков для всех объектов множества Y^{∇}.

На рис. 5.5 показано некоторое бинарное отношение и действие операций Δ и ∇.

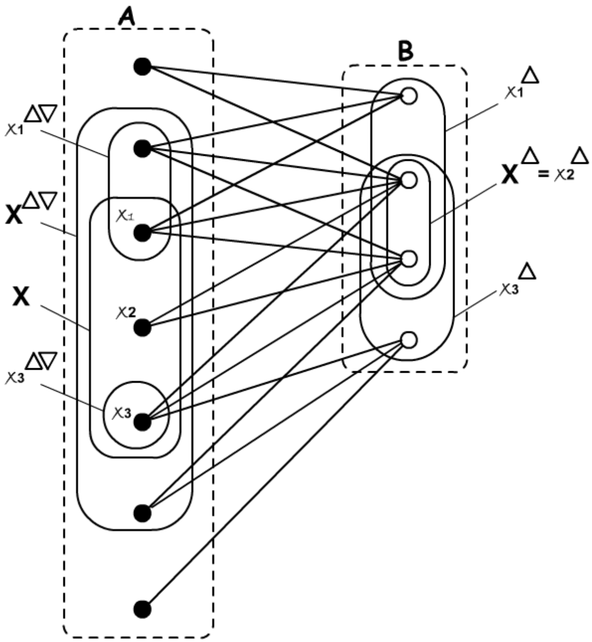


Рис. 5.5. Пример действия операций Δ и ∇

5.3.2. Свойства операций $\Delta\nabla$ и $\nabla\Delta$

Теперь рассмотрим основные свойства операций $\Delta\nabla$ и $\nabla\Delta$.

Теорема 5.1. Пусть дано некоторое бинарное отношение $R \subseteq A \times B$ и произвольные подмножества $X \subseteq A, X' \subseteq A, Y' \subseteq B, Y \subseteq B$. Тогда справедливы следующие три соотношения:

1. $X \subseteq X^{\Delta\nabla}$ $(Y \subseteq Y^{\nabla\Delta})$ — экстенсивность (результат операции не меньше исходного множества);

2. Если $X \subseteq X'$, то $X^{\Delta\nabla} \subseteq X'^{\Delta\nabla}$ (если $Y \subseteq Y'$, то $Y^{\nabla\Delta} \subseteq Y'^{\nabla\Delta}$) — изотонность (сохранение порядка включения);
3. $X^{\Delta\nabla} = (X^{\Delta\nabla})^{\Delta\nabla}$ $(Y^{\nabla\Delta} = (Y^{\nabla\Delta})^{\nabla\Delta})$ — идемпотентность (повторное применение операции не изменяет результат первого применения этой же операции).

Доказательство. Свойство 1. Если $X = Ш$, то $X^{\Delta\nabla} = Ш^{\Delta\nabla} = B^{\nabla}$ и очевидно, что $Ш \subseteq X^{\Delta\nabla}$. Если $X \neq Ш$, то найдется элемент $x \in X$, такой что для всех $y \in X^{\Delta}$ имеет место xRy. Но это означает, что $x \in \bigcap_{x \in X^{\Delta}} R(y) = X^{\Delta\nabla}$. Таким образом, $X \subseteq X^{\Delta\nabla}$. Аналогично доказывается включение $(Y \subseteq Y^{\nabla\Delta})$.

Свойство 2. Покажем сначала, что если $X \subseteq X'$, то $X'^{\Delta} \subseteq X^{\Delta}$ (изменение включения на противоположное). Если $X = Ш$, то $X^{\Delta} = B$ и $X'^{\Delta} \subseteq X^{\Delta} = B$. Если $X \neq Ш$, то

$$X'^{\Delta} = \bigcap_{x \in X'} x^{\Delta} = (\bigcap_{x \in X} x^{\Delta}) \bigcap (\bigcap_{x \in X' \setminus X} x^{\Delta}) \subseteq \bigcap_{x \in X} x^{\Delta} = X^{\Delta}.$$

Таким образом, $X'^{\Delta} \subseteq X^{\Delta}$. Аналогично доказывается включение $Y'^{\nabla} \subseteq Y^{\nabla}$.

Итак, операции Δ и ∇ изменяют включение множеств на противоположное. Поэтому если $X'^{\Delta} \subseteq X^{\Delta}$, то $X^{\Delta\nabla} \subseteq X'^{\Delta\nabla}$ (аналогично, если $Y'^{\nabla} \subseteq Y^{\nabla}$, то $Y^{\nabla\Delta} \subseteq Y'^{\nabla\Delta}$), что и требовалось доказать.

Свойство 3. С одной стороны, в силу свойства 1 имеем $X^{\Delta\nabla} \subseteq (X^{\Delta\nabla})^{\Delta\nabla}$. С другой стороны, в силу этого же свойства имеем $X^{\Delta} \subseteq (X^{\Delta})^{\nabla\Delta}$, откуда с учетом доказанного выше свойства операций Δ и ∇ изменять включение множеств на противоположное имеем $((X^{\Delta})^{\nabla\Delta})^{\nabla} \subseteq (X^{\Delta})^{\nabla}$, т. е. $(X^{\Delta\nabla})^{\Delta\nabla} \subseteq X^{\Delta\nabla}$. Из включения одного множества в другое и наоборот следует их равенство. Таким образом, $X^{\Delta\nabla} = (X^{\Delta\nabla})^{\Delta\nabla}$. Аналогично доказывается, что $Y^{\nabla\Delta} = (Y^{\nabla\Delta})^{\nabla\Delta}$.

На рис. 5.6 приведена схема действия операций Δ и ∇.

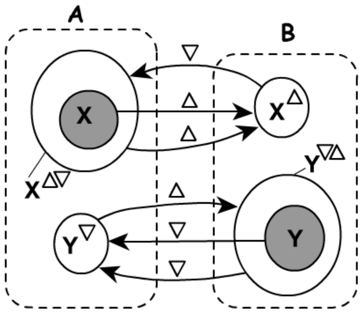


Рис. 5.6. Схема действия операций Δ и ∇

5.3.3. Классы и ко-классы

Среди всевозможных подмножеств X множества A есть такие, которые устойчивы (инвариантны) относительно операции $\Delta\nabla$. Другими словами, устойчивые подмножества остаются неизменными в результате применения к ним операции $\Delta\nabla$, т. е. $X = X^{\Delta\nabla}$. Аналогично, в множестве B имеются подмножества Y, такие что $Y = Y^{\nabla\Delta}$. Так, по свойству 3 (теорема 5.1) устойчивыми являются множества $X^{\Delta\nabla}$ и $Y^{\nabla\Delta}$, где $X \subseteq A, Y \subseteq B$ — произвольные подмножества: сколько бы раз мы ни применяли операцию $\Delta\nabla$ к множеству $X^{\Delta\nabla}$, в результате будем получать одно и то же множество $X^{\Delta\nabla}$ (рис. 5.6).

Кроме множеств вида $X^{\Delta\nabla}$ и $Y^{\nabla\Delta}$, устойчивыми относительно операций $\nabla\Delta$ и $\Delta\nabla$ соответственно являются множества вида X^{Δ}, Y^{∇}, о чем гласит следующая теорема.

Теорема 5.2. Пусть дано некоторое бинарное отношение $R \subseteq A \times B$. Тогда для любых подмножеств $X \subseteq A, Y \subseteq B$ выполняются следующие равенства:

$$X^{\Delta} = (X^{\Delta})^{\nabla\Delta};$$

$$Y^{\nabla} = (Y^{\nabla})^{\Delta\nabla}.$$

Доказательство. Действительно, по свойству 1 имеет место включение $X^{\Delta} \subseteq (X^{\Delta})^{\nabla\Delta}$. С другой стороны, по этому же свойству выполняется включение $X \subseteq X^{\Delta\nabla}$, откуда по свойству 2 получаем $(X^{\Delta\nabla})^{\Delta} \subseteq X^{\Delta}$. Из $X^{\Delta} \subseteq (X^{\Delta})^{\nabla\Delta}$ и $(X^{\Delta\nabla})^{\Delta} \subseteq X^{\Delta}$ следует равенство $X^{\Delta} = (X^{\Delta})^{\nabla\Delta}$. Аналогично доказывается равенство $Y^{\nabla} = (Y^{\nabla})^{\Delta\nabla}$.

Тот факт, что множества вида X^{Δ}, Y^{∇} устойчивы относительно операций соответственно $\nabla\Delta$ и $\Delta\nabla$, иллюстрируется рис. 5.6 (например, если к множеству X^{Δ} применить операцию $\nabla\Delta$, то получим в результате это же множество).

Устойчивые множества вида $X^{\Delta\nabla}$ и $Y^{\nabla\Delta}$ являются подмножествами множеств A и B соответственно, а устойчивые множества вида X^{Δ} и Y^{∇} — подмножествами множеств соответственно B и A. При этом первые устойчивы относительно операций $\Delta\nabla$ и $\nabla\Delta$, а вторые — относительно операций $\nabla\Delta$ и $\Delta\nabla$ соответственно.

Определение 5.10

Пусть дано некоторое бинарное отношение $R \subseteq A \times B$. Обозначим через K множество всех устойчивых подмножеств множества A, а через K' — множество всех устойчивых подмножеств множества B. Элементы этих множеств K и K' будем называть соответственно *классами* и *ко-классами* бинарного отношения R.

5.3.4. Свойства множеств классов и ко-классов

Как устроены множества K классов и K' ко-классов отношения R? Перечислим свойства этих множеств.

- Множество всех классов (ко-классов) образует покрытие множества A (B). В самом деле, каждый элемент $x(y)$ принадлежит по крайней мере одному классу (ко-классу) $x^{\Delta\nabla}$ $(y^{\nabla\Delta})$.
- $A \in K$, $B \in K'$. Действительно, $Ш^{\Delta} = B$ и $Ш^{\nabla} = A$. Равенства выполняются по определению операций Δ и ∇, а то, что $Ш^{\nabla}$ $(Ш^{\Delta})$ есть класс (ко-класс), следует из теоремы 5.2.
- Пересечение классов (ко-классов) есть класс (ко-класс). Это означает, что для подмножеств $X, Y \subseteq A$ выполняется следующее равенство:

  $$X^{\Delta\nabla} \bigcap Y^{\Delta\nabla} = (X^{\Delta\nabla} \bigcap Y^{\Delta\nabla})^{\Delta\nabla};$$

  для подмножеств $X, Y \subseteq B$ имеет место двойственное равенство:

  $$X^{\nabla\Delta} \bigcap Y^{\nabla\Delta} = (X^{\nabla\Delta} \bigcap Y^{\nabla\Delta})^{\nabla\Delta}.$$

  Докажем равенство только для классов (для ко-классов доказательство аналогично). Обозначим $X^{\Delta\nabla} \bigcap Y^{\Delta\nabla}$ через Z. Очевидно, что $Z \subseteq X^{\Delta\nabla}$ и $Z \subseteq Y^{\Delta\nabla}$. Тогда в силу свойств 1 и 2 (теорема 5.1) имеем $Z^{\Delta\nabla} \subseteq X^{\Delta\nabla}$ и $Z^{\Delta\nabla} \subseteq Y^{\Delta\nabla}$, откуда следует $Z^{\Delta\nabla} \subseteq X^{\Delta\nabla} \bigcap Y^{\Delta\nabla} = Z$. В то же время по свойству 1 имеем $Z \subseteq Z^{\Delta\nabla}$. Таким образом, $Z = Z^{\Delta\nabla}$.

  Соответствие $\Delta : K \to K' (\nabla : K' \to K)$ является взаимно однозначным отображением (биекцией), причем $\Delta^{-1} = \nabla (\nabla^{-1} = \Delta)$. В самом деле, пусть $X \in K$, т. е. X — некоторый класс. Тогда $(X^{\Delta})^{\nabla} = X$ и, следовательно, $\Delta\nabla$ — тождественное отображение множества K в себя. Аналогично $\nabla\Delta$ — тождественное отображение множества K' в себя. Следовательно, соответствие $\Delta(\nabla)$ — биекция.

  Между некоторыми классами (ко-классами) выполняется отношение включения $\subseteq$, так что классы (ко-классы) можно частично упорядочить по этому отношению. Тогда частично упорядоченные по включению множества классов и ко-классов антиизоморфны в следующем смысле. Пусть между классами k_1 и k_2 выполняется включение $k_1 \subseteq k_2$. Этим классам взаимно однозначно соответствуют два ко-класса $k_1{}^{\Delta}$ и $k_2{}^{\Delta}$. Тогда между этими ко-классами выполняется противоположное включение: $k_2{}^{\Delta} \subseteq k_1{}^{\Delta}$. Наоборот, если между ко-классами k_1', k_2' выполняется включение $k_1' \subseteq k_2'$, то между соответствующими классами $k_1'^{\nabla}, k_2'^{\nabla}$

выполняется включение $k_2^{\nabla} \subseteq k_1^{\nabla}$. Доказательство данного свойства предлагается читателю в качестве упражнения.

На рис. 5.7 показано некоторое отношение, а также множества всех его классов и ко-классов. При этом множества (классы и ко-классы) расположены так, что множество, объемлющее другое, находится выше, а непосредственно связанные включением множества соединяются линиями. Взаимно однозначное соответствие Δ между классами и ко-классами показано стрелками. Обратное соответствие ∇ между ко-классами и классами можно изобразить противоположно направленными стрелками (на рис. 5.7 не показаны). Нетрудно заметить, что структуры множеств классов и ко-классов подобны (изоморфны) с точностью до наоборот. Поэтому и говорят, что они *антиизоморфны*.

Как построить множества классов и ко-классов на практике? Если множества A и B отношения $R \subseteq A \times B$ конечны и невелики по объему, то возможен следующий алгоритм:

1. Множество $B\,(A)$ — ко-класс (класс).
2. Для всех элементов $x \in A(y \in B)$ строим ко-классы (классы) $x^{\Delta}\,(y^{\nabla})$.
3. Попарные пересечения ко-классов (классов), полученных на предыдущем шаге данного алгоритма, являются ко-классами (классами).
4. Между сформированными классами и ко-классами устанавливаем взаимно однозначное соответствие: от каждого класса k проводим стрелку к ко-классу $k' = k^{\Delta}$.

Возможен и другой алгоритм:

1. Из множества всех подмножеств множества $A(B)$ выбираем все только устойчивые относительно $\Delta\nabla(\nabla\Delta)$ подмножества, т. е. такие подмножества $X \subseteq A(Y \subseteq B)$, что выполняется равенство $X = X^{\Delta\nabla}\,(Y = Y^{\nabla\Delta})$. Это есть множество классов (ко-классов).
2. Между сформированными классами и ко-классами устанавливаем взаимно однозначное соответствие: от каждого класса k проводим стрелку к ко-классу $k' = k^{\Delta}$.

На рис. 5.8 показан еще один пример отношения, а также его структуры множеств классов и ко-классов. Обратите внимание, что в данном случае пустое множество не является ни классом, ни ко-классом.

Графическое изображение взаимосвязанных структур классов и ко-классов можно сделать более компактным, если соответствующие друг другу классы и ко-классы изображать рядом (рис. 5.9).

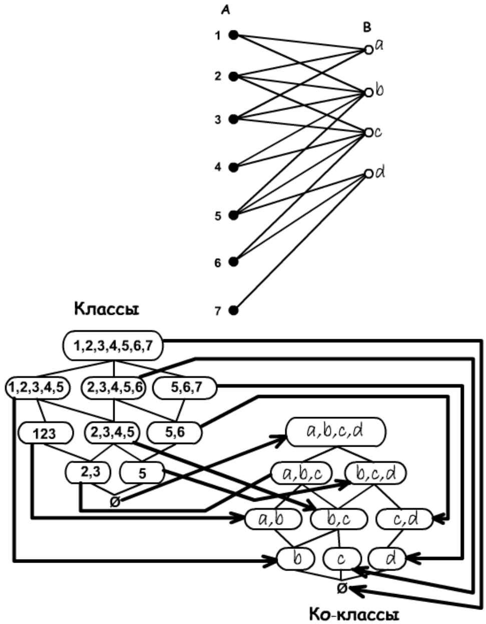


Рис. 5.7. Бинарное отношение (вверху) и множества всех его классов и ко-классов, упорядоченные по отношению включения.
Стрелками показано взаимно однозначное соответствие Δ

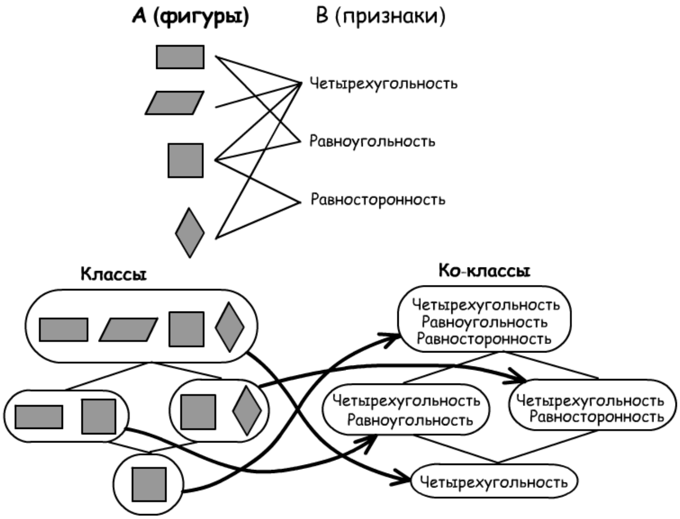


Рис. 5.8. Отношение между фигурами и признаками (вверху) и множества всех его классов и ко-классов, упорядоченные по отношению включения.
Стрелками показано взаимно однозначное соответствие Δ

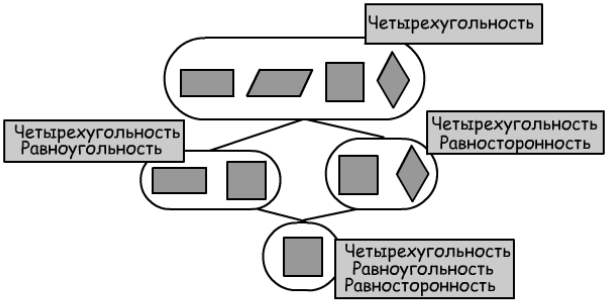


Рис. 5.9. Еще один способ графического представления взаимосвязанных структур классов и ко-классов

Итак, мы теперь умеем представлять бинарные отношения в виде взаимно однозначно соответствующих друг другу множеств классов и ко-классов. Но обратим ли данный переход? Иначе говоря, можно ли по множествам классов и ко-классов и по взаимно однозначному соответствию между ними восстановить исходное отношение между объектами? Оказывается, это возможно, и далее мы рассмотрим, почему и как.

Пусть даны множество K классов, множество K' ко-классов и взаимно однозначное соответствие между ними $\Delta : K \to K' (\nabla : K' \to K)$. Для соответствия Δ определим бинарное отношение $\tilde{R}$, а для соответствия ∇ — бинарное отношение $\tilde{R}'$ следующим образом:

- если $k \in K$ и $k' = k^{\Delta}$, то для всяких $x \in k$ и $y \in k'$ положим $x \tilde{R} y$;
- двойственно, если $k' \in K'$ и $k = k^{\nabla}$, то для всяких $x \in k$ и $y \in k'$ положим $x \tilde{R}' y$.

Теорема 5.3. $\tilde{R} = R = \tilde{R}'$.

Иначе говоря, отношения $\tilde{R}$ и $\tilde{R}'$, определенные указанным выше образом через соответствия Δ и ∇, совпадают с исходным отношением R. А это и означает, что переход от отношения R к соответствию $\Delta : K \to K' (\nabla : K' \to K)$ обратим, т. е. не приводит к потере информации об исходном отношении.

Доказательство. Рассмотрим только равенство $\tilde{R} = R$, поскольку равенство $R = \tilde{R}'$ доказывается аналогично.

Пусть $x \tilde{R} y$, тогда найдутся класс k и ко-класс k', содержащие соответственно элементы x и y (т. е. $x \in k, y \in k'$). При этом $k' = k^{\Delta}$. Поскольку $k' = \bigcap\limits_{z \in k} z^{\Delta}$, то $y \in x^{\Delta}$ и, следовательно, xRy.

Обратно, допустим xRy, тогда $x \in x^{\Delta\nabla}$, $x^{\Delta\nabla}$ — класс. При этом $x^{\Delta\nabla\Delta} = x^{\Delta}$ — ко-класс. Поскольку xRy, то $y \in x^{\Delta}$. Таким образом, нашлись класс $x^{\Delta\nabla}$ и ко-класс x^{Δ}, содержащие элементы x и y соответственно и при этом связанные соответствием Δ (т. к. $(x^{\Delta\nabla})^{\Delta} = x^{\Delta}$). Следовательно, $x\tilde{R}y$. В итоге получаем $\tilde{R} = R$, что и требовалось доказать.

Примечание

Любая операция, обладающая свойствами экстенсивности, изотонности и идемпотентности (см. теорему 5.1), называется *операцией замыкания*, а множества, устойчивые (инвариантные) относительно этой операции, называются *замкнутыми*. Таким образом, операции $\Delta\nabla$ и $\nabla\Delta$ являются операциями замыкания. Множества всех замкнутых подмножеств множества A и всех замкнутых подмножеств множества B образуют структуры, называемые *полными решетками*, в которых наибольшими элементами являются соответственно A и B, а операция умножения означает теоретико-множественное пересечение. Соответствия Δ и ∇ определяют *антиизоморфизм* (дуальный изоморфизм), или *связи Галуа* между полными решетками замкнутых подмножеств множеств A и B. Дополнительно по данному вопросу см. *разд. 7.3*.

5.4. Бинарные отношения на одном множестве

Бинарное отношение на одном множестве — это отношение между двумя объектами, взятыми из одного и того же множества. Другими словами, область задания такого отношения есть множество вида $A \times A$, где A — некоторое множество объектов. Отношение $\langle A \times B, R \rangle$ называют *отношением, заданным на одном множестве*, если $A = B$. Говорят также, что отношение $R \subseteq A \times A$ задано на множестве A, хотя правильнее говорить, что оно задано на декартовом произведении $A \times A$ двух одинаковых множеств. Вместе с тем, бинарные отношения на одном множестве часто обозначают как $\langle A, R \rangle$. Бинарные отношения на одном множестве могут иметь особые свойства и для них могут быть определены специальные операции, к рассмотрению которых мы сейчас и перейдем.

5.4.1. Операции над отношениями

Все операции, определенные для произвольного бинарного отношения, применимы и к бинарным отношениям, заданным на одном множестве. Вместе с тем, для последних могут быть определены еще и другие операции.

Сужение

Определение 5.11

Сужением бинарного отношения $R \subseteq A \times A$ на множество $B \subseteq A$ называется отношение $R' \subseteq B \times B$, определяемое как $R' = R \bigcap (B \times B)$.

Граф отношения R' получается путем удаления из графа отношения R всех вершин из множества $A \setminus B$, а также всех дуг, которые связывают все элементы множества B с элементами множества $A \setminus B$ (рис. 5.10).

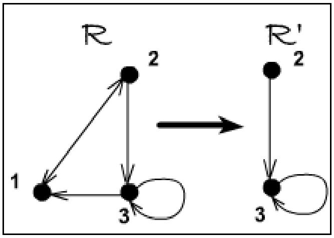


Рис. 5.10. Сужение отношения на множество {2, 3}

Транзитивное замыкание

Определение 5.12

Транзитивным замыканием (транзитивизацией) бинарного отношения $R \subseteq A \times A$ называется отношение $R \subseteq A \times A$, определяемое следующим образом: xRy тогда и только тогда, когда существуют элементы $z_1, z_2, ..., z_n$, такие что xRz_1 и z_1Rz_2 и … и z_nRy. Иначе говоря, xRy только тогда, когда найдется цепочка элементов, связанных отношением R, через которые можно попасть из x в y.

Пусть, например, отношение R определено на множестве населенных пунктов следующим образом: xRy тогда и только тогда, когда из пункта x в пункт y можно попасть, не проходя через другие пункты. Тогда транзитивное замыкание R отношения R есть отношение, определяемое так: xRy тогда и только тогда, когда можно попасть из пункта x в пункт y непосредственно или через промежуточные пункты; при этом из x в y можно попасть непосредственно, если xRy (рис. 5.11).

Транзитивное замыкание можно определить и через операцию композиции отношений *(см. разд. 5.1.3)*:

$$R = R \bigcup R \ \bigcup \ldots \bigcup R^n \bigcup \overset{2}{\ldots},$$

где $R^n = \underbrace{R \circ R \circ \ldots \circ R}_{n}$ (не путайте с n-й степенью множества, т. е. с декартовым произведением).

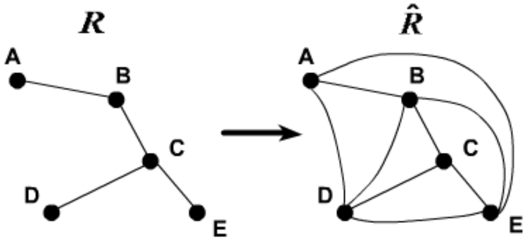


Рис. 5.11. Транзитивное замыкание (транзитивизация) отношения

5.4.2. Свойства отношений

Бинарные отношения, заданные на одном множестве, могут иметь некоторые важные свойства, такие как рефлексивность, симметричность и др. Сочетания этих свойств определяют специальные типы часто встречающихся в жизни отношений, такие как равенство, сходство и порядки различных видов. В данном разделе мы рассмотрим свойства отношений, чтобы затем определить основные типы отношений. Я привожу сначала традиционные определения свойств, а затем — утверждения вида "…тогда и только тогда, когда…", в которых эти свойства выражаются через операции Δ и ∇ (т. е. через классы и ко-классы), рассмотренные в *разд. 5.3.4*. Доказательства этих утверждений предлагается выполнить читателю в качестве упражнения.

Рефлексивность

Если каждый объект находится в некотором отношении сам с собой, то это отношение называется рефлексивным. Точнее, отношение $R \subseteq A \times A$ называется рефлексивным, если для любого элемента $x \in A$ выполняется xRx.

Отношения, например, равенства (=) и нестрогого порядка ($\leq$) на множестве чисел являются рефлексивными, поскольку для любого числа x имеют место $x = x$ и $x \leq x$. В графах рефлексивных отношений все вершины имеют петли, а в матрицах главная диагональ всюду содержит 1.

Отношение $R \subseteq A \times A$ рефлексивно тогда и только тогда, когда для всех $x \in A$ выполняется $x \in x^{\Delta}$ (или $x \in x^{\nabla}$).

Антирефлексивность

Если ни один объект не находится сам с собой в отношении, то последнее называется *антирефлексивным*. Точнее, отношение $R \subseteq A \times A$ называется антирефлексивным, если для любого элемента $x \in A$ не выполняется xRx.

Например, отношение строгого порядка (<) на множестве чисел антирефлексивно. В графах таких отношений все вершины не имеют петель, а в матрицах главная диагональ всюду содержит 0.

Отношение $R \subseteq A \times A$ антирефлексивно тогда и только тогда, когда для всех $x \in A$ выполняется $x \notin x^{\Delta}$ (или $x \notin x^{\nabla}$).

Обратите внимание: просто нерефлексивное отношение не обязательно является антирефлексивным. Отношение называется *нерефлексивным*, если оно не является рефлексивным, т. е. найдется по крайней мере один объект x такой, что не выполняется xRx (т. е. $\exists x((x, x) \notin R)$. В случае антирефлексивного отношения соотношение xRx не выполняется для всех x (т. е. $\forall x((x, x) \notin R)$. Таким образом, если отношение антирефлексивно, то оно и нерефлексивно, но обратное утверждение неверно. На рис. 5.12 приведены примеры рефлексивного, антирефлексивного и нерефлексивного отношений.

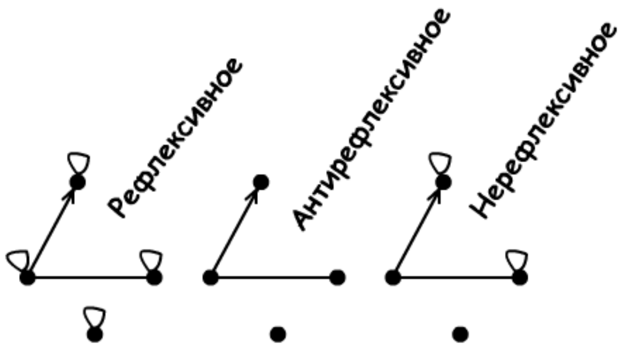


Рис. 5.12. Примеры рефлексивного, антирефлексивного и нерефлексивного отношений

Симметричность

Отношение $R \subseteq A \times A$ называется *симметричным*, если для любых двух объектов $x, y \in A$ из того, что xRy, следует yRx (или, что то же самое, $xR^{-1}y$), и наоборот. Иначе говоря, отношение R симметрично, если $R = R^{-1}$.

Примечание

Напомню, что R^{-1} — отношение, обратное к R.

Например, отношение равенства (=) на множестве всех чисел и "иметь общих родителей" на множестве людей симметричны. В графах симметричных отношений, если из вершины x в вершину y ведет стрелка, то и из вершины y в вершину x также ведет стрелка; вместо двух противоположно направленных стрелок обычно рисуют одну двунаправленную стрелку или вообще не указывают направлений. Матрица симметричного отношения симметрична относительно главной диагонали.

Отношение $R \subseteq A \times A$ симметрично тогда и только тогда, когда для любого подмножества $X \subseteq A$ выполняется равенство $X^{\Delta} = X^{\nabla}$. Таким образом, для симметричного отношения множества всех классов и всех ко-классов совпадают.

Примечание

В определении симметричного отношения можно не требовать выполнения равенства $R = R^{-1}$. Достаточно потребовать включения $R \subseteq R^{-1}$(т. е. если выполнено xRy, то выполнено и yRx), а затем доказать теорему, что отношение R симметрично тогда и только тогда, когда $R = R^{-1}$.

Действительно, по новому определению мы уже имеем $R \subseteq R^{-1}$. Осталось доказать, что верно $R^{-1} \subseteq R$. По свойству изотонности операции обращения *(см. разд. 5.1.2)* из $R \subseteq R^{-1}$ следует, что $R^{-1} \subseteq (R^{-1})^{-1}$, откуда получаем $R^{-1} \subseteq R$. Итак, если отношение симметрично, то $R = R^{-1}$. Обратное утверждение очевидно.

Асимметричность

Отношение $R \subseteq A \times A$ называется *асимметричным*, если для любых объектов $x, y \in A$ из того, что выполнено xRy, следует, что не выполнено yRx, и, наоборот, из выполнения yRx следует невыполнение xRy. Иначе говоря, из xRy и yRx по меньшей мере одно не выполнено (возможно также, что и оба не выполнены). Таким образом, отношение R асимметрично, если $R \cap R^{-1} = Ш$.

Например, отношение строгого порядка (<) на множестве чисел является асимметричным, поскольку для любых двух чисел x, y верно либо $x < y$, либо $y < x$, либо x и y не сравнимы по данному отношению (равные числа). В графах асимметричных отношений нет двунаправленных стрелок или двух противоположно направленных стрелок, соединяющих одни и те же вершины.

Отношение $R \subseteq A \times A$ асимметрично тогда и только тогда, когда для любого подмножества $X \subseteq A$ выполняется равенство $X^{\Delta} \cap X^{\nabla} = Ш$.

Если отношение R асимметрично, то оно антирефлексивно.

Докажем это методом от противного, предположив, что для некоторого x имеет место xRx. Тогда было бы верно и $xR^{-1}x$. Выполнение одновременно xRx и $xR^{-1}x$ означает, что $x(R \bigcap R^{-1})x$ и, следовательно, отношение $R \bigcap R^{-1}$ не пусто, т. е. отношение R не является асимметричным, что противоречит условию нашего утверждения.

Примечание

Обратите внимание: просто несимметричное отношение не обязательно асимметрично.

Антисимметричность

Отношение $R \subseteq A \times A$ называется *антисимметричным*, если для любых объектов $x, y \in A$ соотношения xRy и yRx выполняются только тогда, когда $x = y$ (т. е. x и y — одинаковые объекты).

Введем так называемое *единичное*, или *диагональное* отношение $E \subseteq A \times A$: два объекта $x, y \in A$ находятся в отношении E (т. е. выполняется xEy), если и только если $x = y$. По существу, диагональное отношение — это отношение тождества. Тогда отношение $R \subseteq A \times A$ антисимметрично, если $R \bigcap R^{-1} \subseteq E$.

Например, отношение нестрогого порядка ($\leq$) на множестве чисел является антисимметричным, поскольку для любых двух чисел x, y верно либо $x \leq y$, либо $y \leq x$ и $x \leq x$. В отличие от асимметричного отношения, в графе антисимметричного отношения некоторые вершины могут иметь петли.

Отношение $R \subseteq A \times A$ антисимметрично тогда и только тогда, когда для любого объекта $x \in A$ выполняется равенство $x^{\Delta} \bigcap x^{\nabla} = \{x\}$.

Очевидно, что если отношение асимметрично, то оно антисимметрично, но обратное, вообще говоря, неверно. Антисимметричное, как и асимметричное, отношение несимметрично. На рис. 5.13 приведены примеры симметричного, асимметричного, антисимметричного и несимметричного отношений.

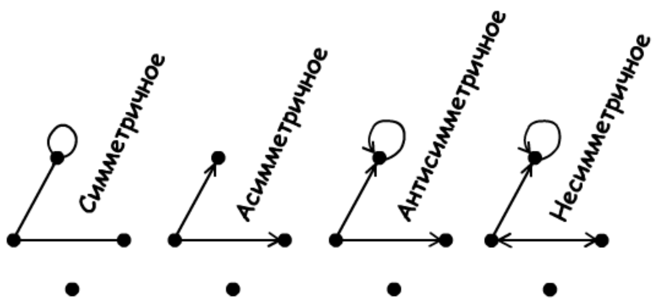


Рис. 5.13. Примеры симметричного, асимметричного, антисимметричного и несимметричного отношений

Транзитивность

Отношение $R \subseteq A \times A$ называется *транзитивным*, если для любых объектов $x, y, z \in A$ из того, что выполняются соотношения xRz и zRy, следует выполнение соотношения xRy. Иначе говоря, отношение R транзитивно, если композиция $R \circ R \subseteq R$. По индукции отсюда следует, что если $xRz_1, z_1Rz_2, ..., z_nRy$, то xRy.

Например, отношения строгого порядка (<) на множестве чисел, а также отношение включения $\subseteq$ между множествами являются транзитивными. В графах транзитивных отношений, если из вершины x в вершину y ведет путь, проходимый по стрелкам, то существует стрелка, непосредственно идущая из x в y.

Отношение R транзитивно тогда и только тогда, когда оно совпадает со своим транзитивным замыканием: $R = \hat{R}$.

На рис. 5.14 показаны примеры транзитивного и нетранзитивного отношений.

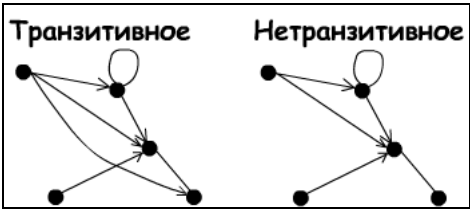


Рис. 5.14. Примеры транзитивного и нетранзитивного отношений

Свойство транзитивности можно определить, правда, несколько более громоздким способом, и в терминах операций Δ и ∇: отношение R транзитивно тогда и только тогда, когда для любого подмножества $X \subseteq A$ выполняется $\bigcup_{x \in X^{\Delta}} x^{\Delta} \subseteq X^{\Delta}$ ($\bigcup_{x \in X^{\nabla}} x^{\nabla} \subseteq X^{\nabla}$) $\cdot$

Примечание

Опеределения всех свойств, кроме транзитивности, через операции Δ и ∇ (т. е. через классы и ко-классы) довольно просты и наглядны. Однако свойство транзитивности в таком же выражении выглядит достаточно сложным. Напомню, что операции Δ и ∇ были определены через сечения отношения 2-го рода: $X^{\Delta} = \bigcap_{x \in X} R(x)$ (двойственно определяется операция ∇); эти операции порождают классы и ко-классы отношения.

Рассмотрим теперь операции $\wedge$ и $\vee$, определенные через сечения отношения 1-го рода: $X^{\wedge} = \bigcup_{x \in X} R(x)$ (двойственно определяется операция $\vee$). Обозначим через K и K' соответственно множества всех классов и ко-классов отношения, полученных
с помощью операций $\wedge$ и $\vee$. Тогда можно доказать, что отношение транзитивно, если и только если для любого класса $k \in K$ (ко-класса $k' \in K$) выполняется $k^{\vee} \subseteq k$ $(k'^{\wedge} \subseteq k')$.

Примечание

Обратите внимание на то, что свойство транзитивности, в отличие от других свойств, формулируется как некая зависимость между тремя объектами, в то время как само отношение выполняется или не выполняется между двумя объектами.

5.4.3. Сочетания свойств отношений

Бинарное отношение на одном множестве может иметь одновременно несколько свойств, рассмотренных в предыдущих разделах. Некоторые сочетания этих свойств определяют особенно интересные отношения, наиболее часто встречающиеся как на практике, так и в математических теориях. Эти отношения мы подробно рассмотрим в следующих разделах, а сейчас лишь укажем соответствия между названиями отношений и наборами их свойств.

Множество особенных отношений (лучше сказать — типов отношений) содержит следующие элементы:

- толерантность;
- эквивалентность;
- квазипорядок;
- нестрогий (частичный) порядок;
- строгий порядок.

Множество свойств отношений содержит такие элементы:

- рефлексивность;
- антирефлексивность;
- симметричность;
- асимметричность;
- антисимметричность;
- транзитивность.

Соответствие между типами отношений и их свойствами является, по существу, определением типов отношений. Данное соответствие само является бинарным отношением и, следовательно, может быть представлено в виде матрицы, графа и связанных структур классов и ко-классов (рис. 5.15). В графе данного соответствия штриховыми линиями отмечено то, что асимметричность и антисимметричность для строгого порядка выводимы из остальных его свойств и, следовательно, их не обязательно включать в определение этого отношения. В скобках указаны условные обозначения типов отношений и их свойств, использованные в изображении структур классов и ко-классов. Из рис. 5.15 видно, что некоторые отношения являются частными случаями других. Так например, эквивалентность является частным случаем толерантности, поскольку она кроме свойств толерантности (рефлексивности и симметричности) имеет еще и свойство транзитивности. В то же время эквивалентность — частный случай квазипорядка, поскольку дополнительно к его свойствам обладает еще и симметричностью.

	Рефлексивность	Антирефлексивность	Симметричность	Асимметричность	Антисимметричность	Транзитивность
Толерантность	+		+			
Эквивалентность	+		+			+
Квазипорядок	+					+
Нестрогий порядок	+				+	+
Строгий порядок		+		+	+	+

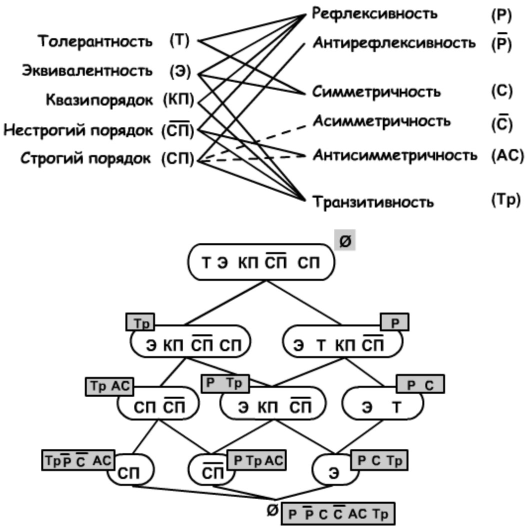


Рис. 5.15. Соответствие между типами бинарных отношений и их свойствами, представленное в трех различных формах

5.4.4. Сходство, или толерантность

Определение и примеры

Определение 5.13

Толерантностью называется рефлексивное и симметричное бинарное отношение на одном множестве.

Это математическое уточнение отношения сходства (похожести) между объектами, хотя основное значение латинского слова tolerantia — терпение. Данное аксиоматическое определение отношения сходства ввел английский математик Зиман, изучавший модели зрительного аппарата, открыв тем самым новый подход к изучению отношений. Дело в том, что традиционный подход заключался в предварительном определении меры сходства (расстояния между объектами), на основе которой затем исследовались взаимное расположение и группировка объектов. Новый подход открывал возможность изучать сходство независимо от того, как конкретно оно определено. Большой вклад в изучение сходства по-новому внес в конце 1960-х годов наш соотечественник Ю. А. Шрейдер. Рассмотрим, почему определенное выше отношение толерантности моделирует интуитивное понятие сходства.

Во-первых, очевидно, что любой объект похож на самого себя, поэтому мы принимаем аксиому, что отношение сходства рефлексивно. Во-вторых, если первый объект похож на второй, то второй похож на первый. Другими словами, два объекта сходны или не сходны независимо от порядка, в котором они сравниваются. Поэтому мы принимаем аксиому симметричности сходства. В принципе, этих двух свойств (рефлексивности и симметричности) уже достаточно, чтобы определить отношение сходства. Разумеется, мы могли бы добавить еще какое-нибудь свойство, например, транзитивность. Однако тогда мы получили бы частный (более специфический) случай отношения сходства. Поэтому остановимся на более общем случае. Но не слишком ли велика эта общность? Быть может, конкретного, содержательного отношения сходства, обладающего только рефлексивностью и симметричностью, просто не существует? Оказывается, это не так. Далее мы рассмотрим содержательные примеры отношения сходства, обладающего только рефлексивностью и симметричностью.

- Как известно, глаз обладает ограниченной разрешающей способностью. Так, если взять несколько точек, расположенных на прямой на некотором расстоянии s друг от друга, то при достаточно малом s мы не сможем различить две соседние точки, но достаточно удаленные точки окажутся различимыми.

- Обычно люди, "взвешивая" предмет в руке, не ощущают разницу в 1 г, и поэтому они не различают предметы массой, например, 10 и 11, 11 и 12, 12 и 13 г, но могут различить 10 и 13 г.
- Схожие объекты могут накапливать незначительные различия так, что в их ряду можно найти совершенно непохожие объекты (рис. 5.16).

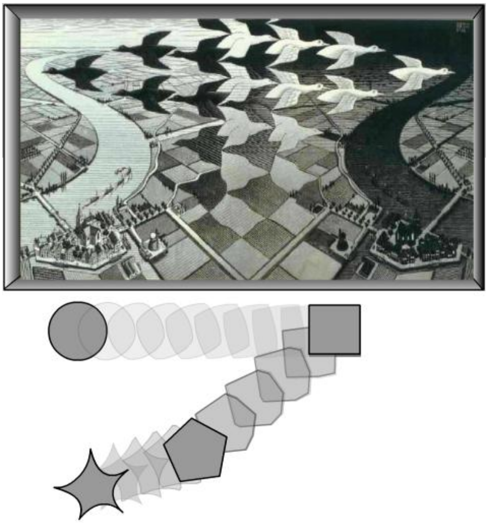


Рис. 5.16. Гравюра голландского художника Эсхера "День и ночь" (вверху) и последовательные трансформации фигур (внизу) показывают, как накопление незначительных различий в сходных объектах приводит к совершенно непохожим объектам

- В множестве четырехбуквенных русских слов, являющихся нарицательными существительными в именительном падеже, два слова будем считать похожими, если они отличаются не более чем одной буквой (точнее, одним вхождением буквы). Тогда можно превратить слово "муха" в слово "слон" (рис. 5.17).
- На рис. 5.18 показано отношение на множестве геральдических и мифических существ. Для упрощения рисунка петли во всех вершинах графа, указывающие на рефлексивность отношения, не показаны. Данное отношение рефлексивно и симметрично, следовательно, оно является толерантностью. Нетрудно заметить, что это отношение не транзитивно и поэтому оно не является эквивалентностью. Так например, кентавр (человеко-лошадь) похож на пегаса (крылатого коня), пегас похож на грифона (крылатого льва), но кентавр на грифона не похож. Данное отношение сходства определялось

следующим образом: два объекта похожи друг на друга, если у них есть общий признак (свойство). При этом использовались следующие признаки:

* иметь крылья;
* быть похожим на льва;
* быть похожим на лошадь;
* быть похожим на человека.

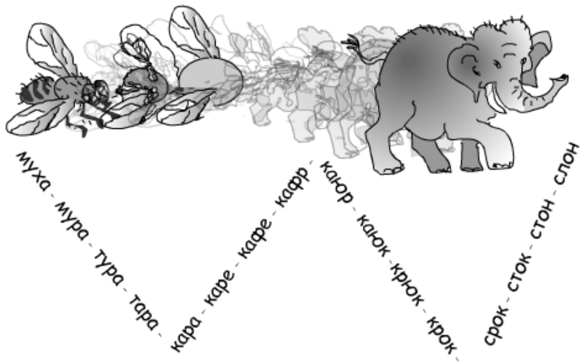


Рис. 5.17. Превращение "мухи" в "слона"

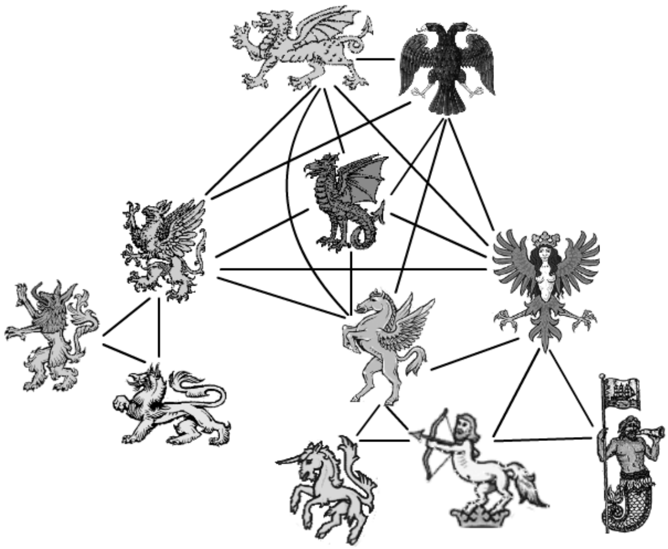


Рис. 5.18. Граф отношения сходства (толерантности) на множестве геральдических и мифических существ (для упрощения рисунка петли в вершинах графа не показаны). Данное отношение рефлексивно и симметрично, но не транзитивно

Если бы к данному списку признаков мы добавили еще какие-нибудь (например, быть похожим на млекопитающее и/или быть похожим на пресмыкающееся), то граф отношения сходства стал бы другим. Однако для того, чтобы новое отношение все же оставалось отношением сходства, а не чем либо иным, необходимо, чтобы оно было рефлексивным и симметричным.

Классы толерантности

Рассматривая произвольное отношение толерантности, можно заметить, что все объекты, между которыми оно задано, можно сгруппировать в классы таким образом, что:

1. В одном классе окажутся все объекты, которые толерантны друг другу (любой объект, толерантный всем объектам данного класса, принадлежит ему); можно сказать, что класс является максимальным подмножеством всех взаимно толерантных объектов.
2. Множество всех таких классов образует покрытие множества всех рассматриваемых объектов, т. е. каждый объект принадлежит хотя бы одному классу.

На рис. 5.19 показана группировка геральдических и мифических существ в классы, обладающие указанными свойствами. Например, в один класс попали пегас, кентавр и единорог, поскольку все они толерантны друг другу (обладают общим признаком схожести с лошадью. При этом кентавр и пегас толерантны и другим объектам, например, алконосту (женщине-птице), но последняя не толерантна одновременно пегасу, кентавру и единорогу, а потому и не входит в их класс. Данные классы, называемые *классами толерантности*, полностью определяют исходное отношение толерантности, поскольку от них всегда можно перейти к исходному графу отношения: все объекты внутри одного и того же класса связаны друг с другом дугами (образуют полные подграфы).

Примечание

Обратите внимание, что указанные классы могут пересекаться (поэтому мы и говорим, что они образуют *покрытие*, а не *разбиение* множества всех объектов).

Итак, класс толерантности содержит все те объекты, которые попарно толерантны друг другу. Объекты, принадлежащие одному и тому же классу толерантности, обладают некими общими признаками, которые и характеризуют (определяют) данный класс толерантности. Объекты, принадлежащие пересечению двух классов толерантности, обладают одновременно признаками, характеризующими и первый, и второй класс. Иначе говоря, они группируются в первый класс по одним общим признакам, а во второй класс — по другим общим признакам. Например, пегас принадлежит классу крылатых, а также классу похожих на лошадь.

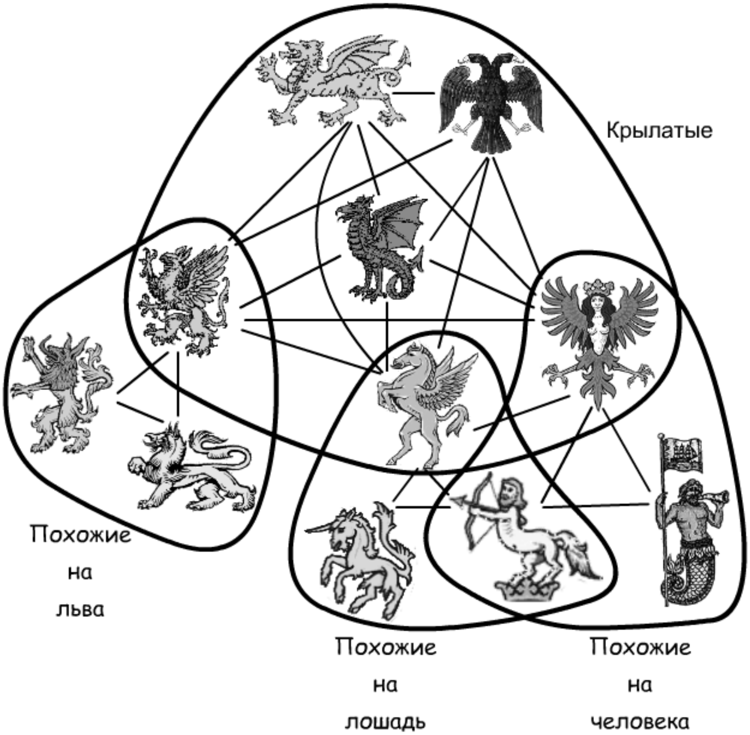


Рис. 5.19. Покрытие множества объектов классами. Объекты из одного класса толерантны (сходны) между собой

Базовые классы толерантности

Как уже отмечалось, множество всех классов толерантности образует покрытие множества объектов. Однако возможен случай, когда не все классы толерантности, а лишь некоторое их подмножество также образует покрытие множества объектов. Например, на рис. 5.19 показаны не все классы толерантности. Так, множество, состоящее из кентавра, пегаса и гарпии, тоже является классом толерантности. Действительно, все объекты этого класса толерантны между собой и не существует другого объекта, толерантного всем этим объектам. Данному классу толерантности соответствует составной признак "быть похожим на человека или на лошадь, или иметь крылья", при этом любые два объекта толерантны благодаря одному элементарному признаку. Так, пегас и гарпия крылаты, гарпия и кентавр имеют человеческую голову, а у кентавра и пегаса общий фрагмент лошадиного туловища. Однако данный класс — лишний с точки зрения возможности восстановления исходного отношения толерантности, поскольку все другие классы, приведенные на рис. 5.19, образуют покрытие множества объектов. Примечательно, что удаление любого из них приводит к тому, что оставшиеся классы уже не обладают этим свойством.

Таким образом, в множестве T всех классов толерантности существует подмножество классов $Б \subseteq T$, элементы которого образуют покрытие множества объектов A и обладающее тем свойством, что при удалении из $Б$ любого элемента оставшиеся элементы уже не образуют покрытия множества A. Такое множество $Б$ классов толерантности называется *базовым*, или *базисом*, а его элементы называются *базовыми классами* толерантности. Например, классы толерантности, показанные на рис. 5.19, образуют базис. Одно и то же отношение толерантности может иметь несколько базисов и тогда предпочтительным является тот из них, который содержит меньше классов. В примере с геральдическими существами базис единственен.

Признаки сходства

Отношение толерантности, как и любое другое бинарное отношение, задается указанием пар объектов, находящихся в данном отношении. Вместе с тем, толерантность можно задать и посредством семейства множеств, образующих покрытие множества объектов: объекты, толерантные друг другу, находятся в некотором множестве из этого семейства и для каждого объекта найдется хотя бы одно такое множество. Точнее, любое покрытие множества объектов порождает (индуцирует) некоторое отношение толерантности. Действительно, пусть дано множество объектов A и семейство множеств $П_1, П_2, \ldots \subseteq A$, образующее покрытие множества A ($A \subseteq П_1 \cup П_2 \cup \ldots$). Определим отношение $\tau \subseteq A \times A$ следующим образом: для любых двух объектов $x, y \in A$ соотношение $x\tau y$ выполняется тогда и только тогда, когда найдется множество $П_i$, такое что $x \in П_i$ и $y \in П_i$ (т. е. множество, содержащее оба объекта). Очевидно, что данное отношение рефлексивно и симметрично и, следовательно, является толерантностью. Заметим также, что множества $П_1, П_2, \ldots$ не обязательно являются классами толерантности, но для каждого $П_i$ найдется содержащий его класс толерантности, т. е. некоторый класс толерантности k, такой что $П_i \subseteq k$. Множества $П_1, П_2, \ldots$, с помощью которых определялось отношение толерантности, естественно интерпретировать как *первичные* признаки (свойства) объектов (все элементы множества Π_i обладают признаком Π_i), а базовые классы этой толерантности можно интерпретировать как *канонические* признаки объектов. На практике первичные признаки можно интерпретировать, например, как свойства объектов, которые непосредственно наблюдаются.

На рис. 5.20 показано покрытие множества из трех объектов тремя множествами, интерпретируемыми как первичные признаки. Признак Π_1 — "быть похожим на лошадь", Π_2 — "быть похожим на человека", Π_3 — "иметь крылья". Это покрытие порождает толерантность, которая имеет в данном случае только один класс толерантности. Ни один из первичных признаков не является классом толерантности. Канонический признак (единственный класс толерантности) можно определить через первичные признаки как $\Pi_1 \bigcup \Pi_2 \bigcup \Pi_3$. Таким образом, канонический признак сходства объектов в рассматриваемом случае описывается так: объекты сходны, если они похожи на лошадь или на человека или имеют крылья.

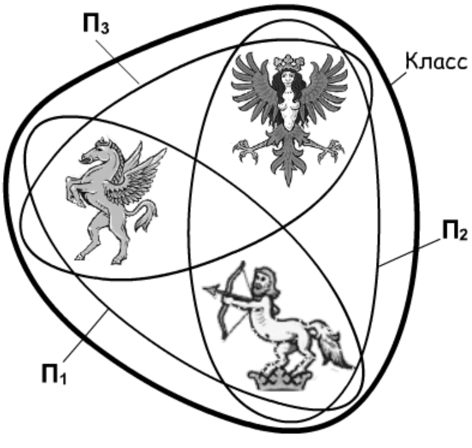


Рис. 5.20. Покрытие множества объектов первичными признаками Π_1, Π_2, Π_3 и класс порожденной толерантности

На рис. 5.21 показано покрытие, каждый элемент которого является классом толерантности, порожденной данным покрытием. Заметим попутно, что классы толерантности k_1, k_2 и k_3 образуют базис (множество базисных классов толерантности) и, таким образом, являются каноническими признаками.

Канонические признаки определяются независимо от того, как задано отношение толерантности, — через первичные признаки или непосредственно как пары сходных объектов. Вместе с тем, канонические признаки всегда могут быть выражены через первичные с помощью операций объединения и пересечения. Об этом говорит следующая теорема.

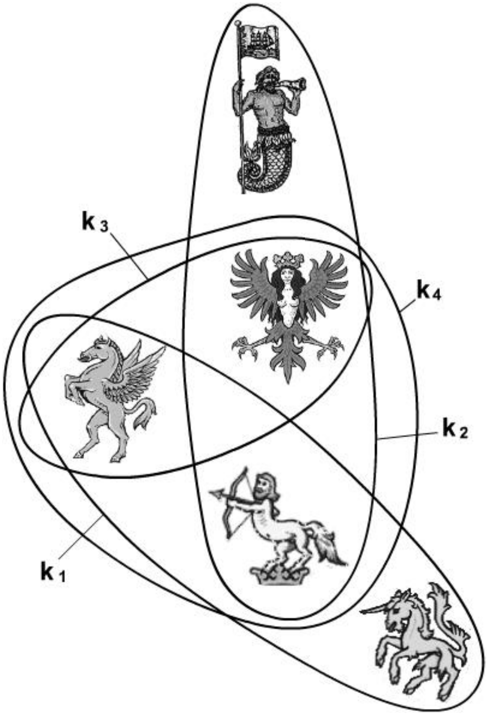


Рис. 5.21. Покрытие, все элементы которого являются классами порожденной толерантности, при этом первые три класса образуют базис

Теорема 5.4. Пусть на множестве объектов A задано покрытие Π, тогда любой класс толерантности, порожденной этим покрытием, всегда может быть выражен через элементы этого покрытия с помощью операций пересечения и объединения.

Доказательство. Возьмем произвольный класс k толерантности τ, пусть $x \in k$. По определению класса для всякого $y \in k$ выполняется $x \tau y$, а по определению толерантности через признаки существует признак $\Pi_{xy} \in \Pi$, такой что $x \in \Pi_{xy}$ и $y \in \Pi_{xy}$. Тогда

1. $x \in \bigcap\limits_{y \in k} \Pi_{xy}$;
2. $\bigcap\limits_{y \in k} \Pi_{xy} \subseteq k$.

Первое следует из того, что $x \in \Pi_{xy}$ для всех признаков Π_{xy}, а второе — из того, что всякий объект $z \in \Pi_{xy}$ толерантен объекту y. Так как y есть произвольный элемент из k, а последний содержит все взаимно толерантные объекты, то $z \in k$. Отсюда вытекает равенство $k = \bigcup_{x \in k}(\bigcap_{y \in k} \Pi_{xy})$, что и требовалось доказать.

Классы толерантности и классы, порожденные операцией $\Delta\nabla$ ($\nabla\Delta$)

Отношение толерантности, как мы в этом только что убедились, характеризуется своими классами толерантности. Вместе с тем оно, как и любое другое бинарное отношение, может быть однозначно представлено и классами (ко-классами), порожденными операцией Δ (∇) (см. *разд. 5.3.3* и *5.3.4*). В общем случае множества тех и других не совпадают. Во избежание путаницы условимся называть классы (ко-классы), порожденные операцией Δ (∇), классами (ко-классами) *отношения* толерантности, а классы, определенные выше в данном разделе, будем называть короче — классами толерантности.

Итак, классы (ко-классы) *отношения* толерантности — это множества, замкнутые (устойчивые) относительно операций $\Delta\nabla$ ($\nabla\Delta$), а классы толерантности — это такие множества, что: 1) каждое из них состоит из всех взаимно толерантных объектов, 2) все вместе они образуют покрытие множества объектов.

Сейчас мы рассмотрим, в чем заключается различие между классами этих типов по существу и какова между ними связь.

Поскольку отношение толерантности рефлексивно и симметрично, то:

1. Для любого объекта $x \in A$ имеет место $x \in x^{\Delta}$ и $x \in x^{\nabla}$ (x^{Δ} и x^{∇} — множества всех тех элементов, которым толерантен объект x).
2. Для любого подмножества объектов $X \subseteq A$ выполняется равенство $X^{\Delta} = X^{\nabla}$, т. е. множества K всех классов и K' всех ко-классов отношения толерантности совпадают (каждый класс является одновременно и ко-классом, и наоборот); взаимно однозначные соответствия (отображения) $\Delta : K \to K'$ и $\nabla : K' \to K$ также совпадают.

Среди свойств классов и ко-классов бинарного отношения, как было показано в *разд. 5.3.4*, имеется такое: неубывающей последовательности классов $k_1 \subseteq k_2 \subseteq k_3 \subseteq ...$ отображение Δ сопоставляет невозрастающую последовательность ко-классов $k_1^{\Delta} \supseteq k_2^{\Delta} \supseteq k_3^{\Delta} \supseteq ...$ (структуры классов и ко-классов антиизоморфны). Таким образом, последовательности классов и ко-классов как бы идут навстречу друг другу и, возможно, совпадают где-то в середине. Такие классы оказываются *устойчивыми* (*инвариантными*) относительно операции Δ. Иначе говоря, устойчивые классы — это такие классы $k \in K$, для которых выполняется равенство $k = k^{\Delta}$. Рассмотрим эти классы.

Лемма 5.1. Устойчивый класс отношения толерантности содержит все объекты, которые толерантны друг другу.

Доказательство. Пусть k — устойчивый класс отношения толерантности $\tau \subseteq A \times A$ и x, y — два произвольных объекта из этого класса. Поскольку $k^{\Delta} = \bigcap_{z \in k} z^{\Delta}$ (по определению операции Δ) и $x, y \in k^{\Delta}$ (т. к. $k = k^{\Delta}$ — устойчивый класс), то $y \in x^{\Delta}$ и $x \in y^{\Delta}$. Следовательно, $y \tau x$ и $x \tau y$. Таким образом, все объекты класса k толерантны между собой. Но, быть может, найдется элемент, толерантный всем объектам класса k и при этом не входящий в этот класс? Покажем, что это не так. В самом деле, поскольку $k = k^{\Delta}$, то для любого объекта $z \notin k$ справедливо $z \notin k^{\Delta}$, т. е. $z \notin \bigcap_{x \in k} x^{\Delta}$. Следовательно, для любого $z \notin k$ найдется $x \in k$, такой что $z \notin x^{\Delta}$, а это означает, что соотношение $x \tau z$ не выполняется (т. е. x и y не толерантны). Таким образом, класс k содержит все толерантные друг другу объекты.

Лемма 5.2. Устойчивые классы отношения толерантности на конечном множестве образуют покрытие этого множества.

Доказательство. Выберем из множества всех объектов A произвольный элемент x и рассмотрим класс x^{Δ}. Очевидно, что одноэлементныйг класс $x^{\Delta} = \{x\}$ является устойчивым. Если x^{Δ} содержит больше одного элемента, то выберем из него элемент $y \neq x$. Далее из множества $x^{\Delta} \setminus \{x, y\}$ выберем элемент z, который толерантен одновременно и x, и y ($x, y \in z^{\Delta}$), затем из $x^{\Delta} \setminus \{x, y, z\}$ выбираем элемент, толерантный одновременно элементам x, y, z. Подобным образом поступаем до тех пор, пока в $x^{\Delta} \setminus \{x, y, z, ...\}$ не останется элементов. Поскольку множество A конечно, то конечно и множество x^{Δ}, и, следовательно, описанная выше процедура формирования множества $\{x, y, z, ...\}$ когда-нибудь завершится. Обозначим сформированное множество через k. Так как все элементы множества k толерантны друг другу, то $k \subseteq k^{\Delta}$.

Предположим теперь, что существует элемент s, такой что $s \in k^{\Delta}$ и $s \notin k$ (т. е. мы предполагаем, что множество k неустойчиво относительно операции Δ). Тогда с учетом $\{x\} \subseteq k$ получим $k^{\Delta} \subseteq x^{\Delta}$ и, следовательно, $s \in x^{\Delta}$. Поскольку $s \in k^{\Delta}$, то объект s толерантен всем элементам множества k и, следовательно, по определению k объект s должен входить в k. Но это противоречит предположению и, таким образом, множество k^{Δ} не может содержать элементы, отличные от элементов множества k, поэтому $k^{\Delta} = k$, а это и означает, что k — устойчивый класс.

Примечание

Лемма 5.2 оказывается верной и при бесконечном множестве A, но для ее доказательства в этом случае требуется использовать метод трансфинитной индукции, который мы в данной книге не рассматриваем.

С учетом договоренности о понимании фраз "класс отношения толерантности" и "класс толерантности" из лемм 5.1 и 5.2 следует

Теорема 5.5. Устойчивые классы отношения толерантности являются классами толерантности.

Примечание

Еще раз подчеркнем, что классы *отношения* толерантности — это подмножества объектов, устойчивые относительно операции $\Delta\nabla$, а классы толерантности — это классы отношения толерантности, устойчивые относительно операции Δ.

Итак, классы толерантности образуют подмножество множества классов отношения толерантности. Обозначим через T множество всех классов толерантности, тогда $T \subseteq K$.

Пусть $\varphi_T{:}A \to T$ — соответствие, сопоставляющее любому объекту $x \in A$ множество $\varphi_T(x)$ всех тех классов толерантности из T, в которые он входит. Поскольку множество T образует покрытие множества A, то каждому объекту $x \in A$ соответствует непустое множество классов толерантности из T.

Теорема 5.6. Объекты x и y толерантны тогда и только тогда, когда $\varphi_T(x) \bigcap \varphi_T(y) \neq$ Ш (т. е. когда существуют классы толерантности, которым принадлежат оба этих объекта).

Доказательство. Пусть $x\tau y$ (где τ — отношение толерантности), тогда $y \in x^{\Delta}$. Следовательно, начав с множества $\{x, y\}$, можно построить устойчивый класс k отношения толерантности такой, что $x, y \in k$. Тогда $k \in \varphi_T(x)$ и $k \in \varphi_T(y)$, а это и означает, что $\varphi_T(x) \bigcap \varphi_T(x) \neq$ Ш.

Обратно, пусть $\varphi_T(x) \bigcap \varphi_T(x) \neq$ Ш. Тогда существует класс k, такой что $k \in \varphi_T(x)$ и $k \in \varphi_T(y)$, а потому $x \in k$, $y \in k$ и $x\tau y$.

Данная теорема показывает, что толерантность вполне определяет свои классы толерантности, а последние однозначно определяют толерантность. Подобным образом можно доказать, что такими же свойствами обладает подмножество базовых классов толерантности.

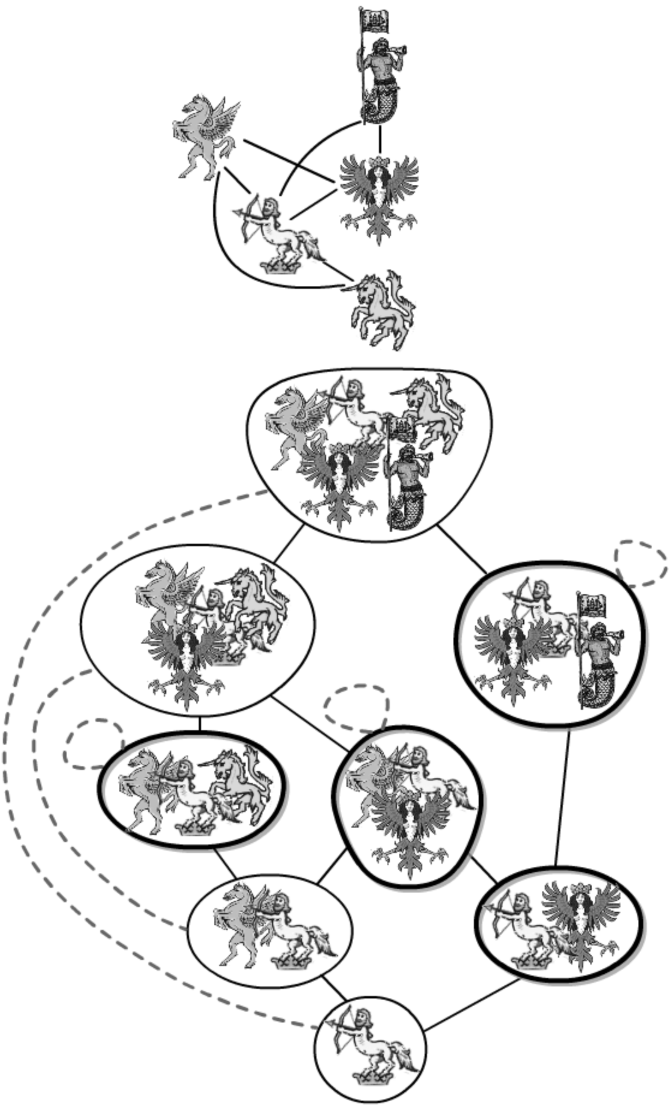


Рис. 5.22. Отношение толерантности и структура его классов, порожденных операцией Δ.
Соответствие Δ между классами указано штриховыми линиями.
Классы толерантности выделены жирными замкнутыми линиями

На рис. 5.22 показано отношение толерантности и структура его классов, порожденных операцией Δ. Как и раньше, в вершинах графа отношения для упрощения не показаны петли, указывающие на рефлексивность. Соответствие Δ между классами и ко-классами показано штриховыми линиями.

Примечание

Напомню, что для симметричного отношения множества всех классов и всех ко-классов, а также соответствия Δ и ∇ между ними, совпадают, поэтому штриховые линии нарисованы без стрелок.

Среди классов *отношения* толерантности есть устойчивые относительно операции Δ, на рисунке такие классы выделены жирными линиями, а также петлевыми штриховыми линиями. Это — классы толерантности, т. е. множества, содержащие все *взаимно* толерантные объекты. Наличие петли около изображения класса как раз и указывает на его устойчивость: соответствие Δ сопоставляет этому классу сам этот класс (операция Δ порождает из этого класса ко-класс, совпадающий с исходным).

Ядро толерантности

Со всяким отношением всегда ассоциируется еще одно, являющееся отношением эквивалентности. Что такое отношение эквивалентности, мы подробно рассмотрим в *разд. 5.4.5*, а сейчас определим отношение эквивалентности, ассоциированное с толерантностью и называемое *ядром толерантности*.

Обозначим ядро толерантности τ через ~, тогда $x{\sim}y$, если $x^{\Delta} = y^{\Delta}$ (или, что то же самое в силу симметричности τ, $x^{\nabla} = y^{\nabla}$). Иначе говоря, два объекта принадлежат ядру, если они толерантны одним и тем же объектам.

Как мы увидим в *разд. 5.4.5*, отношение эквивалентности можно представить как множество непересекающихся классов объектов, такое что все взаимно эквивалентные объекты принадлежат одному и тому же классу, неэквивалентные объекты принадлежат различным классам и для каждого объекта найдется свой класс эквивалентности, которому он принадлежит. Определенное выше ядро толерантности является эквивалентностью и, следовательно, оно вполне представимо множеством не пересекающихся классов, называемых ядрами. Говоря о том, что ядра не пересекаются, мы имеем в виду, что пересечение двух ядер либо пусто, либо они полностью совпадают.

Примечание

Ядро — это отношение, а ядра — классы этого отношения. На рис. 5.23 показан пример отношения толерантности и его ядра. Так, два бескрылых льва принадлежат ядру данной толерантности, поскольку они толерантны одним и тем же объектам. По той же причине два дракона и орел принадлежат одному и тому же ядру. Остальные объекты принадлежат ядру только сами собой, а потому остальные ядра одноэлементные.

Отношение толерантности, все классы ядра которой одноэлементны, называется *безъядерным*.

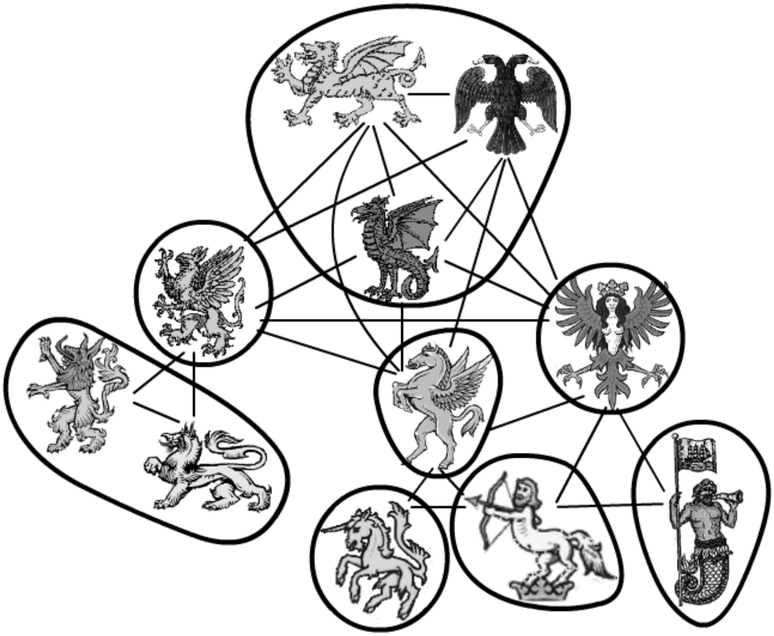


Рис. 5.23. Отношение толерантности и его ядра

Примечание

В общем случае толерантность не определяется своими ядрами. Исключение составляет толерантность, обладающая свойством транзитивности, которая в этом случае есть отношение эквивалентности.

5.4.5. Одинаковость, или эквивалентность

Определение и примеры

Определение 5.14

Эквивалентностью называется рефлексивное, симметричное и транзитивное бинарное отношение на одном множестве.

Эквивалентность является частным случаем толерантности, поскольку обладает всеми ее свойствами, а также еще одним — транзитивностью. Это математическое уточнение отношения одинаковости между объектами. Одинаковыми могут быть признаны и вполне различимые объекты, если рассматривать их с некоторой точки зрения, принимая во внимание лишь некоторые их признаки, существенные в данном контексте и/или для данного наблюдателя.

Для одного наблюдателя два объекта одинаковы, а для другого — нет. Таким образом, отношение одинаковости между одними и теми же объектами весьма субъективно. Вместе с тем, аксиоматическое определение отношения эквивалентности фиксирует объективное содержание понятия одинаковости. Рассмотрим, почему определенное выше отношение эквивалентности моделирует интуитивное понятие одинаковости.

Во-первых, если два объекта одинаковы, то они и похожи, т. е. толерантны друг другу. Следовательно, мы должны принять, что отношение одинаковости рефлексивно и симметрично. Однако если объекты похожи, то это еще не означает, что они одинаковы. Для одинаковости объектов необходимо еще что-то, а именно транзитивность этого отношения. Неискушенный читатель легко принимает за аксиомы наличие у одинаковости свойств рефлексивности и симметричности. Однако, как мне кажется, необходимость транзитивности не столь очевидна. Поэтому я приведу несложные рассуждения для обоснования ее включения в определение одинаковости, или эквивалентности, объектов.

Понятие одинаковости (эквивалентности) является обобщением понятия равенства, которое мы всюду явно или неявно использовали в данной книге. Например, множество определяется как совокупность различных объектов. Следовательно, предполагается, что мы умеем определять, какие объекты различны, а какие — нет. Неразличимые объекты равны, а различимые — не равны. В любом множестве каждый его элемент равен только самому себе. Таким образом, равенство между элементами множества понимается как абсолютное тождество: элемент множества уникален и равен (тождествен) только самому себе. Но именно это и есть исконное содержание понятия равенства, которое мы и возьмем за основу.

Попробуем сначала формально определить отношение равенства, а затем както обобщить его, чтобы получить в итоге интуитивно ясное представление отношения эквивалентности. Однако на первом же этапе реализации этого плана мы обнаруживаем следующее осложнение. Отношение равенства, обычно обозначаемое символом "=", которое мы хотим определить, само используется в определениях чего-либо. Мы рискуем войти в порочный круг. Поэтому отношение равенства = так и оставим в его обычном интуитивном понимании, а определим отношение $\equiv$, которое должно унаследовать от отношения = самые существенные его свойства. Отношение $\equiv$ также назовем *равенством*; из контекста будет ясно, о каком именно отношении идет речь, $\equiv$ или =.

Пусть A_E — произвольное непустое множество. Определим на A_E отношение $\equiv$ следующим образом: для любых элементов $x, y \in A$ $x \equiv y$ тогда и только тогда, когда $x = y$. Очевидно, что граф данного отношения представляет собой совокупность изолированных вершин, соответствующих элементам множества A_E, и каждая вершина x имеет петлю, изображающую тот факт, что выполняется $x \equiv x$. Мы можем также представить данное отношение совокупностью одноэлементных множеств, каждое из которых содержит некоторый элемент $x \in A_E$ и, наоборот, для каждого элемента из A_E в этой совокупности найдется множество, содержащее данный единственный элемент. Множества такой совокупности (назовем их классами равенства), образуют покрытие множества A_E, но при этом не пересекаются, т. е. образуют разбиение множества A_E. Нетрудно убедиться, что это отношение рефлексивно, симметрично и транзитивно.

Теперь обобщим определенное выше отношение равенства $\equiv$ так, чтобы получить отношение эквивалентности. Сама собой напрашивается идея такого обобщения: классы отношения эквивалентости могут содержать несколько элементов, а не только один, как в частном случае отношения равенства $\equiv$. Рассмотрим, как можно построить такое отношение.

Назовем A_E *множеством эталонов*, а его элементы — эталонами. Далее, сформируем множество A, содержащее произвольные объекты, но обладающее следующим свойством: каждый элемент $x \in A$ находится в некотором отношении ET с некоторым единственным элементом множества A_E. Естественно интерпретировать отношение (всюду определенное соответствие) ET как отношение "иметь эталон". Соотношение $xETy$ читается как "x имеет эталон y". Итак, каждый объект из A имеет некоторый эталон, причем единственный. Каждому эталону соответствует подмножество объектов, имеющих данный эталон, а любые два множества объектов, имеющих различные эталоны, не пересекаются.

Теперь на множестве объектов A определим отношение эквивалентности: два объекта эквивалентны тогда и только тогда, когда они имеют один и тот же эталон. Граф такого отношения состоит из полных подграфов: все вершины графа, соответствующие подмножеству всех объектов, имеющих общий эталон, связаны друг с другом. Пример отношения "иметь эталон" и индуцированное им отношение эквивалентности между объектами показаны на рис. 5.24.

Итак, мы определили отношение эквивалентности через вспомогательное отношение "иметь эталон". Данная эквивалентность представляется совокупностью непересекающихся множеств (классов эквивалентности), такой что все объекты из одного класса эквивалентны друг другу, а принадлежащие различным классам — нет. Нетрудно убедиться, что это отношение рефлексивно, симметрично и транзитивно. Таким образом, определенная нами эквивалентность является обобщением отношения равенства.

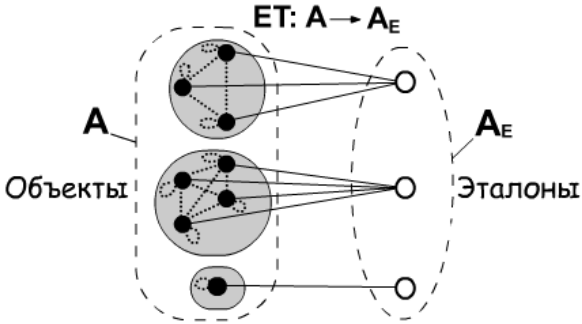


Рис. 5.24. Отношение "иметь эталон" и индуцированное им отношение эквивалентности (показаны пунктиром)

Различимые объекты признаются эквивалентными или неэквивалентными с определенной точки зрения, или по определенному основанию. Так, объекты могут сравниваться по весу, размерам, цвету и другим признакам. Огромное разнообразие признаков, которые могут приниматься во внимание при сравнении объектов, все же ограниченно. Ограничение состоит в том, что признаки должны быть выбраны таким образом, чтобы результат сравнения объектов по этим признакам приводил к разбиению всех объектов на непересекающиеся классы.

Примечание

Напомню, что в случае толерантности (сходства) такое требование к признакам не предъявляется.

Отношение эквивалентности может задаваться и без помощи сравнения с какими бы то ни было эталонами, а непосредственным разбиением множества объектов на непересекающиеся классы, содержащие эквивалентные объекты. Такое отношение состоит из одного или нескольких полных подотношений: все объекты из одного и того же класса связаны друг с другом.

Наконец, отношение эквивалентности можно задать, как и любое другое, непосредственным указанием всех пар эквивалентных объектов, следя за тем, чтобы задаваемое отношение было рефлексивным, симметричным и транзитивным одновременно.

На рис. 5.25 показано разбиение множества геральдических и мифических существ на классы эквивалентности. В качестве признаков классов были выбраны такие, чтобы классы не пересекались. Хотя львы, грифоны и драконы похожи друг на друга, все же по признаку "быть млекопитающим" львы и грифоны ближе друг к другу, чем к драконам. Драконы, идентифицируемые по форме хвоста, относятся к классу пресмыкающихся. К "млекопитающим" я отнес и гарпию, очень похожую на птицу (некие фрагменты ее тела явно указывают на то, что она может вскармливать свое потомство молоком). Человек с рыбьим хвостом, очевидно, может питаться молоком, хотя не исключено, что его мать относится к рыбам. Двуглавый орел ни к одному из уже перечисленных классов не относится и в одиночестве входит в отдельный класс "птиц". На данном множестве объектов отношение эквивалентности можно определить и иначе, важно лишь то, чтобы оно было рефлексивным, симметричным и транзитивным или (что то же самое) чтобы классы эквивалентности не пересекались.

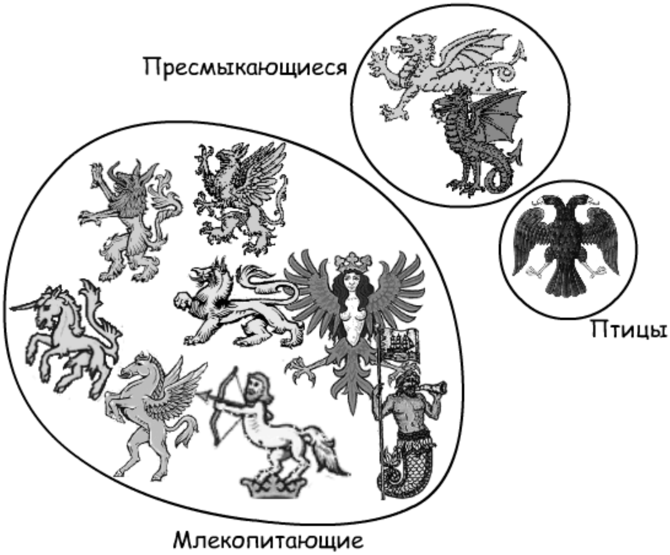


Рис. 5.25. Классы эквивалентности на множестве геральдических и мифических существ

Рассмотрим еще несколько примеров отношения эквивалентности.

- Пусть на множестве объектов A определена функция (однозначное отображение) $f : A \to R$, где R — множество действительных чисел *(см. разд. 5.2)*. Определим на множестве A бинарное отношение ε : $x\varepsilon y$ тогда и только тогда, когда $R(x) = R(y)$. Данное отношение является эквивалентностью и, чтобы убедиться в этом, необходимо проверить, является ли оно рефлексивным, симметричным и транзитивным.

  Изменим немного определение отношения: $x\varepsilon y$ тогда и только тогда, когда $| R(x) - R(y) | \le \delta$, где δ — некоторое положительное число. Иначе говоря, $x\varepsilon y$ выполняется в том и только том случае, если равенство $R(x) = R(y)$ выполняется, как говорят, с точностью до δ. Нетрудно убедиться, что ε — толерантность, а не эквивалентность, поскольку в общем случае свойство транзитивности может быть потеряно.

- Пусть $f(x)$ и $g(x)$ — некоторые алгебраические выражения с переменной x. Определим отношение ε следующим образом: $f(x)\varepsilon g(x)$ тогда и только тогда, когда с помощью заданного набора правил преобразований алгебраических выражений можно из $f(x)$ получить $g(x)$ и, наоборот, из $g(x)$ получить $f(x)$. Например, следующие три выражения находятся в отношении ε:
  - $(x+1)^2$;
  - $(x+1)(x+1)$;
  - x^2+2x+1.

  Отношение ε является эквивалентностью и ее изучению посвящена значительная часть школьной алгебры.
- Пусть $f(x)$ и $g(x)$ — некоторые логические предложения, составленные из одноместных предикатов и операций алгебры логики (конъюнкции, дизъюнкции и отрицания). Отношение равносильности *(см. разд. 2.1 и 2.2)* между логическими предложениями является отношением эквивалентности.

Итак, мы показали, что эквивалентность можно определить двумя способами:

1. Аксиоматически — как отношение, обладающее свойствами рефлексивности, симметричности и транзитивности.
2. Посредством разбиения множества объектов на непересекающиеся классы.

Применяя второй способ, мы приходим к рефлексивному, симметричному и транзитивному отношению, т. е. такому, которое определяется первым способом. В данной книге мы воспользуемся первым способом и покажем, что он приводит к разбиению множества объектов на непересекающиеся классы.

Как уже отмечалось, отношение эквивалентности является частным случаем отношения толерантности. Можно сказать, что эквивалентность — это транзитивная толерантность. Поэтому естественно ожидать проявления в эквивалентности свойств толерантности, а изучение эквивалентности можно провести по тому же плану, что и изучение толерантности: рассмотреть, что такое классы эквивалентности, как они связаны с классами, порожденными операцией $\Delta\nabla$, и т. д. Забегая вперед, скажу, что эквивалентность устроена проще, чем толерантность.

Классы эквивалентности

Классы эквивалентности, как и классы толерантности, это классы *отношения* эквивалентности, устойчивые относительно операции $\Delta(\nabla)$. Обозначим множество всех устойчивых классов через T и будем называть их *классами типа* T. Множество всех классов отношения эквивалентности, порожденных операцией $\Delta\nabla$, обозначим K. Для толерантности, как нам уже известно, имеет место включение $T \subseteq K$, а для эквивалентности соотношение между T и K дает следующая теорема.

Теорема 5.7. Все классы отношения эквивалентности $\varepsilon \subseteq A \times A$, кроме, возможно, A и Ø, являются устойчивыми классами типа T.

Иначе говоря, $T = K'$, где K' — либо равно K, либо равно K за исключением классов A и Ø.

Доказательство. Пусть $k \in K$ $(k \neq Ш, k \neq A)$ — класс, порожденный операцией $\Delta\nabla$. Если $x \in k^{\Delta}$, то для любых $y, z \in k$ имеет место $y\varepsilon x$ и $z\varepsilon x$. В силу симметричности и транзитивности отношения ε имеют место также $y\varepsilon z$ и $z\varepsilon y$. Таким образом, в любом классе k все элементы эквивалентны друг другу. Кроме того, все элементы класса k эквивалентны всем элементам класса k^{Δ} и, следовательно, класс k^{Δ} содержит все элементы класса k. Таким образом, $k \subseteq k^{\Delta}$.

Обратно, если $x \in k$, то для любых $y, z \in k^{\Delta}$ имеет место $x\varepsilon y$ и $x\varepsilon z$. В силу симметричности и транзитивности отношения ε имеют место также $y\varepsilon x$ и $z\varepsilon x$. Таким образом, в любом классе k^{Δ} все элементы эквивалентны друг другу. Кроме того, все элементы класса k^{Δ} эквивалентны всем элементам класса k и, следовательно, класс k содержит все элементы класса k^{Δ}. Таким образом, $k^{\Delta} \subseteq k$.

Из соотношений $k \subseteq k^{\Delta}$ и $k^{\Delta} \subseteq k$ следует равенство $k = k^{\Delta}$, а это и означает, что любой класс, за исключением, возможно, A и Ø, устойчив относительно операции Δ.

Из теоремы 5.7 вытекают два следствия, доказать которые поручается читателю:

1. Если множество A всех элементов, на котором определено отношение эквивалентности ε, является устойчивым относительно операции Δ классом (т. е. $A^{\Delta} = A$), то множество K всех классов *отношения* ε состоит из единственного элемента — класса A. При этом $\varepsilon = A \times A$. Иначе говоря, если A — устойчивый класс отношения эквивалентности ε, то последнее является полным отношением.

2. Если $\varepsilon \neq A \times A$, то пустое множество Ø и множество A являются классами *отношения* эквивалентности ε, причем $Ш^{\Delta} = A$. Так что классы Ø и A неустойчивы относительно операции Δ. Иначе говоря, если эквивалентность ε не является полным отношением, то среди классов этого отношения есть A и Ø, которые неустойчивы относительно операции Δ.

Следующая теорема дает два фундаментальных свойства классов отношения эквивалентности.

Теорема 5.8. Пусть $\varepsilon \subseteq A \times A$ — отношение эквивалентности. Тогда для всех $x, y \in A$ имеют место следующие соотношения:

1. $x \in x^{\Delta}$ — каждый элемент принадлежит классу отношения эквивалентности, порожденному данным элементом.
2. $x\varepsilon y$ тогда и только тогда, когда $x^{\Delta} = y^{\Delta}$, — эквивалентные объекты порождают равные классы отношения эквивалентности и, наоборот, элементы, порождающие равные классы отношения эквивалентности, эквивалентны.

Доказательство. Первое свойство вытекает непосредственно из свойства рефлексивности отношения ε. Докажем второе соотношение. Пусть $x\varepsilon y$, тогда $y^{\Delta} \subseteq x^{\Delta}$. В самом деле, из того, что некоторый элемент $z \in y^{\Delta}$ (а это означает, что $z\varepsilon y$) и $x\varepsilon y$, в силу транзитивности ε вытекает $x\varepsilon z$ или, что то же самое, $z \in x^{\Delta}$. Симметричность ε позволяет подобным образом доказать обратное включение $x^{\Delta} \subseteq y^{\Delta}$. Таким образом, $x^{\Delta} = y^{\Delta}$. Мы доказали, что если $x\varepsilon y$, то $x^{\Delta} = y^{\Delta}$.

Обратно, пусть $x^{\Delta} = y^{\Delta}$. Тогда $x \in y^{\Delta}$ по первому свойству ($x \in x^{\Delta}$), но это означает, что $x\varepsilon y$.

Из первого свойства следует, что каждый объект принадлежит некоторому классу эквивалентности, а из второго — что два класса эквивалентности либо совпадают, либо не пересекаются. Таким образом, множество всех различных классов эквивалентности образует разбиение множества всех объектов.

С другой стороны, каждому разбиению $\mathfrak{R}$ множества A можно сопоставить некоторую эквивалентность ε^{*}, классами которой являются классы данного разбиения: для любых $x, y \in A$ $x\varepsilon^{*}y$ выполняется тогда и только тогда, когда существует класс $R \in \mathfrak{R}$, содержащий элементы x и y. При этом говорят, что разбиение $\mathfrak{R}$ определяет эквивалентность ε^{*}. Говорят также, что произвольная эквивалентность ε определяет разбиение, классами которого являются классы данного разбиения.

Следующая теорема содержит фундаментальный результат о взаимно однозначном соответствии между эквивалентностью и разбиением.

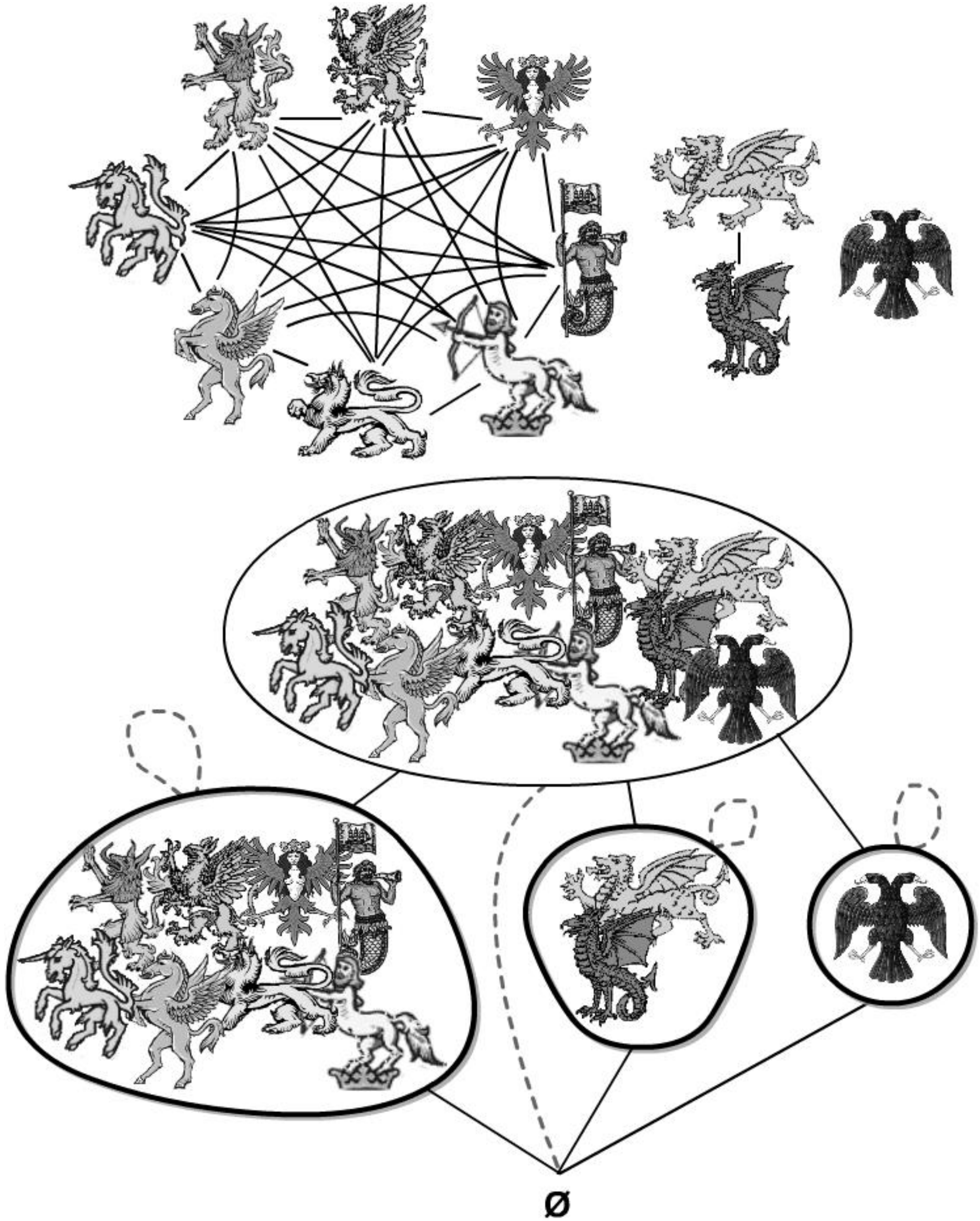


Рис. 5.26. Отношение эквивалентности и структура его классов, порожденных операцией Δ . Соответствие Δ между классами указано штриховыми линиями.
Классы эквивалентности выделены жирными замкнутыми линиями

Теорема 5.9. Если какое-либо отношение эквивалентности $\varepsilon \subseteq A \times A$ определяет разбиение $\Re$ множества A, то отношение эквивалентности, определяемое этим разбиением $\Re$, совпадает с ε, и, наоборот, если некоторое разбиение $\Re$ множества A определяет отношение эквивалентности $\varepsilon \subseteq A \times A$, то разбиение множества A, определяемое этим отношением, совпадает с $\Re$.

Доказательство. Пусть на множестве A задано отношение ε, определяющее разбиение $\Re$ этого множества, и пусть $\Re$ определяет некоторое отношение эквивалентности ε^*. Покажем, что $\varepsilon = \varepsilon^*$. Допустим, что $x\varepsilon y$, тогда $x \in x^\Delta, y \in x^\Delta$ и $x^\Delta \in \Re$. Из определения отношения ε^* следует, что $x\varepsilon^* y$. Таким образом, установлено, что $\varepsilon \subseteq \varepsilon^*$.

Обратно, если $x\varepsilon^* y$, то в $\Re$ найдется такой класс R, что $x \in R$ и $y \in R$. Но R есть некоторый класс эквивалентности ε и поэтому $x\varepsilon y$. Следовательно, $\varepsilon^* \subseteq \varepsilon$.

Из $\varepsilon \subseteq \varepsilon^*$ и $\varepsilon^* \subseteq \varepsilon$ вытекает равенство $\varepsilon = \varepsilon^*$.

На рис. 5.26 показан граф отношения эквивалентности (петли у вершин не показаны), а также структура всех классов отношения эквивалентности. Соответствие Δ между классами и ко-классами показано штриховыми линиями.

Примечание

Напомню, что для симметричного отношения множества всех классов и всех ко-классов, а также соответствия Δ и ∇ между ними совпадают, поэтому штриховые линии нарисованы без стрелок.

Среди классов *отношения* эквивалентности есть классы, устойчивые относительно операции Δ (на рисунке такие классы выделены жирными линиями, а также петлевыми штриховыми линиями). Это классы эквивалентности. Наличие петли около изображения класса как раз и указывает на его устойчивость: соответствие Δ сопоставляет этому классу сам этот класс (операция Δ порождает из этого класса ко-класс, совпадающий с исходным).

Ядро эквивалентности

У всякого произвольного бинарного отношения есть ядро — отношение эквивалентности, ассоциированное с данным бинарным отношением.

Примечание

Напомню, что два объекта находятся в ядре некоторого отношения, если они находятся в этом отношении с одними и теми же объектами.

В *разд. 5.4.4* мы рассматривали ядро толерантности. Сейчас мы выясним, что такое ядро отношения эквивалентности, т. е. какая еще эквивалентность ассоциируется с данной эквивалентностью.

Ядро (обозначим его через ~) произвольной эквивалентности ε совпадает с ней самой. В самом деле, по определению ядра $x \sim y$ тогда и только тогда, когда $x^{\Delta} = y^{\Delta}$, а в силу теоремы 5.8 $x\varepsilon y$ тогда и только тогда, когда $x^{\Delta} = y^{\Delta}$, откуда и следует, что $x \sim y$ тогда и только тогда, когда $x\varepsilon y$.

Множество всех классов эквивалентности $\varepsilon \subseteq A \times A$ называют также *фактор-множеством* множества A и обозначают через A/ε. Например, множество всех классов ядра $Я_R$ некоторого бинарного отношения $R \subseteq A \times B$ есть фактор-множество множества A.

Задание эквивалентности с помощью эталонов

В самом начале *разд. 5.4.5* мы уже рассматривали, как отношение эквивалентности может быть задано посредством отношения "быть эталоном". Однако там мы предполагали, что объекты и эталоны принадлежат разным множествам, подобно тому как гири и взвешиваемые с их помощью предметы — разные множества. Теперь рассмотрим случай, когда объекты и эталоны образуют одно и то же множество.

Пусть на множестве объектов A задано некоторое разбиение $\mathfrak{R} = \{R_1, R_2, ...\}$ на непересекающиеся классы $R_1, R_2, ...$. Мы уже знаем, что такое разбиение задает на множестве A некоторое отношение эквивалентности ε. В каждом классе R_i разбиения выберем по одному какому-нибудь объекту x_i и назовем такой объект *эталоном* данного класса, или его *типичным представителем*. Например, в классе лошадей различных пород можно выбрать некоторую типичную лошадь. Среди собак это сделать труднее (я обычно представляю немецкую овчарку). Впрочем, не важно, что именно вы выберете в качестве эталона класса. Раз имеется разбиение множества объектов на классы, то в качестве эталона класса может выступать любой его элемент.

Определим на множестве A отношение ET "иметь эталон": $xETy$ тогда и только тогда, когда y есть эталон класса, которому принадлежит x. Соотношение $xETy$ читается как "x имеет в качестве эталона y". Очевидно, обратное отношение ET^{-1} можно интерпретировать как "быть эталоном". Соотношение yET^{-1} читается как "y — эталон для x".

Отношение ET не симметрично и не рефлексивно: соотношение $xETx$ выполняется только тогда, когда объект x — эталон некоторого класса (эталон сам себе эталон). Пример отношения "иметь эталон" показан на рис. 5.27.

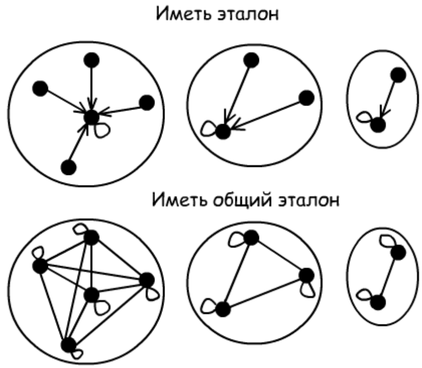


Рис. 5.27. Отношение "иметь эталон" и определенное через него отношение эквивалентности "иметь общий эталон"

Очевидно, отношение эквивалентности ε, заданное через разбиение $\Re$, можно определить и через отношение ET "иметь эталон" следующим образом: $x\varepsilon y$, если и только если $ET(x) = ET(y)$ (объекты x и y имеют один и тот же эталон). Таким образом, ε — отношение "иметь общий эталон".

Представление эквивалентности через отношение "иметь эталон" более компактно, чем непосредственно множеством всех пар эквивалентных друг другу объектов (это хорошо видно на рис. 5.27). Так, только один класс отношения эквивалентности, состоящий из n объектов, представляется полным графом с $n(n-1)$ дугами, если учитывать и петли, а таблица, представляющая этот класс, будет содержать столько же строк. Отношение "иметь эталон" для одного класса эквивалентности потребует лишь n стрелок в графе и n строк в таблице.

5.4.6. Порядки

Что такое порядок?

Располагая элементы некоторого множества друг за другом, мы тем самым упорядочиваем их или устанавливаем между ними некоторое отношение *порядка*. Простейшим примером является естественный порядок натуральных чисел. Его естественность заключается в том, что для любых двух натуральных чисел мы знаем, какое из них следует за другим или какое из них больше другого, поэтому мы можем расположить натуральные числа в последовательности так, что большее число будет расположено, например, правее меньшего: 1, 2, 3, … . Разумеется, последовательность элементов можно выписывать в любом направлении, а не только слева направо. Само понятие натуральных чисел уже содержит в себе идею упорядоченности. Можно сказать, что натуральные числа были созданы для явного воплощения этой идеи. Устанавливая некоторое относительное расположение элементов какого-либо множества, мы тем самым задаем на нем некоторое бинарное отношение порядка, которое в каждом конкретном случае может иметь свое название, например, "быть меньше", "быть старше", "содержаться в", "следовать за" и т. д. Символические обозначения порядка также могут быть разнообразными, например, $\leq$, $\subseteq$, $\rightarrow$ и т. д.

Главным отличительным признаком отношения порядка является наличие у него свойства транзитивности. Так, если мы имеем дело с последовательностью каких-то объектов $x_1, x_2, ..., x_n, ...$, упорядоченных, например, по отношению $\leq$, то из того, что выполняется $x_1 \leq x_2 \leq ... \leq x_n \leq ...$, должно следовать, что для любой пары x_i, x_j элементов этой последовательности также выполняется $x_i \leq x_j$ (рис. 5.28).

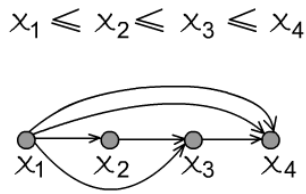


Рис. 5.28. Упорядоченная последовательность элементов и граф соответствующего отношения порядка

Обратите внимание, что для пары элементов $x_i \leq x_j$ в графе отношения мы проводим стрелку от вершины x_i к вершине x_j, т. е. от меньшего элемента к большему. Вместе с тем, транзитивность еще не достаточна, чтобы обладающее ею отношение было порядком. Вспомним, что и эквивалентность транзитивна, но она не ассоциируется нашей интуицией с каким-либо порядком. Впрочем, эквивалентность с точки зрения порядка иногда рассматривают как тривиальный или вырожденный случай отношения порядка. В этом случае любые два объекта либо безразличны для наблюдателя (принадлежат одному классу эквивалентности), либо не сравнимы (принадлежат различным классам эквивалентности). Для того чтобы отношение было порядком, необходимо сочетание транзитивности с другими свойствами или с отсутствием определенных свойств. Так, наличие несимметричности вместе с транзитивностью делают отношение похожим на то, что мы интерпретируем как порядок.

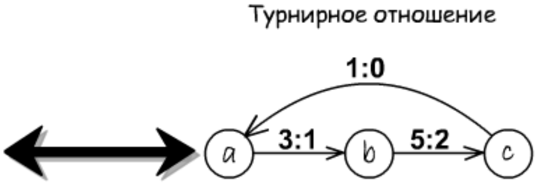
На практике нередко требуется упорядочить элементы какого-нибудь множества, исходя из имеющегося отношения между ними, которое порядком не является. Так, на основе турнирной таблицы соревнований футбольных команд, в которой зафиксированы результаты проведенных матчей, требуется упорядочить команды по местам с учетом *всех* матчей, чтобы выявить призеров турнира. В каждой клетке таблицы зафиксирован результат матча как соотношение между голами, забитыми в ворота противника соответствующими командами. Таким образом, легко определить, кто выиграл, или кто был сильнее, в каждом конкретном матче. Если матч окончился ничьей, то можно считать, например, что в нем победили (проиграли) оба участника. Однако нам требуется указать, кто выиграл весь турнир и, следовательно, занял первое место по итогам соревнований в целом, кто был вторым, третьим и т. д. Иначе говоря, нам необходимо установить отношение порядка, а лучше — строгого порядка, чтобы не делить одно и то же место между несколькими командами.

На рис. 5.29 в качестве примера показана турнирная таблица с итогами всех трех матчей между тремя командами a, b, c и соответствующий ей граф, в котором стрелка направлена от команды-победителя к проигравшей команде, а около стрелки указан счет матча. Это турнирное отношение асимметрично, что необходимо для порядка, но не транзитивно. Так что турнирное отношение не является порядком. Из таблицы и графа видно, что каждая команда одержала одну победу и потерпела одно поражение, следовательно, сумма всех побед не может определить победителя турнира. С другой стороны, команда a победила команду b, которая, в свою очередь, победила c. Хочется признать, что a сильнее c. Однако это — попытка навязать турнирному отношению транзитивность, которой у него нет, но которая должна быть у другого отношения "быть сильнее". Далее мы замечаем, что команда c выиграла у a, что, естественно, разрушает наше начальное представление о том, кто кого сильнее. Остается попробовать учесть отношение или разность забитых и пропущенных мячей. Как по отношению, так и по разности забитых и пропущенных мячей в данном примере получается, что сильнейшей является команда b, на втором месте — a, на третьем — c. Это означает, что команда b сильнее и a, и c, а команда a сильнее только команды c. Таким образом, нам удалось построить асимметричное и транзитивное отношение "быть сильнее", являющееся строгим порядком.

Заметим, что в общем случае может оказаться, что места по отношению мячей и по их разности могут распределиться неодинаково. Какая победа кажется более убедительной, со счетом 3:1 или 5:2? Я не специалист по футболу, но мне кажется предпочтительным счет 5:2. Тогда выясняется, что моему предпочтению соответствует критерий разности забитых и пропущенных мячей ((5 – 2 = 3) > (3 – 1 = 2)). Те, кому более убедительным победным счетом кажется 3:1, руководствуются, видимо, критерием отношения мячей ((3 : 1 = 3) > (5 : 2 = 2,5)).

Турнирная таблица

	a	b	c
a		3:1	0:1
b	1:3		5:2
c	1:0	2:5	



	Забитые и пропущенные мячи		Место в турнире
	отношение	разность	
a	3/2=1.5	3-2=1	2
b	5/3=1.67	5-3=2	1
c	1/2=0.5	1-2=-1	3

Рис. 5.29. Упорядочение команд по данным турнирной таблицы

Квазипорядок

Определение 5.15

Квазипорядком называется рефлексивное и транзитивное бинарное отношение на одном множестве.

Квазипорядок является обобщением эквивалентности, поскольку обладает всеми ее свойствами за исключением симметричности. Напомню попутно, что толерантность также есть обобщение эквивалентности, но по другой причине: из-за необязательности транзитивности.

Квазипорядок — наиболее общее отношение из группы порядков различных видов, которые мы рассмотрим позже. На рис. 5.30 (вверху) показан пример отношения квазипорядка на множестве мифических и вполне реальных существ (петли у вершин графа отношения не показаны, чтобы не загромождать рисунок). Формирование этого отношения подчинялось следующему правилу: объекты связаны данным отношением, если один из них может быть получен путем некоторой модификации другого, причем эта связь симметрична только для мифических существ. Иначе говоря, мифическое существо может быть получено как из мифического, так и из реального, но реальное существо не может быть получено из мифического. В данном примере реальными существами являются кони, а пегас, кентавр и единорог могли произойти и от коней, и друг от друга. Нетрудно заметить, что, грубо говоря, в данном квазипорядке много эквивалентности и мало порядка. Сходством с эквивалентностью квазипорядок обязан симметричным (двунаправленным) связям между некоторыми элементами, а сходством с порядком — несимметричным (однонаправленным) связям.

Поскольку отношение квазипорядка в целом не симметрично, его множества классов K и ко-классов K' не совпадают (рис. 5.30). Тем не менее, для восстановления квазипорядка по K, K' и по отображению $\Delta : K \to K'$ ($\nabla : K' \to K$) достаточно ограничиться знанием лишь одного множества — K или K'. Об этом и говорит следующая теорема.

Теорема 5.10. Пусть $R \subseteq A \times A$ — отношение квазипорядка, K — множество всех его классов, $\varphi : A \to K$ — естественное соответствие, сопоставляющее каждому объекту множества A множество всех классов, которым этот объект принадлежит. Тогда xRy выполняется, если и только если $\varphi(y) \subseteq \varphi(x)$.

Двойственную теорему можно сформулировать для ко-классов K' и естественного соответствия $\psi : A \to K'$: xRy выполняется, если и только если $\psi(x) \subseteq \psi(y)$.

Примечание

Доказательство попытайтесь вывести самостоятельно.

Таким образом, квазипорядок полностью определяется своими классами (ко-классами). Для восстановления отношения квазипорядка по его классам (ко-классам) необходимо выполнить следующее:

1. Все элементы внутри класса (ко-класса) связать друг с другом стрелками в обоих направлениях.
2. Для каждой пары классов $k_1 \subseteq k_2$ провести стрелки от всех элементов меньшего класса ко всем элементам большего класса; в случае ко-классов поступаем противоположным образом: стрелки проводятся от всех элементов большего ко-класса ко всем элементам меньшего ко-класса.

В действительности можно ограничиться только классами вида x^{∇} или ко-классами вида x^{Δ}, т. е. классами или ко-классами, порожденными одиночными элементами множества A. Для квазипорядка, показанного на рис. 5.30, множество всех классов (ко-классов) совпадает с множеством всех классов (ко-классов), порожденных одиночными элементами.

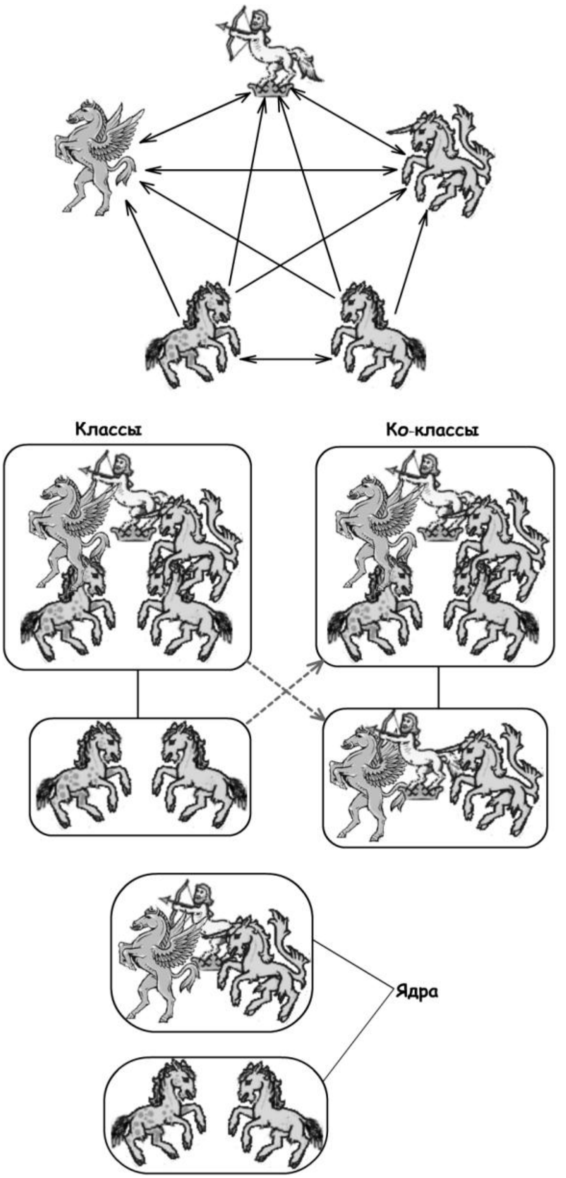


Рис. 5.30. Квазипорядок, структура его классов и ко-классов, а также ядра (петли у вершин графа не показаны)

Квазипорядок — несимметричное отношение, поэтому ядро для него определяется так: $x \sim y$ тогда и только тогда, когда $x^{\Delta} = y^{\Delta}$ и $x^{\nabla} = y^{\nabla}$.

Примечание

Напомню, что в случае симметричных отношений (например, толерантности и эквивалентности) условия $x^{\Delta} = y^{\Delta}$ и $x^{\nabla} = y^{\nabla}$ равносильны друг другу.

На рис. 5.30 показан пример квазипорядка (вверху) и классы ядра (внизу), которые называют ядрами.

Теорема 5.11. Пусть $R \subseteq A \times A$ — квазипорядок. Элементы $x, y \in A$ входят в одно ядро квазипорядка тогда и только тогда, когда $xR \bigcap R^{-1} y$ (или, в другой записи, $(x, y) \in R \bigcap R^{-1}$).

Доказательство. Пусть элементы x, y принадлежат одному и тому же ядру. Тогда выполняются равенства $x^{\Delta} = y^{\Delta}$ и $x^{\nabla} = y^{\nabla}$. Так как в силу рефлексивности квазипорядка $y \in y^{\Delta}, y \in y^{\nabla}$ и $x \in x^{\Delta}, x \in x^{\nabla}$, то $x \in y^{\Delta}, x \in y^{\nabla}$. Поэтому yRx, xRy и, следовательно, $xR^{-1}y$. Из того, что выполняются xRy и $xR^{-1}y$, следует, что выполняется $xR \bigcap R^{-1} y$.

Обратно, пусть $xR \bigcap R^{-1} y$. Следовательно, xRy и yRx. Если некоторый $z \in x^{\Delta}$, то с учетом yRx и транзитивности R имеем yRz, т. е. $z \in y^{\Delta}$. Аналогично, из $z \in y^{\Delta}$ и xRy следует $z \in x^{\Delta}$. Таким образом, $x^{\Delta} = y^{\Delta}$. Таким же способом доказывается равенство $x^{\nabla} = y^{\nabla}$. Итак, мы показали, что из $xR \bigcap R^{-1} y$ вытекает $x^{\Delta} = y^{\Delta}$ и $x^{\nabla} = y^{\nabla}$, т. е. то, что элементы x, y принадлежат одному и тому же ядру.

Редукция порядка

Для порядков важную роль играет так называемая *операция редукции отношения* (от лат. reductio — упрощение). Эта операция в некотором смысле обратна операции транзитивного замыкания *(см. разд. 5.4.1)* и позволяет упрощать граф отношения порядка.

Определение 5.16

Редукцией отношения $R \subseteq A \times A$ называется отношение $R^r \subseteq A \times A$, такое что $R^r = R \setminus R^2$, где $R^2 = R \circ R$ — композиция отношения R самого с собой *(см. разд. 5.1.3)*.

Редукция данного отношения получается путем изъятия из графа этого отношения таких дуг (x, y), что найдется более длинный путь из вершины x в вершину y. Иначе говоря, $xR^r y$ выполняется тогда и только тогда, когда выполняется xRy, и не существует промежуточного элемента z такого, что xRz и zRy, т. е. $xR^r y$ означает, что x и y непосредственно связаны.

На практике, когда хотят изобразить отношение порядка, довольно часто рисуют граф его редукции, причем не оговаривая это явно. В результате на рисунке мы видим нетранзитивное отношение, которое порядком не является. На рис. 5.31 (слева) показаны два примера редукции отношения порядка.

Граф отношения порядка можно еще более упростить, если воспользоваться методом так называемых *диаграмм Хассе*. Диаграмма Хассе строится следующим образом. Меньшие по порядку элементы располагают ниже, а бóльшие — выше. Поскольку одного такого правила недостаточно для изображения, проводят линии, показывающие, какой из двух элементов больше, а какой меньше другого. При этом достаточно нарисовать лишь линии для непосредственно следующих друг за другом элементов. Примеры диаграмм Хассе показаны на рис. 5.31.

В отличие от графа редукции отношения, который можно как угодно поворачивать в плоскости, в диаграмме Хассе можно не указывать стрелки. Диаграмму Хассе тоже можно поворачивать в плоскости, но не произвольно. При поворотах необходимо сохранять относительное положение (выше—ниже) вершин диаграммы.

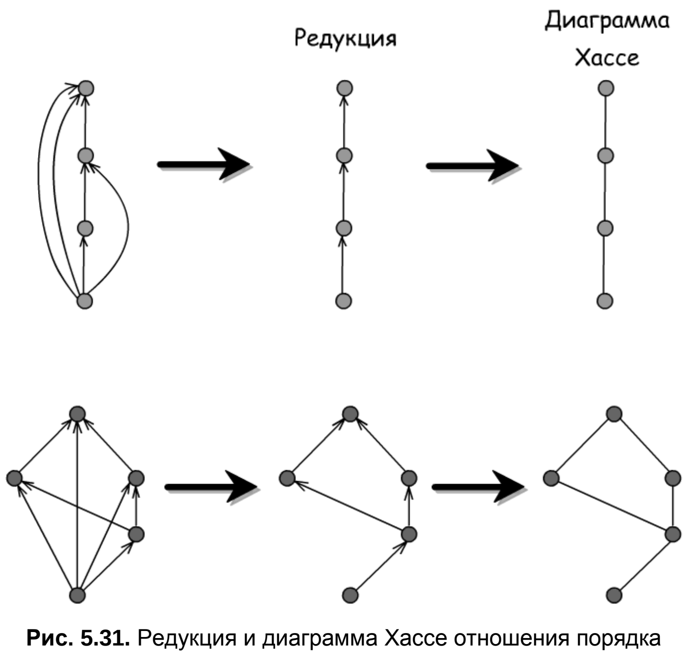


Рис. 5.31. Редукция и диаграмма Хассе отношения порядка

Частичный или нестрогий порядок

ОПРЕДЕЛЕНИЕ 5.17

Квазипорядок, все ядра которого одноэлементны, называют *безъядерным*, или *частичным* порядком.

Эпитет "частичный" выражает тот факт, что, возможно, не все элементы множества сравнимы в данном отношении.

Теорема 5.12. Частичный порядок антисимметричен.

Доказательство. Пусть элементы x, y принадлежат одному ядру. Тогда, по теореме 5.11, $xR \bigcap R^{-1} y$. Так как R — безъядерный квазипорядок, а x, y принадлежат одному ядру, то $x = y$. Таким образом, отношение $R \bigcap R^{-1}$ содержит только пары вида (x, x) и, следовательно, $R \bigcap R^{-1} \subseteq E$, где E — диагональное отношение (отношение тождества). А это и означает, что отношение R антисимметрично.

Итак, частичный порядок — это безъядерный антисимметричный квазипорядок, т. е. рефлексивное, транзитивное и антисимметричное бинарное отношение.

Типичными примерами отношения частичного порядка являются отношения "не больше", "не меньше", "не старше". Частица "не" в названиях отношений служит для выражения их рефлексивности. Отношение "не больше" совпадает с отношением "меньше либо равно", а отношение "не меньше" то же самое, что и "больше либо равно". В связи с этим частичный порядок еще называют *нестрогим* порядком. Часто отношение частичного (нестрогого) порядка обозначают символом "$\leq$".

Частичный порядок можно получить из квазипорядка путем перехода к отношению порядка между его ядрами. Это показано на рис. 5.32 (петли у вершин графов не показаны). Все элементы одного ядра, как известно, связаны между собой исходным отношением, а потому эквивалентны.

Отношение включения $\subseteq$ между подмножествами некоторого множества также является частичным порядком. Очевидно, что не любые два подмножества сравнимы по этому отношению. На рис. 5.33 показан частичный порядок по включению на множестве всех подмножеств множества $\{1,2,3\}$. Стрелки на графе, которые должны быть направлены вверх, не показаны.

Множества, на которых задан частичный порядок, называют *частично упорядоченными*, или просто *упорядоченными* множествами.

Нетрудно заметить, что если на множестве задан некоторый частичный порядок R, то и обратное ему отношение R^{-1} также является частичным порядком.

Так, если $\leq$ — частичный порядок, то $\geq$ — тоже частичный порядок. Для упорядоченных множеств выполняется принцип двойственности: если для множества, упорядоченного по отношению R, выполняется некоторое утверждение $C(R)$, то выполняется и двойственное утверждение, получающееся из C заменой всюду R на R^{-1}.

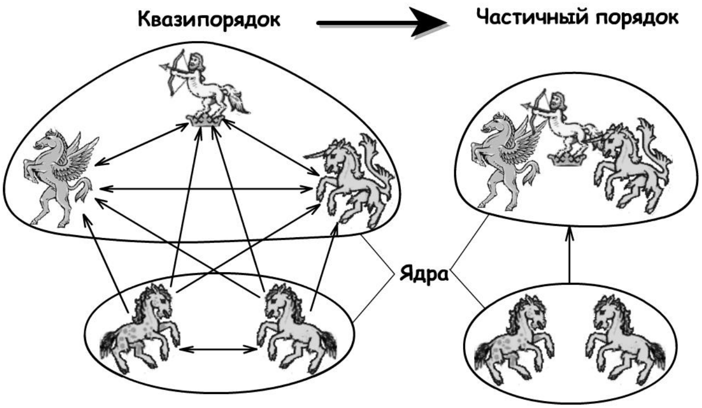


Рис. 5.32. Переход от квазипорядка к частичному порядку на его ядрах

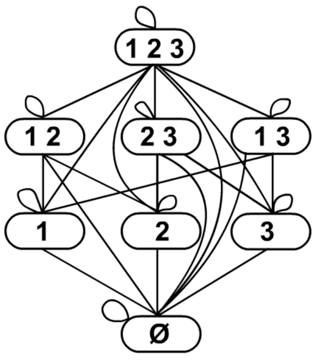


Рис. 5.33. Частичный порядок включения на множестве всех подмножеств множества {1, 2, 3} (стрелки не показаны)

Пусть на множестве A задан частичный порядок $\leq$. Элемент x называется *максимальным* (*минимальным*) элементом множества A, если из того, что $x \leq y$ ($y \leq x$), следует равенство $x = y$. Иначе говоря, элемент x является максимальным (минимальным), если для любого элемента y или неверно, что $x \leq y$ ($y \leq x$), или выполняется $x = y$. Таким образом, максимальный (минимальный) элемент больше (меньше) всех отличных от него элементов, с которыми он находится в отношении $\leq$.

Элемент x называется *наибольшим* (*наименьшим*), если для любого $y \in A$ выполняется $y \leq x$ ($x \leq y$).

В частично упорядоченном множестве может быть несколько минимальных и/или максимальных элементов, но наименьших и наибольших элементов не может быть больше одного. Наименьший (наибольший) элемент является одновременно и минимальным (максимальным), но обратное утверждение неверно. На рис. 5.34 слева показан частичный порядок с двумя минимальными и двумя максимальными элементами, а справа — частичный порядок с наименьшим и наибольшим элементами.

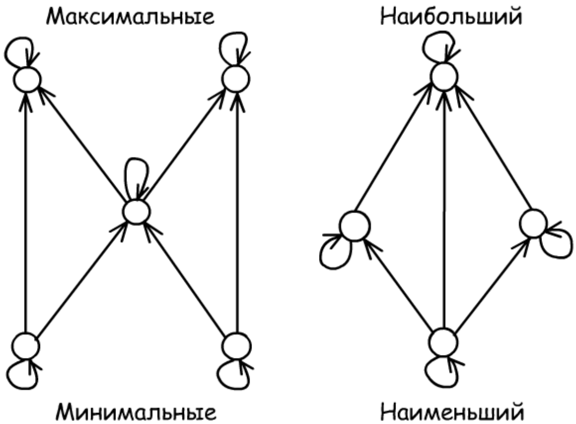


Рис. 5.34. Примеры частичных порядков

Теорема 5.12. В конечном частично упорядоченном множестве всегда существуют минимальный и максимальный элементы.

Для доказательства этой теоремы требуется показать, что для любого элемента y частично упорядоченного множества найдется минимальный (максимальный) элемент x, такой что $x \leq y$ ($y \leq x$). Если для некоторого y нет такого x, то y и есть минимальный (максимальный) элемент. В противном случае аналогичным образом рассматриваем элемент $z \leq y$ ($y \leq z$) и т. д. Эта процедура когда-нибудь завершится в силу конечности множества.

В множестве 2^A всех подмножеств произвольного множества A (не обязательно конечного), упорядоченного по отношению включения $\subseteq$, существуют наибольший и наименьший элементы: само множество A является наибольшим элементом, а пустое множество Ø — наименьшим. Однако в множестве всех непустых подмножеств множества A уже нет наименьшего элемента, зато каждое одноэлементное подмножество является минимальным.

Упорядоченное множество, у которого есть наибольший и наименьший элементы, называется *ограниченным*. На рис. 5.35 показан пример бесконечного ограниченного множества. Разумеется, изобразить бесконечное множество на конечной странице нельзя, но можно показать принцип его построения. Здесь петли около вершин не показаны для упрощения рисунка. По той же причине не показаны дуги, обеспечивающие отображение свойства транзитивности. Другими словами, на рис.5.35 представлена диаграмма Хассе отношения порядка.

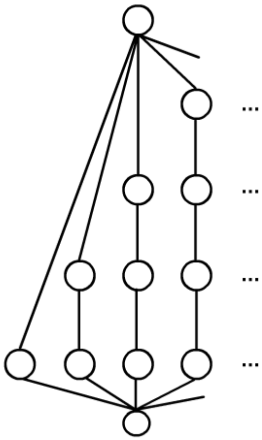


Рис. 5.35. Бесконечное ограниченное частично упорядоченное множество (стрелки и петли не показаны), изображенное в виде диаграммы Хассе

Бесконечные множества могут не иметь максимальных, или минимальных, или тех и других элементов. Например, множество натуральных чисел (1, 2, 3, …) имеет наименьший элемент 1, но не имеет максимальных. Множество всех действительных чисел с естественным порядком не имеет ни наименьшего, ни наибольшего элемента. Однако его подмножество, состоящее из всех чисел $x \leq 5$, имеет наибольший элемент (число 5), но не имеет наименьшего.

Линейный порядок

Элементы x и y частично упорядоченного множества называются *сравнимыми*, если $x \leq y$ или $y \leq x$. В противном случае они не сравнимы.

Определение 5.18

Упорядоченное множество, в котором любые два элемента сравнимы, называется линейно упорядоченным, а порядок — линейным порядком. Линейный порядок еще называют совершенным порядком.

Например, множество всех действительных чисел с естественным порядком $\leq$, а также все его подмножества, линейно упорядочены. Множество 2^A всех подмножеств множества A, упорядоченного по отношению включения $\subseteq$, будет линейно упорядоченным только тогда, когда оно содержит не больше одного элемента.

Линейно упорядоченные подмножества частично упорядоченного множества A называют *цепями* в A.

Строгий порядок

Определение 5.19

Отношение называется *строгим порядком*, если оно антирефлексивно и транзитивно.

Из антирефлексивности и транзитивности логически следует асимметричность, а из асимметричности и транзитивности следует рефлексивность. Попробуйте доказать это самостоятельно в качестве упражнения.

Таким образом, строгий порядок можно определить двумя равносильными (эквивалентными) способами:

- как антирефлексивное и транзитивное отношение;
- как асимметричное и транзитивное отношение.

Обычно строгий порядок обозначают символами "<" и ">", чтобы напоминать об антирефлексивности. Строгий порядок можно получить из частичного (нестрогого) порядка путем удаления из последнего всех пар одинаковых элементов, т. е. пар вида (x, x). Очевидно, что граф строгого порядка получается из графа частичного порядка путем удаления всех петель.

Очевидно, что отношение, обратное строгому порядку, также является строгим порядком. Если $<$ — строгий порядок, то обратное отношение обычно обозначают не как $<^{-1}$, а как $>$.

По редукции строгого порядка всегда можно восстановить исходный строгий порядок, о чем гласит следующая теорема.

Теорема 5.13. Транзитивное замыкание редукции строгого порядка R на конечном множестве совпадает с R, т. е. $\widehat{R^r} = R$.

Доказательство. По определению редукции $R^r \subseteq R$. Следовательно, $\hat{R^r} \subseteq \hat{R}$. Но, как отмечалось в *разд. 5.4.2*, для транзитивного отношения выполняется равенство $R = \hat{R}$, поэтому $\hat{R^r} \subseteq R$.

Обратно, пусть xRy. Если выполняется $xRz_1,..., z_nRy$, то в силу транзитивности и асимметричности отношения R в последовательности $x, z_1,..., z_n, y$ любые два элемента различны.

Рассмотрим все возможные последовательности элементов $z_1,..., z_n$ ($n \geq 0$), такие что выполняется $xRz_1,..., z_nRy$. Поскольку множество, на котором задано отношение R, конечно, то и указанных последовательностей конечное количество. Значит, среди них есть хотя бы одна последовательность максимальной длины. Выберем одну из них. Из того, что последовательность $z_1,..., z_n$ имеет максимальную длину, следует $xR^r z_1,..., z_nR^r y$. В самом деле: если, например, не выполняется $z_i R^r z_{i+1}$, то $z_i R^2 z_{i+1}$, где $R^2 = R \circ R$ (композиция). Это означает, что существует такой элемент u, что $z_i Ru$ и uRz_{i+1}. Но тогда последовательность $z_1,..., z_i, u, z_{i+1},..., z_n$ имеет бóльшую длину, что противоречит нашему выбору. Итак, имеет место $xR^r z_1,..., z_nR^r y$, откуда по определению транзитивного замыкания $x \hat{R^r} y$ и $R \subseteq \hat{R^r}$. В итоге имеем $\hat{R^r} = R$.

Итак, строгий порядок всегда можно обратимо представить его редукцией, а восстановить исходное отношение можно с помощью транзитивного замыкания редукции.

Рассмотрим отношение линейного строгого порядка, т. е. такой строгий порядок <, в котором для любой пары различных элементов x и y имеет место либо $x < y$, либо $y < x$. Последние два соотношения не могут выполняться одновременно в силу асимметричности строгого порядка. Следующие две теоремы о свойствах линейного строгого порядка приведем без доказательства.

Теорема 5.14. Пусть на конечном непустом множестве A задано отношение линейного строгого порядка <. Тогда существует единственный элемент $x \in A$, такой что для любого $y \in A$, отличного от x, выполнено отношение $x < y$. Иначе говоря, в A существует наименьший элемент.

Теорема 5.15. Пусть на конечном непустом множестве A задано отношение линейного строгого порядка <. Тогда элементы множества A можно перенумеровать (составить последовательность элементов $x_1, x_2,..., x_n$), так что соотношение $x_i < x_j$ будет выполняться тогда и только тогда, когда $i < j$.

Рассмотрим пример задания линейного строгого порядка. Пусть на множестве A определена инъективная *(см. разд. 5.2)* функция $f : A \to \Re$, принимающая значения из множества $\mathfrak{R}$ действительных чисел. Зададим отношение < на множестве A следующим образом: $x < y$, если $f(x) < f(y)$. Данное отношение антирефлексивно, поскольку не может быть $f(x) < f(x)$. Транзитивность также очевидна. Для любых различных элементов x и y из множества A выполняется либо $f(x) < f(y)$, либо $f(y) < f(x)$, т. к. функция f является инъекцией. Следовательно, определенный нами порядок < является линейным. Функция f взаимно однозначно отображает множество A на некоторое подмножество множества $\mathfrak{R}$ действительных чисел, так что соотношение $x < y$ для любых элементов из A равносильно неравенству $f(x) < f(y)$. Так, если функция f сопоставляет объекту, например, его вес (длину, объем и т. п.) $f(x)$, мы получаем описанный выше порядок.

Древесный порядок

Объекты самой различной природы могут быть упорядочены иерархически. Вот несколько примеров.

- Части книги упорядочены так, что книга содержит главы, главы содержат разделы, а разделы состоят из подразделов.
- Папки в файловой системе компьютера вложены друг в друга, образуя ветвящуюся структуру.
- Отношение родители—дети может быть изображено в виде так называемого *генеалогического древа*, которое показывает, кто чьим предком (или отпрыском) является.

Иерархический порядок называют древесным (древовидным) потому, что его граф похож на дерево. Древесный порядок является частным случаем строгого порядка.

Определение 5.20

Отношение $R \subseteq A \times A$ называется *древесным порядком*, если

1. R — строгий порядок.
2. Для любых $x, y \in A$ из того, что xRy и xRz, следует yRz или zRy.
3. В множестве A есть наибольший элемент, который называется *корнем дерева*.

На рис. 5.36 слева показан граф древесного порядка, который называют просто *деревом*. На этом же рисунке справа показана *редукция древесного порядка*. Обычно при изображении древесного порядка рисуют именно его редукцию, которая больше напоминает дерево.

Примечание

Напомню, что строгий порядок (а древесный порядок является строгим порядком) взаимно однозначно представляется своей редукцией.

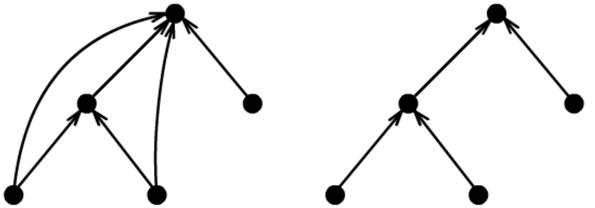


Рис. 5.36. Древесный порядок (слева) и его редукция (справа)

Для различных элементов древесного порядка были введены специальные названия, дающие определенные удобства при обсуждении свойств древесных порядков.

- Для любого элемента $x \in A$ множество x^{Δ} называется множеством *предков* (*предшественников*) элемента x; если $y \in x^{\Delta}$, то y называется *предком* элемента x.
- Для любого элемента $x \in A$ множество x^{∇} называется множеством *потомков* (*последователей*) элемента x; если $y \in x^{\nabla}$, то y — *потомок* x.
- Если $xR^{r}y$, то y называется *родителем* x, а x — *дочерним элементом* для y; здесь R^{r} — редукция древесного порядка. Таким образом, родитель — это непосредственный предок, а дочерний элемент — непосредственный потомок.
- Если $x^{\nabla} =$ Ш, то x называется *концевым элементом*, или *листом* (у листьев нет потомков). Листья являются минимальными элементами древесного порядка.
- Если $x^{\Delta} =$ Ш, то x называется *корнем* дерева (у корня нет предков).

На рис. 5.37 показана редукция древесного порядка с указанием перечисленных выше элементов.

Применительно к древесным порядкам часто используется понятие *яруса*. Нулевой ярус состоит из корня дерева, первый ярус состоит из всех дочерних элементов корня дерева, второй — из всех дочерних элементов вершин первого яруса и т. д. Например, в дереве на рис. 5.37 пять ярусов. Количество дочерних элементов вершины дерева называют *основанием* этой вершины. Например, в дереве на рис. 5.37 основание вершины x равно 3; листья имеют основание 0. Деревья, все вершины которых, за исключением листьев, имеют основание 2, называют *бинарными*.

Важное свойство древесного порядка дает следующая теорема.

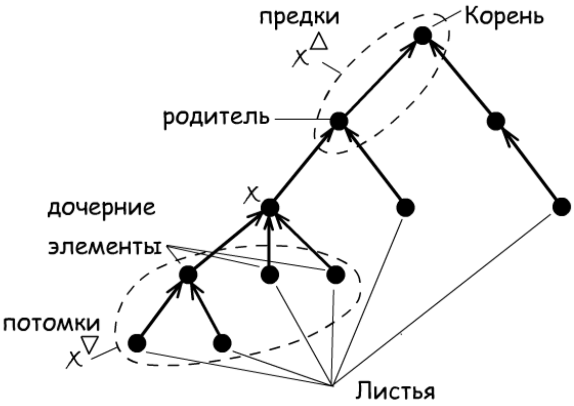


Рис. 5.37. Как называются элементы древесного порядка

Теорема 5.16. В конечном множестве с древесным порядком каждый элемент, за исключением корня, имеет только одного родителя.

Доказательство. Пусть A — конечное множество и $R \subseteq A \times A$ — древесный порядок. Предположим, что утверждение теоремы неверно. Тогда найдутся такие элементы $x, y, z \in A$, что $xR^r y$ и $xR^r z$ и $y \neq z$. В силу свойства 2 определения 5.20 должно выполняться или yRz, или zRy. Но тогда по определению редукции или неверно, что $xR^r z$, или неверно, что $xR^r y$, что противоречит определению родителя.

Итак, мы доказали, что у любого элемента не может быть больше одного родителя. Покажем теперь, что у любого элемента, за исключением корня, есть родитель. В самом деле, если x не имеет родителя (т. е. для всех y не выполняется соотношение $xR^r y$), то для всех y не выполняется соотношение $x\hat{R^r}\, y$, а значит и xRy, поскольку для строгого порядка по теореме 5.13 имеет место равенство $\hat{R^r} = R$ (напомню, что $\hat{R^r}$ — транзитивное замыкание редукции R^r отношения R). Однако корень дерева, обозначим его через x_0, — наибольший элемент, и, следовательно, выполняется соотношение xRx_0. Таким образом, если $x \neq x_0$, то x имеет родителя.

На рис. 5.38 показано отношение строгого порядка (точнее, его редукция), которое не является древесным порядком, хотя и очень похоже на него. Дело в том, что некоторые элементы в этом порядке имеют больше одного родителя.

Нетрудно заметить, что подмножество, состоящее из некоторого элемента и всех его потомков, само упорядочено древесным порядком. Дерево, соответствующее этому порядку, называют *поддеревом* исходного дерева.

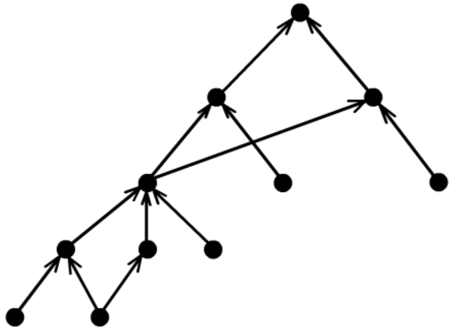


Рис. 5.38. Строгий порядок, не являющийся древесным

ПРИМЕЧАНИЕ

Обратите внимание: если в определении 5.20 убрать третье условие о существовании корня, то определяемый порядок превратится в *лес* — совокупность не связанных между собой деревьев.

Очевидно, что отношение, обратное древесному порядку, также является древесным порядком. Однако в графе такого отношения стрелки направлены не к корню дерева, а, наоборот, к его листьям. При этом корень дерева оказывается не наибольшим, а наименьшим элементом, а листья — не минимальными, а максимальными элементами. Графы таких древесных порядков часто изображают так, чтобы корень находился внизу, а листья — наверху. На диаграммах Хассе, соответствующих древесному порядку, стрелки вообще не изображают.

5.5. Отображения отношений

В *разд. 5.2* мы рассматривали соответствия между двумя множествами. Бинарные отношения также можно рассматривать как соответствия между элементами двух множеств. Однако на множествах, между которыми есть некоторое соответствие, могут быть заданы какие-то бинарные отношения, и тогда возникает задача сопоставить или сравнить их друг с другом. Если мы хотим понять, что такое аналогия в математическом смысле этого слова, то нам следует разобраться, прежде всего, с отображениями множеств, на каждом из которых задано некоторое отношение. В данном разделе понятие гомоморфизма как раз и есть математическое уточнение (экспликация) понятия аналогии. Мы начнем именно с этого понятия, которое затем разовьем путем приобщения к нему дополнительных свойств, получив в результате то, что называется корреспонденцией, изоморфизмом и т. д.

Итак, множества можно сопоставлять друг другу. Это означает, что элементам одного множества сопоставляются элементы другого множества. Символически соответствие, обозначим его через f, между множествами A и B записывается так: $f: A \to B$. Возьмем некоторый элемент $x \in A$, тогда $f(x)$ является некоторым элементом множества B, который соответствует элементу x по данному соответствию f. Элемент $f(x)$, как мы уже знаем, называется образом элемента x. Разумеется, так можно говорить, только если для данного x соответствие f определено. Теперь мы предполагаем, что на множествах A и B определены некоторые отношения и при всем этом задано соответствие $f: A \to B$.

Пусть на множестве A задано некоторое бинарное отношение R, а на множестве B — бинарное отношение Q. Предположим еще, что между множествами A и B имеется некоторое соответствие $f: A \to B$. Сопоставить отношения R и Q — значит, указать, какую информацию о выполнении отношения Q несет тот факт, что между некоторыми элементами множества A выполняется отношение R.

Таким образом, нас интересует не просто соответствие $f: A \to B$ между множествами, а соответствие $f: <A, R> \to <B, Q>$ между множествами с отношениями. Последняя запись означает, что между множествами A и B имеется соответствие f, а R и Q — бинарные отношения, заданные на этих множествах. Если для элемента $x \in A$ соответствие определено, то $f(x)$ — образ этого элемента в множестве B, т. е. $f(x) \in B$. Пусть для элементов $x, y \in A$ соответствие f определено. Тогда сопоставление отношений R и Q означает выяснение того, что можно сказать о выполнении соотношения $f(x)Qf(y)$, если известно, что выполняется xRy.

На рис. 5.39 показано соответствие между двумя бинарными отношениями. Отношения отображаются стрелками, а соответствие между множествами — штриховыми линиями. Штриховые линии связывают между собой прообразы (слева) и образы (справа) соответствия. Хотя данные отношения заданы на различных множествах, между ними легко обнаружить некое подобие. В данном случае отношение между геометрическими фигурами можно интерпретировать как схему сборки человечка, а отношение между частями человечка — как ее конкретную реализацию. Нетрудно заметить, что в данном примере не все реализованное было предусмотрено схемой, но все, что есть в схеме, было реализовано. В частности, в схеме сборки не была предусмотрена шапка.

Соответствия между множествами A и B могут обладать или не обладать некоторыми "хорошими" свойствами. Одним из таких хороших свойств является всюду определенность соответствия, т. е. существование образа $f(x)$

для каждого элемента $x \in A$. Если при этом образ $f(x)$ является одноэлементным множеством, то соответствие f называют *отображением* (*однозначным соответствием*, или *функцией*). Например, соответствие на рис. 5.39 является отображением.

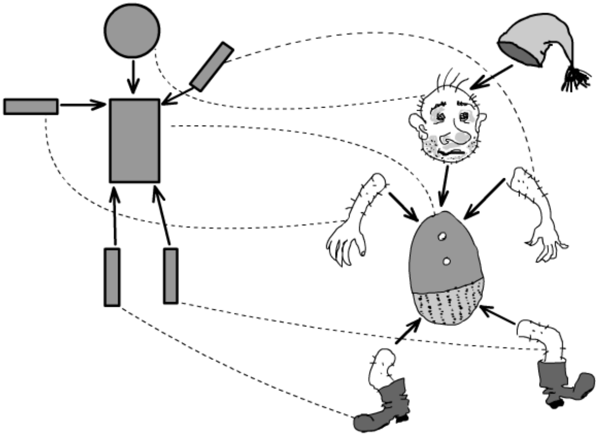


Рис. 5.39. Пример соответствия между двумя бинарными отношениями

Далее, отображения могут быть взаимно однозначными (биекциями, биективными), инъективными и сюръективными. Определения таких отображений даны в *разд. 5.2*. При наличии между множествами A и B какого-либо из перечисленных отображений можно говорить о подобии отношений в большей или меньшей степени.

5.5.1. Гомоморфизм отношений

Определение 5.21

Отображение $f : A \to B$ называется *гомоморфным* (от греч. homos — подобный, morphe — форма, вид) *отображением* (или просто *гомоморфизмом*) бинарного отношения $< A, R >$ в бинарное отношение $< B, Q >$, если для всех $x, y \in A$ из xRy следует $f(x)Qf(y)$.

Иначе говоря, при гомоморфизме, если два прообраза находятся в отношении R, то их образы должны находиться в отношении Q.

На рис. 5.40 приведены два примера гомоморфных отображений. Обратите внимание: во-первых, множество B может иметь больше элементов, чем множество A, причем элементы из B, не являющиеся образами никаких элементов из A, могут быть связаны с другими элементами из B произвольным образом; во-вторых, даже если все элементы из B имеют прообразы, они могут иметь и дополнительные связи друг с другом. При всем при этом отображение f остается гомоморфным.

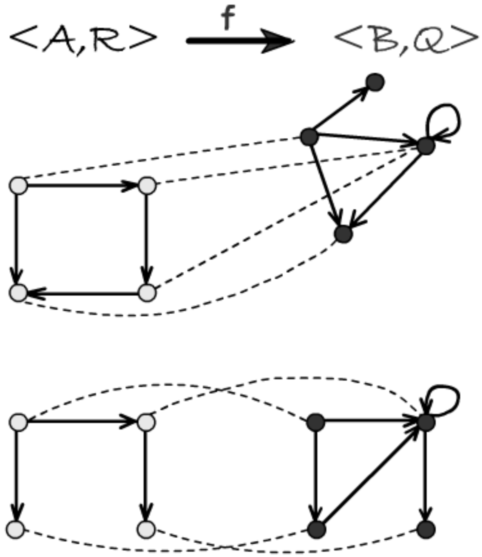


Рис. 5.40. Примеры гомоморфизмов отношений

Отображение, показанное на рис. 5.39, также является гомоморфным. Легко проверить, что для любой пары связанных геометрических фигур найдется связанная пара частей тела человечка.

Может показаться, что при гомоморфизме отношение, в которое происходит отображение, сложнее (богаче) исходного отношения. Однако это не так. На рис. 5.41 показан гомоморфизм, при котором правое отношение проще левого. В данном случае правое отношение можно интерпретировать как схему сборки памятника (в виде бюста) человечку, отношение между частями тела которого показано слева.

Рассмотрим еще несколько примеров.

- Пусть на множестве R заданы строгий порядок и взаимно однозначное отображение $f: A \to A$, сопоставляющее каждому элементу x тот же самый элемент. Тогда отображение $f :< A, R^r > \to < A, R >$ редукции R^r отношения R является гомоморфизмом.

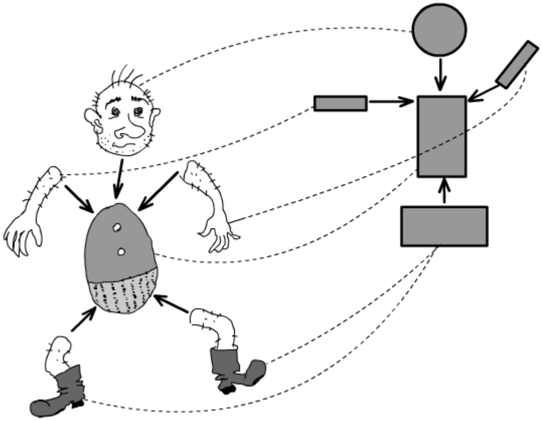


Рис. 5.41. Гомоморфное отображение схемы сборки человечка в схему сборки его "бюста". Это же отображение является и корреспонденцией

- Пусть R — произвольное бинарное отношение на множестве A, а $f : A \to A$ — взаимно однозначное отображение, сопоставляющее каждому элементу x тот же самый элемент. Тогда отображение

  $f :< A, \hat{R} >\to< A, R >$ транзитивного замыкания отношения R в исходное отношение R является гомоморфизмом.
- Пусть ε — отношение эквивалентности на множестве A, а K_ε — множество всех классов этой эквивалентности. Определим на множестве K_ε тривиальное отношение тождества Q: для всех $k, k' \in K_\varepsilon$ соотношение kQk' выполняется тогда и только тогда, когда $k = k'$. Иначе говоря, единственное свойство отношения Q — рефлексивность; граф такого отношения представляется в виде изолированных вершин с петлями. Далее, определим взаимно однозначное соответствие $f : A \to K_\varepsilon$, сопоставляющее каждому элементу множества A единственный класс эквивалентности, которому он принадлежит. Тогда отображение $f :< A, \varepsilon >\to< K_\varepsilon, Q >$ является гомоморфизмом.

Примечание

Можно доказать, что отображение отношений $f :< A, R >\to< B, Q >$ является гомоморфизмом тогда и только тогда, когда для любого класса (ко-класса) k_R отношения R найдется класс (ко-класс) k_Q отношения Q, такой что $f(k_R) \subseteq k_Q$ и $f(k_R{}^\Delta) \subseteq k_Q{}^\Delta$ ($f(k_R{}^\nabla) \subseteq k_Q{}^\nabla$).

Теорема 5.17. Пусть $e : A \to A$ — тождественное отображение множества A (т. е. для любого $x \in A$ имеет место $e(x) = x$, а R и Q — некоторые отношения на A). Тогда отображение $e :< A, R > \to < A, Q >$ является гомоморфизмом в том и только в том случае, когда $R \subseteq Q$.

Иллюстрация данного утверждения показана на рис. 5.40 (внизу), если считать, что левое и правое множества это одно и то же множество.

5.5.2. Корреспонденция отношений

Корреспонденция отношений в некотором смысле противоположна гомоморфизму. При корреспонденции, если два образа находятся в отношении Q, то любые их прообразы находятся в отношении R. Термин "корреспонденция" для таких отображений ввел С. Шаумян.

Определение 5.22

Отображение $f : A \to B$ называется *корреспонденцией* бинарного отношения $< A, R >$ в бинарное отношение $< B, Q >$, если для всех $x, y \in B$ из xQy следует $f^{-1}(x)Rf^{-1}(y)$. Здесь f^{-1} — отображение, обратное отображению f, а $f^{-1}(x)Rf^{-1}(y)$ означает, что для любой пары элементов $x' \in f^{-1}(x), y' \in f^{-1}(y)$ выполняется $x'Ry'$.

Дело в том, что прообразы $f^{-1}(x), f^{-1}(y)$ не обязательно одноэлементны. Если каждый элемент множества B имеет не больше одного прообраза, то в этом случае определение корреспонденции можно упростить: отображение $f : A \to B$ является корреспонденцией, если для всех $x, y \in B$ из $f(x)Qf(y)$ следует xRy.

На рис. 5.42 показан пример корреспонденции отношений.

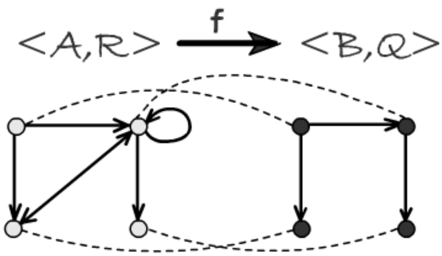


Рис. 5.42. Пример корреспонденции отношений

Следующая теорема аналогична теореме 5.17.

Теорема 5.18. Пусть $e : A \to A$ — тождественное отображение множества A (т. е. для любого $x \in A$ имеет место $e(x) = x$, а R и Q — некоторые отношения на A). Тогда отображение $e :< A, R > \to < A, Q >$ является корреспонденцией в том и только в том случае, когда $R \subseteq Q$.

Иллюстрация данного утверждения показана на рис. 5.42, если считать, что левое и правое множества это одно и то же множество.

Если отображение $f : A \to B$ взаимно однозначно (биективно), то f тогда и только тогда является корреспонденцией, когда обратное отображение $f^{-1} : B \to A$ является гомоморфизмом. Если обратное отображение не существует, то понятие корреспонденции не сводится к понятию гомоморфизма.

Отображение, показанное на рис. 5.41, является не только гомоморфизмом, но и корреспонденцией.

Определение 5.23

Отображение, являющееся одновременно гомоморфизмом и корреспонденцией, называется *k-гомоморфизмом*.

Оказывается, что любой гомоморфизм может быть расширен до k-гомоморфизма, о чем и говорит следующая теорема.

Теорема 5.19. Пусть $f : A \to B$ — отображение и $< B, Q >$ — некоторое отношение на множестве B. Тогда

1. Существует единственное отношение $< A, S >$, такое что $f :< A, S > \to < B, Q >$ является k-гомоморфизмом.
2. Для любого отношения $< A, V >$, для которого f является гомоморфизмом отношения $< A, V >$ в $< B, Q >$, выполняется включение $V \subseteq S$.

Доказательство.

1. Определим отношение $< A, S >$ условием:

   xSy, если $f(x)Qf(y)$. (*)

   Очевидно, что отображение $f :< A, S > \to < B, Q >$ является k-гомоморфизмом. Теперь докажем единственность такого отношения S. Предположим, что f является также k-гомоморфизмом некоторого отношения $< A, S' >$ в отношение $< B, Q >$. Допустим сначала, что $xS'y$. Так как f — гомомомоморфизм отношения $< A, S' >$ в отношение $< B, Q >$, то $f(x)Qf(y)$. Из условия (*) следует xSy. Таким образом, $S' \subseteq S$. Обратно, пусть xSy. По условию (*) определения отношения S должно быть $f(x)Qf(y)$. Поскольку f является корреспонденцией отношения $< A, S' >$ в отношение $< B, Q >$, то выполняется $xS'y$. Следовательно, $S \subseteq S'$. Таким образом, $S' = S$.

2. Включение $V \subseteq S$ доказывается так же, как мы доказывали включение $S' \subseteq S$ в первой части доказательства.

Примечание

Можно доказать, что отображение отношений $f :< A, R > \to < B, Q >$ является корреспонденцией тогда и только тогда, когда для любого класса (ко-класса) k_Q отношения Q найдется класс (ко-класс) k_R отношения R, такой что $f^{-1}(k_Q) \subseteq k_R$ и $f^{-1}(k_Q{}^{\Delta}) \subseteq k_R{}^{\Delta}$ ($f^{-1}(k_Q{}^{\nabla}) \subseteq k_R{}^{\nabla}$).

5.5.3. Изоморфизм и другие морфизмы отношений

В зависимости от свойств отображения $f : A \to B$ множества A в множество B гомоморфизм $f :< A, R > \to < B, Q >$ отношения $< A, R >$ в отношение $< B, Q >$ получил специальные названия.

Определение 5.24

Если отображение $f : A \to B$ взаимно однозначно (биективно), то гомоморфизм $f :< A, R > \to < B, Q >$ называется *изоморфизмом*. Если f — сюръекция, то гомоморфизм f называется *эпиморфизмом*. Если f — инъекция, то гомоморфизм f называется *мономорфизмом*. Изоморфизм, эпиморфизм, мономорфизм f называются соответственно *k*-изоморфизмом, *k*--эпиморфизмом, *k*-мономорфизмом, если f одновременно является корреспонденцией.

На рис. 5.39 приведен мономорфизм, а если бы в множестве частей человечка шапка отсутствовала, то тогда это был бы изоморфизм и даже *k*-изоморфизм. На рис. 5.40 внизу показан изоморфизм, а на рис. 5.41 — эпиморфизм и даже *k*-эпиморфизм.

Теперь рассмотрим еще раз связь между классами и ко-классами бинарного отношения *(см. разд. 5.3.4)*. Примеры этой связи показаны на рис. 5.7 и 5.8. Итак, пусть дано некоторое бинарное отношение $R \subseteq A \times B$. Обозначим через K и K' множества соответственно классов и ко-классов данного отношения. Как известно, между этими множествами, существует взаимно однозначное отображение $\Delta : K \to K'$, которое определяется через отношение R (точнее, через сечения 2-го рода этого отношения). Множества классов и ко-классов можно упорядочить по отношению включения $\subseteq$. Таким образом, имеются два отношения: $< K, \subseteq >$ и $< K', \subseteq >$. Отображение $\Delta :< K, \subseteq > \to < K', \subseteq >$ этих отношений является *k*-изоморфизмом потому, что:

1. $\Delta : K \to K'$ — биекция.
2. Для любых классов $x, y \in K$ из включения $x \subseteq y$ следует включение ко-классов $\Delta(y) \subseteq \Delta(x)$ (в обозначениях данной книги — $y^{\Delta} \subseteq x^{\Delta}$);

таким образом, Δ является гомоморфизмом, а с учетом п. 1 — изоморфизмом.

3. Для любых ко-классов $x', y' \in K'$ из включения $x' \subseteq y'$ следует включение классов $\Delta^{-1}(y') \subseteq \Delta^{-1}(x')$ (в обозначениях данной книги — $y'^{\nabla} \subseteq x'^{\nabla}$); таким образом, Δ является корреспонденцией, а с учетом п. 2 — *k*-изоморфизмом.

Множества классов и ко-классов являются частично упорядоченными по отношению включения множествами, в которых есть наибольший и наименьший элементы, причем пересечение любых классов (ко-классов) является классом (ко-классом). Такие частично упорядоченные множества называются *полными решетками* (*структурами*). В теории структур отображения, которые здесь назывались *k*-изоморфизмами, называются просто изоморфизмами. Таким образом, отображение Δ между частично упорядоченными по включению множествами классов и ко-классов называют просто *изоморфизмом*. Однако если между классами выполняется включение $x \subseteq y$, то между их образами (ко-классами) выполняется противоположное включение $\Delta(y) \subseteq \Delta(x)$, и, наоборот, если между ко-классами выполняется $x' \subseteq y'$, то между их прообразами выполняется противоположное включение $\Delta^{-1}(y') \subseteq \Delta^{-1}(x')$. Чтобы отметить данную противоположность включений, изоморфизм Δ называют *антиизоморфизмом*, или *дуальным изоморфизмом*.

Глава 6

От *n*-арных отношений к реляционным базам данных

Конечная цель, которую всегда нужно иметь в виду, состоит в том, чтобы достичь правильной точки зрения на основания.

К. Вейерштрасс

В *главе 5* мы рассматривали бинарные (2-арные) отношения между элементами двух множеств, изображая их для наглядности в виде графов. Тернарные (3-арные) отношения между тремя множествами также можно представлять графами, вершины которых помечены элементами из любых двух множеств, а дуги — элементами из оставшегося третьего множества. При $n > 3$ графическое представление отношения становится весьма проблематичным. Во всяком случае, такое изображение теряет главное свое достоинство — быть наглядным. Если же еще и количество элементов в множествах велико, то при любом n, даже при $1 < n \leq 3$, графическая форма представления отношений становится громоздкой и непригодной для практических целей их анализа. Что же делать? Остается универсальная табличная форма представления отношений, которая достаточно ясна и компактна даже при больших n. В прямоугольной таблице каждому столбцу соответствует множество элементов одного из множеств, а каждой строке — сочетание или кортеж тех элементов из различных множеств, которые находятся в данном отношении.

Итак, *n*-арное отношение можно представить в виде таблицы. С другой стороны, табличное представление данных используется на практике сплошь и рядом, поскольку это удобно и наглядно. Наконец, большие объемы данных хранятся на компьютерах также в табличной форме. Одна или несколько взаимосвязанных или совсем не связанных таблиц образуют то, что называют *реляционной базой данных* (англ. relation — отношение, связь). Проектирование реляционной базы данных — это проектирование таблиц и связей между ними. Если таблицы устроены так, что представляют некоторые отношения, то проектирование базы данных в целом, а также правила манипулирования данными (извлечение нужных данных, добавление новых, удаление устаревших и изменение уже имеющихся данных) можно водрузить на прочный математический фундамент и далее в конструкторско-инженерной деятельности руководствоваться стройной теорией.

Таблицы, из которых состоит любая реляционная база данных, не просто вместилища данных, а формы, представляющие некоторые отношения, но отношения являются не чем иным, как множествами. Все запросы к базе данных, направленные на извлечение из нее нужных записей, интерпретируются как инструкции по выполнению тех или иных операций, являющихся, в конечном счете, операциями алгебры множеств и исчисления предикатов.

Чтобы понять суть задачи создания реляционной базы данных, а также операций над данными, достаточно рассмотреть на теоретическом (абстрактном) уровне всего лишь несколько основных положений теории отношений, чем мы и займемся в данной главе. Это быстрее и, в конечном счете, лучше, как мне кажется, нежели разбирать большое количество конкретных примеров и ситуаций.

6.1. Отношения и таблицы

Что такое отношение, мы достаточно подробно рассмотрели в *главах 4* и *5*. Там мы отметили, что отношения, по существу, это множества. Вместе с тем, их можно представлять в виде графов, матриц и таблиц. Выбор той или иной формы представления обусловлен, в первую очередь, удобством. Удобством для чего? В данной книге мы выбирали ту форму, которая наиболее наглядна при изучении сути отношений, а именно — графовую. Это было действительно так, пока мы изучали относительно простые (маломерные) бинарные отношения. Однако выбор формы обусловлен не только легкостью зрительного восприятия, но и возможностью применения математического аппарата, специально разработанного для исследования данной формы. Так, используя графовую форму, мы получаем возможность применения многочисленных средств из арсенала глубоко разработанной теории графов. Положив в основу матричную форму представления отношений, мы получаем возможность применения богатейшего материала алгебры. Аналогично, прибегая к табличной форме представления n-арных отношений, мы должны учитывать ее специфику.

Напомню вкратце, что такое n-арное отношение. Пусть имеется n множеств (не обязательно различных) $A_1, A_2, ..., A_n$. Тогда любое подмножество R декартова произведения $A_1 \times A_2 \times ... \times A_n$ есть некоторое отношение, заданное на множествах $A_1, A_2, ..., A_n$. Элементами отношения R являются последовательности (кортежи) вида $(x_1, x_2, ..., x_n)$, такие что $x_1 \in A_1, x_2 \in A_2, ..., x_n \in A_n$. Кортеж это не множество, а упорядоченная последовательность элементов. Если кортеж $(x_1, x_2, ..., x_n) \in R$, то говорят, что элементы $x_1, x_2, ..., x_n$ находятся в отношении R. Если $n = 2$ (бинарное отношение), то кроме выражения вида $(x_1, x_2) \in R$ мы можем использовать и такое: $x_1 R x_2$. Но при $n > 2$ такую форму записи составить нельзя. Поскольку отношение R есть множество, в нем не может быть двух и более одинаковых кортежей. Количество элементов n в кортеже называется его *длиной*. Два кортежа $(x_1, x_2, ..., x_n)$ и $(x'_1, x'_2, ..., x'_m)$ одинаковы тогда и только тогда, когда

1. Они имеют равные длины ($n = m$),
2. $x_1 = x'_1, x_2 = x'_2, ..., x_n = x'_n$.

Понятно, что в отношении все кортежи имеют одинаковую длину.

Не нужно быть особенно прозорливым, чтобы понять, что определенное выше *n*-арное отношение можно представить в виде прямоугольной таблицы с n столбцами, соответствующими множествам $A_1, A_2, ..., A_n$.

Стоит только добавить в эту таблицу новые кортежи или, наоборот, удалить некоторые из имеющихся, или изменить значение хотя бы одной ячейки таблицы (элемент кортежа), как отношение, представляемое данной таблицей, тут же изменится. Таким образом, если мы хотим, чтобы таблица представляла некоторое фиксированное отношение, мы должны потребовать, чтобы ее содержимое было неизменным. Однако с практической точки зрения данное ограничение чересчур обременительно, поскольку в жизни мы имеем дело, как правило, с не до конца заполненными таблицами.

Теперь рассмотрим таблицы. Очевидно, что таблицы создают для того, чтобы хранить и представлять в них данные. Сначала таблица пуста. В процессе ввода данных она представляет то одно, то другое отношение. О каком-то фиксированном отношении, представляемом таблицей, можно будет говорить только тогда, когда таблица будет окончательно заполнена данными. Заметим, что если в таблице окажутся хотя бы две одинаковые строки, то набор ее строк уже не будет образовывать множество, а значит, и представлять какое бы то ни было отношение.

Итак, получается, что отношения могут представляться таблицами, а таблицы могут представлять отношения. Правда, для полного соответствия между отношениями и таблицами требуется кое-что согласовать. Требование, чтобы в таблице не было одинаковых строк, кажется вполне естественным: на практике мы должны избегать дублирования строк в таблицах. Остается требование, согласно которому таблица даже в стадии наполнения данными должна представлять одно и то же, первоначально заданное отношение. Для выполнения этого требования необходимо задать правила заполнения таблицы, которые бы определяли, какие данные допустимы, а какие — нет. В теории реляционных баз данных такие правила называются *ограничениями целостности отношения*. Обычно достаточно четко оговаривается, данные какого типа можно вводить в тот или иной столбец таблицы. Например, в столбец для записи чисел нельзя вводить буквы. Кроме того, данные из одной и той же строки, но из различных столбцов могут быть как-то взаимно обусловлены. Например, если в одном столбце записываются доходы, а в другом — расходы, то значение столбца для записи баланса не может быть каким угодно, оно равно разности значений первых двух столбцов. Все эти правила заполнения таблицы, даже если она пока пуста, задают ограничения на данные, которые в ней могут находиться, а значит, определяют некоторое отношение, которое эта таблица будет представлять в результате окончательного заполнения.

С другой стороны, отношение можно задать не обязательно в явном виде как множество всех его кортежей, а посредством некоторого *n*-местного предиката $R(x_1, x_2,..., x_n)$, переменные которого принимают значения из множеств $A_1, A_2,..., A_n$ соответственно. Данный предикат принимает значение "Истина" на тех и только тех кортежах $(x_1, x_2,..., x_n) \in A_1 \times A_2 \times ... \times A_n$, которые принадлежат отношению R, т. е. когда $(x_1, x_2,..., x_n) \in R$. В противном случае предикат $R(x_1, x_2,..., x_n)$ принимает значение "Ложь". Здесь мы обозначили одной буквой R и предикат, и множество кортежей, определяемое данным предикатом. Предикат $R(x_1, x_2,..., x_n)$ может представлять собой довольно сложное логическое предложение, но как бы оно ни было сформулировано, оно задает ограничение целостности отношения.

Итак, отношение есть множество, элементами которого являются кортежи, которые, в свою очередь, состоят из элементов других множеств. При этом между любым кортежем отношения и самим отношением имеет место отношение принадлежности $\in$, но не включения $\subseteq$. Кортеж представляет собой не множество, а упорядоченную последовательность элементов. Множество всех кортежей, входящих в данное отношение, можно задать с помощью предиката, представляющего собой набор правил, которым должны удовлетворять данные, помещаемые в таблицу.

Таким образом, мы согласовали в общих чертах понятия отношения и таблицы на основе понятия ограничения целостности. Поскольку далее мы будем представлять отношения в виде таблиц, введем специальные обозначения и термины, которые обычно применяются в теории отношений (реляционной теории). Так, говоря о некотором отношении R между элементами множеств $A_1, A_2, ..., A_n$, мы будем использовать запись вида $R(A_1, A_2, ..., A_n)$, называя ее *структурой отношения*. Таким образом, в структуре отношения указываются имя R отношения и перечень множеств $A_1, A_2, ..., A_n$, между элементами которых данное отношение определено. Далее мы будем говорить кратко: отношение R определено на атрибутах $A_1, A_2, ..., A_n$ или между этими атрибутами.

Итак, говоря об отношении $R(A_1, A_2, ..., A_n)$, мы имеем в виду следующее.

- Отношение имеет имя *R*. Например, сведения о товарах, хранящихся на складе, образуют некоторое отношение, которому можно дать имя `СКЛАД`.
- Множества $A_1, A_2, ..., A_n$, для которых определено отношение *R*, имеют различные имена, называемые *атрибутами* отношения. Мы будем считать, что $A_1, A_2, ..., A_n$ — атрибуты отношения. Совокупность всевозможных значений любого множества с именем A_i ($i = 1, 2, ..., n$) называется *доменом* атрибута A_i. Элемент домена атрибута A_i будем называть *значением* атрибута A_i.

Примечание

В отношении не может быть двух и более одинаковых атрибутов, хотя их домены могут быть и одинаковы (в смысле равенства множеств). Например, отношение `СКЛАД` может быть определено для атрибутов `НАИМЕНОВАНИЕ`, `КОЛИЧЕСТВО`, `ЦЕНА` и `ПОСТАВЩИК`. Структуру этого отношения можно записать так: `СКЛАД` (`НАИМЕНОВАНИЕ`, `КОЛИЧЕСТВО`, `ЦЕНА`, `ПОСТАВЩИК`). Каждый атрибут имеет свой домен — множество возможных значений. Например, доменом атрибута `КОЛИЧЕСТВО` может быть множество неотрицательных чисел. Доменом атрибута `ЦЕНА` также может быть множество неотрицательных чисел, но в общем случае домены `КОЛИЧЕСТВО` и `ЦЕНА` могут не совпадать как множества. При табличном представлении отношения каждому атрибуту соответствует заголовок столбца, а домену — множество допустимых значений в этом столбце.

- Порядок перечисления атрибутов в структуре отношения не имеет значения. Иначе говоря, перестановка атрибутов местами оставляет само отношение прежним, хотя вид таблицы, представляющей это отношение, изменяется. Например, отношения `СКЛАД` (`НАИМЕНОВАНИЕ`, `КОЛИЧЕСТВО`, `ЦЕНА`, `ПОСТАВЩИК`) и `СКЛАД` (`ПОСТАВЩИК`, `НАИМЕНОВАНИЕ`, `КОЛИЧЕСТВО`, `ЦЕНА`) считаются эквивалентными.
- Элементами отношения являются кортежи — последовательности значений атрибутов отношения. В отношении не может быть двух и более одинаковых кортежей, а порядок расположения кортежей не имеет значения. Это все — обычные свойства множества как неупорядоченного собрания различающихся объектов. При табличном представлении отношения каж-

дому кортежу взаимно однозначно соответствует строка таблицы. Строки таблицы мы будем называть *записями*.

Внимание

В данной главе мы обозначаем через $R(A)$ отношение, заданное на атрибуте (или множестве атрибутов) A, а не сечение этого отношения через множество A, как это делалось в *главе 5*.

6.2. Операции над отношениями

Коль скоро отношение — это множество (а именно множество кортежей или, другими словами, записей), то к нему применимы все теоретико-множественные операции, рассмотренные ранее. Однако в реляционной теории особую роль играют специальные операции, заложенные в основу стандартного языка запросов к базам данных SQL (Structured Query Language — язык структурированных запросов):

- селекция;
- проекция;
- естественное соединение.

Эти операции выражаются некоторым образом через обычные операции над множествами, такие как объединение, пересечение, вычитание и декартово произведение. Рассмотрим их более подробно.

6.2.1. Селекция

Операция *селекции* (выборки, ограничения) отношения выделяет из него некоторое подмножество кортежей, удовлетворяющих некоторому условию. При работе с базами данных выборка из всей имеющейся совокупности только требуемых данных — наиболее часто применяемая операция. Конкретное определение этой операции зависит от вида условия. Одной из частных разновидностей операции селекции является операция сужения отношения. В *разд. 5.4.1* мы уже рассматривали эту операцию для бинарного отношения, а теперь дадим ее определение для общего случая n-арного отношения.

Пусть имеется некоторое отношение $R(A1, A2, \ldots, An)$. Обозначим домены атрибутов $A1, A2, \ldots, An$ этого отношения через $D1, D2, \ldots, Dn$ соответственно.

Примечание

Напомню, что домен атрибута — это множество его допустимых значений.

Далее, обозначим через $d1, d2, \ldots, dn$ какие-нибудь подмножества доменов $D1, D2, \ldots, Dn$ соответственно (т. е. $di \subset Di$, $i = 1, 2, \ldots, n$). Тогда *сужением* отношения $R(A1, A2, \ldots, An)$ на множество $d1 \times d2 \times \ldots \times dn$ называется отношение $R(A1, A2, \ldots, An) \bigcap d1 \times d2 \times \ldots \times dn$. По существу, данная операция просто выделяет из исходного отношения только те кортежи, в которых элементы (значения атрибутов) принадлежат указанным подмножествам соответствующих доменов.

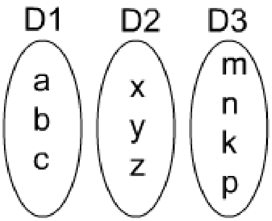
Примечание

В частном случае, когда $di = Di$ для всех $i = 1, 2, \ldots, n$, сужение отношения $R(A1, A2, \ldots, An)$ на множество $d1 \times d2 \times \ldots \times dn$ равно исходному отношению $R(A1, A2, \ldots, An)$.

На рис. 6.1 в качестве примера показаны некоторое отношение $R(A1, A2, A3)$, домены $D1$, $D2$ и $D3$ атрибутов $A1$, $A2$ и $A3$, некоторые подмножества $d1$, $d2$ и $d3$ этих доменов, декартово произведение $d1 \times d2 \times d3$ и результат сужения исходного отношения на множество кортежей $d1 \times d2 \times d3$.

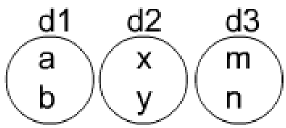
R(A1,A2,A3) : таблица

A1	A2	A3
a	x	m
a	y	n
b	z	k
c	x	n
c	x	p



d1xd2xd3 : таблица

A1	A2	A3
a	x	m
a	y	m
b	x	m
b	y	m
a	x	n
a	y	n
b	x	n
b	y	n



Сужение R(A1,A2,A3) на d1xd2xd3 : таблица

A1	A2	A3
a	x	m
a	y	n

Рис. 6.1. Сужение отношения $R(A1, A2, A3)$ на множество $d1 \times d2 \times d3$

Рассмотрим в качестве примера отношение ЗАРПЛАТА (СОТРУДНИК, ВЫПЛАТА), которое можно представить в виде двухстолбцовой таблицы. В этой таблице в столбце СОТРУДНИК указаны имена сотрудников некоторой фирмы или какого-то ее отделения, а в столбце ВЫПЛАТА — денежная сумма, причитающаяся соответствующему сотруднику.

Допустим, нас интересуют выплаты только Иванову и Петрову — элементам домена атрибута СОТРУДНИКИ. Обозначим через $D_{\text{СОТРУДНИК}}$ множество всех значений столбца СОТРУДНИК. Это и есть домен атрибута СОТРУДНИК. Аналогично, обозначим через $D_{\text{ВЫПЛАТА}}$ домен атрибута ВЫПЛАТА. Тогда интересующее нас отношение можно определить следующим образом:

$$\text{ЗАРПЛАТА (СОТРУДНИК, ВЫПЛАТА)} \bigcap \{\text{Иванов, Петров}\} \times D_{\text{ВЫПЛАТА}}$$

Здесь через {Иванов, Петров} обозначено требуемое подмножество домена атрибута СОТРУДНИК.

На естественном языке запрос на получение указанного множества кортежей выглядит более чем просто: "Выбрать кортежи, в которых значение атрибута СОТРУДНИК равно 'Иванов' или 'Петров'". Эквивалентный запрос можно еще сформулировать и так: "Выбрать кортежи, в которых значение атрибута СОТРУДНИК равно 'Иванов', и те кортежи, в которых значение атрибута СОТРУДНИК равно 'Петров'". Обратите внимание на использование союзов "или" и "и" в этих формулировках запросов. Если в первой формулировке "или" заменить на "и", то искомое множество кортежей будет заведомо пустым, поскольку атрибут СОТРУДНИК в одном и том же кортеже не может иметь два и более различных значения. Вторая формулировка допускает замену "и" на "или", поскольку в ней речь идет о различных значениях атрибута в различных кортежах. Применяя в обычной речи союзы "и" и "или", всегда следует отдавать себе отчет в том, что имеется в виду, — объединение или пересечение соответствующих множеств.

ПРИМЕЧАНИЕ

На языке SQL данный запрос формулируется следующим образом:

```
SELECT * FROM ЗАРПЛАТА WHERE СОТРУДНИК='Иванов' OR СОТРУДНИК='Петров'
```

По-русски это выглядит так: "выбрать кортежи (записи) из таблицы ЗАРПЛАТА, в которых атрибут (столбец) СОТРУДНИК имеет значение 'Иванов' или 'Петров'".

6.2.2. Проекция

Операция *проекции* отношения заключается в удалении из него указанных атрибутов. Пусть дано некоторое отношение $R(A1, A2, \ldots, An)$. Обозначим через **A** множество $A1, A2, \ldots, An$ всех атрибутов отношения, а через **X** — некоторое его подмножество (т. е. $\mathbf{X} \subseteq \mathbf{A}$). Тогда операция проекции отношения $R(\mathbf{A})$ на множество атрибутов $\mathbf{X}$ приводит к отношению $R[\mathbf{X}]$, кортежи которого получаются из кортежей исходного отношения путем удаления значений тех атрибутов, которые не принадлежат $\mathbf{X}$. Для любого кортежа z из отношения $R(\mathbf{A})$ его проекция $z[\mathbf{X}]$ на множество атрибутов $\mathbf{X}$ определяется аналогично: в этом кортеже следует оставить значения только атрибутов из $\mathbf{X}$.

Примечание

Квадратные скобки здесь указывают лишь на то, что рассматривается проекция отношения или кортежа на заключенные в эти скобки атрибуты.

Внимание

В данной главе мы обозначаем через $R[\mathbf{X}]$ проекцию отношения на множество атрибутов $\mathbf{X}$, а не сечение отношения 2-го рода, как мы это делали в *главе 5*. Кроме того, операцию вычитания множеств мы будем обозначать символом "-".

На рис. 6.2 в качестве примера показаны некоторое отношение $R(A1, A2, A3)$ и его проекция $R[A1, A3]$ на множество атрибутов $A1, A3$.

Примечание

На языке SQL проекция отношения (таблицы) на заданное множество атрибутов (столбцов) выполняется с помощью выражения:

```
SELECT список_столбцов FROM имя_таблицы
```

На практике часто операция проекции используется совместно с операцией селекции (выборки) кортежей. В этом случае применяется следующее выражение на языке SQL:

```
SELECT список_столбцов FROM имя_таблицы WHERE условие_выборки
```

R(A1,A2,A3) : таблица

A1	A2	A3
a	x	m
a	y	n
b	z	k
c	x	n
c	x	p

Проекция R(A1,A2,A3) на {A1,A3}

A1	A3
a	m
a	n
b	k
c	n
c	p

Рис. 6.2. Пример проекции отношения

6.2.3. Естественное соединение

Предположим, имеются два отношения, содержащие, кроме прочих, и одинаковые атрибуты. Рассмотрим, например, две таблицы $R1$ (ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА) и $R2$ (ПРЕПОДАВАТЕЛЬ, КАФЕДРА):

- Отношение $R1$ содержит сведения об именах преподавателей и учебных дисциплинах, занятия по которым они проводят. Вообще говоря, один и тот же преподаватель может заниматься несколькими дисциплинами, а по одной и той же дисциплине могут проводить занятия разные преподаватели.
- Отношение $R2$ содержит сведения о приписке преподавателей к кафедрам. Предполагается, что каждый преподаватель может быть приписан только к одной кафедре.

Естественно, может возникнуть задача сведе́ния этих двух отношений в одно (соединить две таблицы в одну) — $R3$ (ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА, КАФЕДРА). На рис. 6.3 показаны примеры отношений $R1$ и $R2$, а также результат их соединения $R3$. Важным обстоятельством, позволившим выполнить соединение двух отношений, является наличие у них общего атрибута ПРЕПОДАВАТЕЛЬ. Обратите внимание: проекция отношения $R3$ на атрибуты {ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА} в точности равна отношению $R1$, а проекция на атрибуты {ПРЕПОДАВАТЕЛЬ, КАФЕДРА} — отношению $R2$. Это означает, что при соединении отношений $R1$ и $R2$ в отношение $R3$ не было привнесено ничего нового и ничего не было потеряно. Другими словами, отношение $R3$ в точности представляет информацию, содержащуюся в отношениях $R1$ и $R2$.

Рассмотренная операция соединения отношений называется операцией *естественного соединения*. Теперь уточним ее определение.

Пусть даны два отношения $R1(\mathbf{A1})$ и $R2(\mathbf{A2})$, где **A1**, **A2** — множества атрибутов (а не одиночные атрибуты). Естественным соединением этих отношений $R1(\mathbf{A1}) * R2(\mathbf{A2})$ называется отношение $R3(\mathbf{A1} \cup \mathbf{A2})$, содержащее те и только те кортежи z, для которых выполняются одновременно следующие два условия:

- $z[\mathbf{A1}] \in R1(\mathbf{A1})$;
- $z[\mathbf{A2}] \in R2(\mathbf{A2})$.

Примечание

Квадратные скобки указывают на то, что рассматривается проекция на заключенные в них атрибуты. Операция естественного соединения здесь обозначена символом "*".

R1 : таблица

ПРЕПОДАВАТЕЛЬ	ДИСЦИПЛИНА
Иванов	Математический анализ
Иванов	Геометрия
Иванов	Теория вероятностей
Михайлов	Органическая химия
Михайлов	Неорганическая химия
Сергеев	Квантовая механика
Сергеев	Атомная физика
Сергеев	Физика твердого тела
Сидоров	Атомная физика
Сидоров	Квановая механика
Сидоров	Термодинамика
Петров	Математический анализ
Петров	Математическая логика
Петров	Теория множеств
Федоров	Топология
Федоров	Математическая логика

R2 : таблица

ПРЕПОДАВАТЕЛЬ	КАФЕДРА
Иванов	101
Михайлов	103
Петров	101
Сергеев	102
Сидоров	102
Федоров	101

R3=R1*R2 : запрос на выборку

ПРЕПОДАВАТЕЛЬ	ДИСЦИПЛИНА	КАФЕДРА
Иванов	Геометрия	101
Иванов	Математический анализ	101
Иванов	Теория вероятностей	101
Михайлов	Неорганическая химия	103
Михайлов	Органическая химия	103
Петров	Теория множеств	101
Петров	Математическая логика	101
Петров	Математический анализ	101
Сергеев	Квантовая механика	102
Сергеев	Атомная физика	102
Сергеев	Физика твердого тела	102
Сидоров	Атомная физика	102
Сидоров	Квановая механика	102
Сидоров	Термодинамика	102
Федоров	Топология	101
Федоров	Математическая логика	101

Рис. 6.3. Пример соединения отношений

R1 : таблица

A1	A2
a1	b1
a1	b2
a2	b1

R2 : таблица

A2	A3
b1	c1
b1	c2
b2	c1

R1xR2 : запрос на выборку

	A1	R1.A2	R2.A2	A3
√	a1	b1	b1	c1
	a1	b2	b1	c1
√	a2	b1	b1	c1
√	a1	b1	b1	c2
	a1	b2	b1	c2
√	a2	b1	b1	c2
	a1	b1	b2	c1
√	a1	b2	b2	c1
	a2	b1	b2	c1

R3=R1*R2 : запрос на выборку

A1	A2	A3
a2	b1	c1
a1	b1	c1
a2	b1	c2
a1	b1	c2
a1	b2	c1

Рис. 6.4. Естественное соединение отношений

Операцию естественного соединения двух отношений можно выразить через операции декартова произведения, проекции и пересечения. Рассмотрим три случая.

- Множества атрибутов отношений $R1(\mathbf{A1})$ и $R2(\mathbf{A2})$ не равны, но пересекаются (т. е. $\mathbf{A1} \neq \mathbf{A2}$, $\mathbf{A1} \cap \mathbf{A2} \neq \varnothing$). В данном случае:

  $$R1(\mathbf{A1}) * R2(\mathbf{A2}) = (R1(\mathbf{A1}) \times R2(\mathbf{A2} - \mathbf{A1})) \cap \\ \cap (R1(\mathbf{A1} - \mathbf{A2}) \times R2(\mathbf{A2})).$$

- Множества атрибутов отношений не пересекаются (т. е. $\mathbf{A1} \cap \mathbf{A2} = \varnothing$). В данном случае естественное соединение равно декартову произведению исходных отношений:

  $$R1(\mathbf{A1}) * R2(\mathbf{A2}) = R1(\mathbf{A1}) \times R2(\mathbf{A2}).$$

- Множества атрибутов отношений равны (т. е. $\mathbf{A1} = \mathbf{A2}$). Тогда естественное соединение равно пересечению исходных отношений:

  $$R1(\mathbf{A1}) * R2(\mathbf{A2}) = R1(\mathbf{A1}) \cap R2(\mathbf{A2}).$$

На рис. 6.4 показаны два исходных отношения, их декартово произведение и естественное соединение. В таблице декартова произведения выделены галочками те кортежи, из которых получены кортежи естественного соединения.

Примечание

Язык SQL предоставляет специальные средства для различных типов соединения таблиц. Естественное соединение отношений $R1$ (ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА) и $R2$ (ПРЕПОДАВАТЕЛЬ, КАФЕДРА), показанных на рис. 6.3, можно выполнить с помощью следующего SQL-выражения:

```
SELECT R1.*, R2.КАФЕДРА FROM R1, R2 WHERE R1.ПРЕПОДАВАТЕЛЬ=R2.ПРЕПОДАВАТЕЛЬ
```

По-русски это звучит так: "Выбрать все столбцы таблицы `R1` и столбец `КАФЕДРА` таблицы `R2` из декартова произведения `R1` и `R2` при условии, что преподаватели из этих таблиц одинаковы".

Здесь подразумевается именно декартово произведение, т. к. после ключевого слова `FROM` указаны две таблицы. В SQL также можно использовать выражение `NATURAL JOIN` для явного указания операции естественного соединения.

При работе с базами данных кроме естественного соединения используются и другие способы комбинации таблиц. Однако в реляционной теории понятие естественного соединения играет исключительно важную роль, поскольку через него определяется понятие корректной декомпозиции одной таблицы на несколько других. Этому посвящен следующий раздел.

6.3. Декомпозиция отношений

Таблица базы данных, представляющая некоторое отношение, может иметь очень большие размеры по количеству записей (строк) и столбцов, а также по общему объему содержащихся в них данных. Эффективность работы с таблицами по мере роста их объема уменьшается. Это обстоятельство вынуждает искать способы декомпозиции одной большой таблицы на несколько таблиц меньших размеров.

Так, в рассмотренном в предыдущем разделе отношении $R3$ (ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА, КАФЕДРА) (см. рис. 6.3) все атрибуты имеют повторяющиеся значения, а атрибут ПРЕПОДАВАТЕЛЬ однозначно определяет атрибут КАФЕДРА, поскольку каждый преподаватель приписан только к одной кафедре. Наличие такой зависимости между атрибутами ПРЕПОДАВАТЕЛЬ и ДИСЦИПЛИНА наталкивает на мысль, что декомпозиция возможна.

Отношения $R1$ (ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА) и $R2$ (ПРЕПОДАВАТЕЛЬ, КАФЕДРА) можно рассматривать как результат декомпозиции отношения $R3$, т. е. как операцию, обратную естественному соединению. Обратите внимание: общий объем данных в отношениях $R1$ и $R2$ меньше объема данных в отношении $R3$. Кроме того, в отношении $R2$ атрибут ПРЕПОДАВАТЕЛЬ не имеет повторяющихся значений.

Очевидно, что отношения, получающиеся в результате декомпозиции исходного отношения, являются проекциями последнего на некоторые подмножества атрибутов. Напомню, что проекция отношения на заданные атрибуты получается из этого отношения путем удаления из него всех атрибутов, кроме заданных. Заметим также, что для декомпозиции отношения на две проекции необходимо, чтобы оно было определено не меньше чем для трех атрибутов.

Декомпозиция может быть корректной или некорректной. Другими словами, она может быть обратимой или необратимой. Обратимость декомпозиции означает, что она является эквивалентным преобразованием исходной информации (форма изменяется, а содержание нет). Необратимость декомпозиции означает неэквивалентность преобразования: в содержание либо что-то вносится, либо что-то теряется. Далее мы дадим определение корректной декомпозиции и покажем пример некорректной декомпозиции.

6.3.1. Корректная декомпозиция

Декомпозиция отношения на некоторые его проекции называется *корректной*, если исходное отношение можно восстановить по этим проекциям с помощью операции естественного соединения. Точнее, отношение $R(\mathbf{A})$ корректно декомпозируется на свои проекции $R[\mathbf{X}]$ и $R[\mathbf{Y}]$ ($\mathbf{X} \subseteq \mathbf{A}$, $\mathbf{Y} \subseteq \mathbf{A}$), если выполняется равенство $R(\mathbf{A}) = R[\mathbf{X}] * R[\mathbf{Y}]$.

Например, декомпозиция отношения $R3$ (ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА, КАФЕДРА) (см. рис. 6.3) на его проекции

$R3$ [ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА] = $R1$ (ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА)

и

$R3$ [ПРЕПОДАВАТЕЛЬ, КАФЕДРА] = $R2$ (ПРЕПОДАВАТЕЛЬ, КАФЕДРА)

является корректной.

Это утверждение просто следует из того, что отношение $R3$ было получено из отношений $R1$ и $R2$ путем их естественного соединения (*см. разд. 6.2.3*).

6.3.2. Пример некорректной декомпозиции

На рис. 6.5 показаны некоторое отношение $R(A1, A2, A3)$, три его проекции $R[A1, A2]$, $R[A2, A3]$ и $R[A1, A3]$, а также естественное соединение этих трех проекций $R[A1, A2] * R[A2, A3] * R[A1, A3]$. Нетрудно заметить, что естественное соединение содержит кортеж $(0, 0, 0)$, которого нет в исходном отношении $R(A1, A2, A3)$. Таким образом, данная декомпозиция некорректна. Можно показать, что это отношение вообще нельзя корректно декомпозировать, т. к. в естественном соединении всегда будет появляться кортеж $(0, 0, 0)$.

R(A1,A2,A3) : таблица

A1	A2	A3
1	0	0
0	1	0
0	0	1

Проекция на A1,A2

A1	A2
1	0
0	1
0	0

Проекция на A2,A3

A2	A3
0	0
1	0
0	1

Проекция на A1,A3

A1	A3
1	0
0	0
0	1

Естественное соединение трех проекций

A1	A2	A3
1	0	0
0	1	0
0	0	1
0	0	0

Рис. 6.5. Пример некорректной декомпозиции отношения

Далее, говоря о декомпозиции отношения, мы будем всегда иметь в виду корректную декомпозицию.

6.3.3. Зависимости между атрибутами

Итак, отношение не всегда можно декомпозировать на две его проекции. Критерием возможности декомпозиции является наличие в отношении некоторых зависимостей между его атрибутами. Это общий критерий. Если в чем-либо обнаруживаются какие-нибудь зависимости, то есть принципиальная возможность представить это нечто в более компактном виде.

Чтобы отношение можно было декомпозировать на две его проекции, достаточно существования в нем так называемых *функциональных зависимостей*. Необходимым и достаточным условием декомпозиции отношения является наличие в нем многозначных зависимостей. Рассмотрим эти зависимости по порядку.

Функциональные зависимости

Функциональные зависимости просты для понимания и обычно легко обнаруживаются в отношении. Между атрибутами *A* и *B* существует функциональная зависимость, если любое значение атрибута *A* однозначно определяет значение атрибута *B*.

Рассмотрим несколько примеров отношений, показанных на рис. 6.3.

- В отношении *R*2 (ПРЕПОДАВАТЕЛЬ, КАФЕДРА) между атрибутами ПРЕПОДАВАТЕЛЬ и КАФЕДРА имеется функциональная зависимость, поскольку каждый преподаватель может быть приписан лишь к одной кафедре (предполагается, что совместительство не допускается). Другими словами, по значению атрибута ПРЕПОДАВАТЕЛЬ можно однозначно определить, на какой кафедре он работает. Однако между атрибутами КАФЕДРА и ПРЕПОДАВАТЕЛЬ в отношении *R*2 нет функциональной зависимости: на одной и той же кафедре могут работать несколько преподавателей.
- В отношении *R*1 (ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА) нет ни одной функциональной зависимости: один и тот же преподаватель может заниматься несколькими дисциплинами, а одну и ту же дисциплину могут вести несколько преподавателей.
- В отношении *R*3 (ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА, КАФЕДРА) существуют две функциональных зависимости:
  - между атрибутами ПРЕПОДАВАТЕЛЬ и КАФЕДРА;
  - между атрибутами ДИСЦИПЛИНА и КАФЕДРА.

Вообще говоря, функциональная зависимость может иметь место не только между отдельными атрибутами, но и между подмножествами атрибутов отношения. Для этого общего случая дадим определение функциональной зависимости.

Пусть дано некоторое отношение $R(\mathbf{A})$ с атрибутами из множества **A**, т. е. **A** — это множество атрибутов, а не отдельный атрибут. Обозначим через **X** и **Y** некоторые подмножества атрибутов ($\mathbf{X} \subseteq \mathbf{A}$, $\mathbf{Y} \subseteq \mathbf{A}$). В отношении $R(\mathbf{A})$ выполняется функциональная зависимость между атрибутами **X** и **Y**, обозначаемая как $\mathbf{X} -> \mathbf{Y}$, если для любых кортежей $z1$ и $z2$ этого отношения из равенства их проекций на множество атрибутов ***X*** следует равенство их проекций на множество атрибутов **Y**. Иначе говоря, функциональная зависимость $\mathbf{X} -> \mathbf{Y}$ выполняется, если справедливо следующее высказывание: для любых $z1$ и $z2$ из равенства $z1[\mathbf{X}] = z2[\mathbf{X}]$ следует равенство $z1[\mathbf{Y}] = z2[\mathbf{Y}]$.

> ***Внимание***
>
> В данной главе для обозначения зависимости между атрибутами используется стрелка $->$, которую не следует путать с символом импликации (логического следования).

Если выполняется функциональная зависимость $\mathbf{X} -> \mathbf{Y}$, то говорят, что атрибуты **X** функционально (однозначно) определяют атрибуты **Y**.

Например, в отношении $R3$ (ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА, КАФЕДРА), показанном на рис. 6.3, имеются следующие функциональные зависимости:

- {ПРЕПОДАВАТЕЛЬ} -> {КАФЕДРА} — преподаватель функционально определяет кафедру;
- {ДИСЦИПЛИНА} -> {КАФЕДРА} — дисциплина функционально определяет кафедру;
- {ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА} -> {КАФЕДРА} — преподаватель и дисциплина функционально определяют кафедру.

Рассмотрим еще один пример. Предположим, отношение ПРОДАЖИ (ТОВАР, КОЛИЧЕСТВО, ЦЕНА, СТОИМОСТЬ) содержит сведения о проданных товарах. Стоимость товара однозначно определяется его количеством и ценой, поскольку существует простая формула: *стоимость* = *количество* × *цена*. Поэтому в данном отношении имеется функциональная зависимость {КОЛИЧЕСТВО, ЦЕНА} -> {СТОИМОСТЬ}.

Если в отношении $R(\mathbf{A})$ выполняется функциональная зависимость $\mathbf{X} -> \mathbf{Y}$ ($\mathbf{X} \subseteq \mathbf{A}$, $\mathbf{Y} \subseteq \mathbf{A}$), то это отношение декомпозируется на две его проекции $R[\mathbf{X} \cup \mathbf{Y}]$ и $R[\mathbf{X} \cup (\mathbf{A} - \mathbf{Y})]$. Иначе говоря, справедливо следующее равенство: $R(\mathbf{A}) = R[\mathbf{X} \cup \mathbf{Y}] * R[\mathbf{X} \cup (\mathbf{A} - \mathbf{Y})]$.

Здесь символом "*" обозначена операция естественного соединения.

Таким образом, исходное отношение $R(\mathbf{A})$ можно восстановить с помощью операции естественного соединения двух своих проекций — $R[\mathbf{X}\bigcup\mathbf{Y}]$ и $R[\mathbf{X}\bigcup(\mathbf{A}-\mathbf{Y})]$.

Например, в отношении $R3$ (ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА, КАФЕДРА), показанном на рис. 6.3, имеется функциональная зависимость {ПРЕПОДАВАТЕЛЬ} –> {КАФЕДРА}. Следовательно, это отношение можно декомпозировать на две проекции:

- $R3$ [ПРЕПОДАВАТЕЛЬ, КАФЕДРА];
- $R3$ [ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА].

В данном примере можно принять такие обозначения:

- $\mathbf{A}$ = {ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА, КАФЕДРА};
- $\mathbf{X}$ = {ПРЕПОДАВАТЕЛЬ};
- $\mathbf{Y}$ = {КАФЕДРА};
- $\mathbf{X}\bigcup\mathbf{Y}$ = {ПРЕПОДАВАТЕЛЬ, КАФЕДРА};
- $\mathbf{A}-\mathbf{Y}$ = {ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА};
- $\mathbf{X}\bigcup(\mathbf{A}-\mathbf{Y})$ = {ПРЕПОДАВАТЕЛЬ} $\bigcup$ {ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА} = {ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА}.

Отношение ПРОДАЖИ (ТОВАР, КОЛИЧЕСТВО, ЦЕНА, СТОИМОСТЬ), имеющее функциональную зависимость {КОЛИЧЕСТВО, ЦЕНА} –> {СТОИМОСТЬ}, можно декомпозировать на такие проекции:

- ПРОДАЖИ [КОЛИЧЕСТВО, ЦЕНА, СТОИМОСТЬ]
- ПРОДАЖИ [ТОВАР, КОЛИЧЕСТВО, ЦЕНА]

В любом отношении существует хотя бы одна функциональная зависимость. Например, для произвольного отношения $R(\mathbf{A})$ всегда выполняется зависимость $\mathbf{A} -> \mathbf{A}$, поскольку по определению в любом отношении не может быть одинаковых кортежей и, следовательно, любой кортеж однозначно определяет сам себя. Такая функциональная зависимость является тривиальной. Точнее, функциональная зависимость $\mathbf{X} -> \mathbf{Y}$ называется *тривиальной*, если $\mathbf{Y} \subseteq \mathbf{X}$.

Наличие в отношении нетривиальной функциональной зависимости является достаточным, но не необходимым условием корректной декомпозиции этого отношения на две проекции. Это означает, что если функциональная зависимость есть, то декомпозиция возможна. Однако корректная декомпозиция может быть выполнена и для некоторых отношений, в которых нет ни одной функциональной зависимости. Например, на рис. 6.4 было показано отношение $R3 = R1 * R2$, в котором нет ни одной нетривиальной функциональной зависимости. Однако оно может быть корректно декомпозировано на две своих проекции $R1$ и $R2$ (см. рис. 6.4).

Приведем несколько свойств функциональных зависимостей.

- Пусть $\mathbf{Y} = \{A1, \ldots, Ak\}$ — некоторое подмножество атрибутов отношения. Если в этом отношении выполняется функциональная зависимость $\mathbf{X} \to \mathbf{Y}$, то выполняется и множество функциональных зависимостей $\{\mathbf{X} \to Ai\}$, где $i = 1, \ldots, k$. Другими словами, коль скоро выполняется функциональная зависимость между множествами атрибутов **X** и **Y**, то выполняются и зависимости между **X** и одиночными атрибутами из **Y**. Обратное утверждение тоже верно: если в отношении выполняются функциональные зависимости $\{\mathbf{X} \to Ai\}$ ($i = 1, \ldots, k$), то выполняется и функциональная зависимость $\mathbf{X} \to \mathbf{Y}$, где $\mathbf{Y} = \{A1, \ldots, Ak\}$.
- В отношении $R(\mathbf{A})$ функциональная зависимость $\mathbf{X} \to \mathbf{Y}$ выполняется тогда и только тогда, когда она выполняется в проекции $R[\mathbf{Z}]$, где $\mathbf{X} \cup \mathbf{Y} \subseteq \mathbf{Z}$.

Многозначные зависимости

Обобщением понятия функциональной (однозначной) зависимости является понятие *многозначной зависимости*. Наличие в отношении многозначной зависимости является необходимым и достаточным условием его декомпозиции на две проекции.

> ***Примечание***
>
> Напомню, что выполнение функциональной зависимости обеспечивает лишь достаточное условие декомпозиции. Вот почему некоторые отношения без функциональных зависимостей все-таки могут быть декомпозированы.

Многозначная зависимость имеет смысл только для отношений с тремя и более атрибутами. Проще всего понятие многозначной зависимости объяснить для отношения с тремя атрибутами. Пусть дано некоторое отношение $R(A1, A2, A3)$, для которого выполняются следующие условия.

- Атрибут $A1$ не определяет однозначно атрибут $A2$, т. е. нет функциональной зависимости $\{A1 \to A2\}$.
- Атрибуты $A2$ и $A3$ не зависят друг от друга, т. е. в отношении $R(A1, A2, A3)$ любое значение $A2$ может сочетаться в кортежах с любым значением $A3$.

Тогда в отношении $R(A1, A2, A3)$ выполняется многозначная зависимость между атрибутами $A1$ и $A2$, которая обозначается как $\{A1 \to\to A2\}$. При этом говорят, что атрибут $A1$ многозначно определяет атрибут $A2$.

Это частное определение многозначной зависимости, возможно, не очень понятное. Поэтому поясним его на примере.

Пусть дано отношение *R* (ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА, УВЛЕЧЕНИЕ), которое содержит сведения о преподавателях, дисциплинах, занятия по которым они проводят, и личных увлечениях (хобби). Предполагается, что один преподаватель может проводить занятия по нескольким дисциплинам, т. е. атрибут ПРЕПОДАВАТЕЛЬ не определяет однозначно атрибут ДИСЦИПЛИНА. Далее, мы считаем, что нет никакой связи между дисциплинами и увлечениями преподавателей. Иначе говоря, если известно, что преподаватель читает студентам, например, математический анализ, то из этого нельзя однозначно определить, чем конкретно он увлекается. И наоборот, если мы знаем о конкретном увлечении преподавателя классической музыкой, то еще не можем однозначно сказать, какой именно дисциплиной он занимается. Все это означает, что в рассматриваемом отношении значения атрибутов ДИСЦИПЛИНА и УВЛЕЧЕНИЕ могут сочетаться произвольным образом.

Итак, атрибуты ДИСЦИПЛИНА и УВЛЕЧЕНИЕ не зависят друг от друга, а атрибут ПРЕПОДАВАТЕЛЬ определяет атрибут ДИСЦИПЛИНА не однозначно. Пример рассмотренного отношения показан на рис. 6.6.

Поскольку в отношении *R* постулируется независимость атрибутов ДИСЦИПЛИНА и УВЛЕЧЕНИЕ, в этом отношении каждое сочетание значений атрибутов ПРЕПОДАВАТЕЛЬ и ДИСЦИПЛИНА должно сочетаться с каждым значением атрибута УВЛЕЧЕНИЕ. В примере отношения, показанном на рис. 6.6, преподаватель Иванов проводит занятия по двум дисциплинам и имеет два хобби. Поэтому ему в таблице посвящено четыре записи (кортежа). Преподаватель Михайлов занят двумя дисциплинами и имеет одно увлечение, поэтому информация о нем представлена в отношении двумя записями. В данном отношении имеется многозначная зависимость ПРЕПОДАВАТЕЛЬ –>> ДИСЦИПЛИНА или, другими словами, атрибут ПРЕПОДАВАТЕЛЬ многозначно определяет атрибут ДИСЦИПЛИНА.

R : таблица

ПРЕПОДАВАТЕЛЬ	ДИСЦИПЛИНА	УВЛЕЧЕНИЕ
Иванов	Математический анализ	Классическая музыка
Иванов	Математический анализ	Футбол
Иванов	Геометрия	Классическая музыка
Иванов	Геометрия	Футбол
Михайлов	Органическая химия	Поп-музыка
Михайлов	Неорганическая химия	Поп-музыка

Рис. 6.6. Отношение с многозначной зависимостью ПРЕПОДАВАТЕЛЬ –>> ДИСЦИПЛИНА

Примечание

Мы бы не говорили о многозначной зависимости ПРЕПОДАВАТЕЛЬ –>> ДИСЦИПЛИНА, если бы между атрибутами ДИСЦИПЛИНА и УВЛЕЧЕНИЕ существовала какая-нибудь зависимость.

Очевидно, что с точки зрения хранения информации такое представление отношения в виде одной таблицы избыточно. И действительно, его можно декомпозировать так, как показано на рис. 6.7.

R : таблица

ПРЕПОДАВАТЕЛЬ	ДИСЦИПЛИНА	УВЛЕЧЕНИЕ
Иванов	Математический анализ	Классическая музыка
Иванов	Математический анализ	Футбол
Иванов	Геометрия	Классическая музыка
Иванов	Геометрия	Футбол
Михайлов	Органическая химия	Поп-музыка
Михайлов	Неорганическая химия	Поп-музыка

Проекция на ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА

ПРЕПОДАВАТЕЛЬ	ДИСЦИПЛИНА
Иванов	Математический анализ
Иванов	Геометрия
Михайлов	Органическая химия
Михайлов	Неорганическая химия

Проекция на ПРЕПОДАВАТЕЛЬ, УВЛЕЧЕНИЕ

Преподаватель	УВЛЕЧЕНИЕ
Иванов	Классическая музыка
Иванов	Футбол
Михайлов	Поп-музыка

Рис. 6.7. Декомпозиция отношения, содержащего многозначную зависимость ПРЕПОДАВАТЕЛЬ –>> ДИСЦИПЛИНА

Поскольку выполняется многозначная зависимость ПРЕПОДАВАТЕЛЬ –>> ДИСЦИПЛИНА, атрибуты ДИСЦИПЛИНА и УВЛЕЧЕНИЕ не зависят друг от друга. Поэтому имеется возможность представить связи ПРЕПОДАВАТЕЛЬ—ДИСЦИПЛИНА и ПРЕПОДАВАТЕЛЬ—УВЛЕЧЕНИЕ в виде отдельных отношений, причем так, чтобы исходное отношение R (ПРЕПОДАВАТЕЛЬ, ДИСЦИПЛИНА, УВЛЕЧЕНИЕ) полностью восстанавливалось с помощью операции естественного соединения.

Обобщим вышесказанное. Отношение $R(A1, A2, A3)$, в котором выполняется многозначная зависимость $A1 \rightarrow\!\!> A2$, можно декомпозировать на две проекции: $R[A1, A2]$ и $R[A1, A3]$. Это означает, что выполняется равенство $R(A1, A2, A3) = R[A1, A2] * R[A1, A3]$.

Следующий пример отличается от рассмотренного ранее лишь уровнем абстракции. Пусть дано некоторое отношение $R(A1, A2, A3)$, представленное в виде таблицы (рис. 6.8). Это представление отношения. Вся информация о зависимостях между его атрибутами не постулируется какими-либо предложениями, а заключена в самом представлении отношения. Иначе говоря, зависимости, если они существуют, могут быть выявлены путем анализа самого содержимого отношения, хотя это и не всегда простая задача.

R(A1,A2,A3) : таблица

A1	A2	A3
a2	b1	c1
a2	b1	c2
a1	b2	c2
a1	b1	c1
a1	b1	c2
a1	b2	c1

Рис. 6.8. Отношение с многозначной зависимостью *A*1 –>> *A*2

В данном отношении выполняется многозначная зависимость $A1 \text{–>>} A2$. Поэтому его можно декомпозировать на две проекции — $R[A1, A2]$ и $R[A1, A3]$, как показано на рис. 6.9.

R(A1,A2,A3) : таблица

A1	A2	A3
a2	b1	c1
a2	b1	c2
a1	b2	c2
a1	b1	c1
a1	b1	c2
a1	b2	c1

Проекция на A1 , A2

A1	A2
a1	b1
a1	b2
a2	b1

Проекция на A1 , A3

A1	A3
a1	c1
a1	c2
a2	c1
a2	c2

Рис. 6.9. Декомпозиция отношения *R*(*A*1, *A*2, *A*3)
с многозначной зависимостью *A*1 –>> *A*2

В рассмотренном отношении (рис. 1.8) кроме многозначной зависимости $A1 -\!\!>> A2$ выполняется еще и многозначная зависимость $A1 -\!\!>> A3$. Следовательно, исходное отношение можно декомпозировать на проекции $R[A1, A3]$ и $R[A1, A2]$, т. е. на те же самые, что позволяет получить зависимость $A1 -\!\!>> A2$.

В общем случае многозначная зависимость определяется не только между одноэлементными множествами атрибутов.

Ранее уже отмечалось, что наличие многозначной зависимости в отношении является необходимым и достаточным условием декомпозиции отношения на две его проекции. Если в отношении $R(\mathbf{A})$ выполняется многозначная зависимость $\mathbf{X} -\!\!>> \mathbf{Y}$, то это отношение можно декомпозировать на две проекции $R[\mathbf{X}, \mathbf{Y}]$ и $R[\mathbf{X}, \mathbf{Z}]$. Это утверждение о том, что многозначная зависимость *необходима* для декомпозиции.

То, что наличие многозначной зависимости является еще и *достаточным* условием декомпозиции, вытекает из следующего. Пусть отношение $R(\mathbf{A})$ декомпозируется на две проекции — $R[\mathbf{X}]$ и $R[\mathbf{Y}]$ (это означает, что $R(\mathbf{A}) = R[\mathbf{X}] * R[\mathbf{Y}]$). Тогда в отношении $R(\mathbf{A})$ выполняются две многозначные зависимости $\mathbf{X} \bigcap \mathbf{Y} -\!\!>> \mathbf{X} - \mathbf{Y}$ и $\mathbf{X} \bigcap \mathbf{Y} -\!\!>> \mathbf{Y} - \mathbf{X}$. Иначе говоря, если нам удалось декомпозировать отношение на две проекции, то это означает, что в отношении выполняются по крайней мере две многозначные зависимости.

На рис. 6.10 представлены некоторое отношение $R(A1, A2, A3)$ и две его проекции — $R[A1, A2]$ и $R[A2, A3]$. Нетрудно проверить, что выполняется равенство $R(A1, A2, A3) = R[A1, A2] * R[A2, A3]$. Следовательно, отношение декомпозируется на эти проекции. Но тогда в отношении $R(A1, A2, A3)$ выполняются две многозначные зависимости:

- $A2 -\!\!>> A1$ — поскольку $A1, A2 \bigcap A2, A3 = A2$ и $A1, A2 - A2, A3 = A1$;
- $A2 -\!\!>> A3$ — поскольку $A2, A3 - A1, A2 = A3$.

С учетом первой зависимости исходное отношение можно декомпозировать на проекции $R[A2, A1]$ и $R[A2, A3]$, показанные на рис. 6.10. С учетом второй зависимости это отношение можно декомпозировать на такие же проекции.

Нетрудно заметить, что в определении многозначной зависимости множества атрибутов **Y** и **Z** играют симметричные роли: если что-либо справедливо для **Y**, то это же справедливо и для **Z**, и наоборот. Поэтому можно сформулировать следующее утверждение: многозначная зависимость $\mathbf{X} -\!\!>> \mathbf{Y}$ выполняется в отношении $R(\mathbf{A})$ тогда и только тогда, когда в $R(\mathbf{A})$ выполняется многозначная зависимость $\mathbf{X} -\!\!>> \mathbf{Z}$.

R(A1,A2,A3) : таблица

A1	A2	A3
a1	b1	c1
a1	b1	c2
a1	b2	c1
a2	b1	c1
a2	b1	c2

Проекция на A1, A2

A1	A2
a1	b1
a1	b2
a2	b1

Проекция на A2,A3

A2	A3
b1	c1
b1	c2
b2	c1

Рис. 6.10. Из данной декомпозиции отношения следует наличие в нем зависимостей $A2 \;{-}{>}{>}\; A1$ и $A2 \;{-}{>}{>}\; A3$

Многозначные (а также функциональные) зависимости сохраняются в проекциях отношения, которые не затрагивают их левой части. Другими словами, если в отношении $R(\mathbf{A})$ выполняется многозначная зависимость $\mathbf{X} {-}{>}{>} \mathbf{Y}$, то в любой проекции $R[\mathbf{U}]$, для которой $\mathbf{X} \subseteq \mathbf{U}$, выполняется и многозначная зависимость $\mathbf{X} {-}{>}{>} \mathbf{Y} \cap \mathbf{U}$. Для функциональных зависимостей верно и обратное утверждение: если в проекции выполняется функциональная зависимость, то эта же зависимость выполняется и во всем отношении. Для многозначных зависимостей это утверждение неверно: из выполнимости многозначной зависимости в некоторой проекции отношения еще не следует ее выполнимость во всем отношении.

6.3.4. Правила вывода зависимостей

Все утверждения, сформулированные для многозначных зависимостей, справедливы и для функциональных зависимостей. Поэтому далее мы будем рассматривать только многозначные зависимости.

Если в отношении было каким-либо образом замечено выполнение одних зависимостей, то другие могут быть выведены логически из тех, что обнаружены. В основе декомпозиции, как известно, находится та или иная зависимость. Таким образом, при наличии нескольких зависимостей мы можем выбирать декомпозицию. Не всякая декомпозиция окажется удачной с точки зрения экономного представления данных. Не стоит также забывать и о том, что результаты декомпозиции (проекции исходного отношения) могут использоваться как исходный материал для дальнейшей декомпозиции: то, что хорошо на первом этапе декомпозиции, может плохо сказаться на всей многоэтапной декомпозиции. Иначе говоря, оптимизация декомпозиции в целом не означает, что каждый отдельный ее этап должен быть наилучшим.

Примечание

Данный раздел посвящен выводам одних зависимостей из факта наличия других и адресуется тем, кто хочет глубже познать реляционную теорию либо поупражняться в логике и алгебре множеств. Этот материал не из легких и может быть пропущен при первом чтении данной главы.

В систему правил вывода многозначных зависимостей входят четыре основных правила. Прежде чем их перечислить, договоримся об обозначениях. Через **A** будем обозначать множество всех атрибутов отношения, а через **X**, **Y** и **Z** — подмножества атрибутов ($\mathbf{X} \subseteq \mathbf{A}$, $\mathbf{Y} \subseteq \mathbf{A}$, $\mathbf{Z} \subseteq \mathbf{A}$). Разумеется, каждое из них может содержать и только единственный элемент — какой-нибудь один атрибут отношения. Для обозначения многозначной зависимости используется, как и раньше, символ "$->>$". Однако все приведенные далее утверждения справедливы не только для многозначных зависимостей, но и для функциональных зависимостей.

Итак, для зависимостей между атрибутами отношения выполняются следующие правила вывода:

1. *Правило дополнения.* Если выполняется зависимость $\mathbf{X} ->> \mathbf{Y}$, то выполняется и зависимость $\mathbf{X} ->> \mathbf{Z}$ для всех **Z**, таких, что $\mathbf{Y} \cap \mathbf{Z} = \mathbf{X}$ и $\mathbf{X} \cup \mathbf{Y} \cup \mathbf{Z} = \mathbf{A}$.
2. *Правило рефлексивности*. Если $\mathbf{X} \subseteq \mathbf{Y}$, то $\mathbf{X} ->> \mathbf{Y}$.
3. *Правило пополнения*. Если выполняется зависимость $\mathbf{X} ->> \mathbf{Y}$ и для некоторых подмножеств атрибутов **V** и **W** справедливо включение $\mathbf{V} \subseteq \mathbf{W}$, то выполняется зависимость $\mathbf{X} \cup \mathbf{W} ->> \mathbf{Y} \cup \mathbf{V}$.
4. *Правило транзитивности*. Если выполняются две зависимости — $\mathbf{X} ->> \mathbf{Y}$ и $\mathbf{Y} ->> \mathbf{Z}$, то выполняется и зависимость $\mathbf{X} ->> \mathbf{Z} - \mathbf{Y}$.

Из этих четырех основных правил можно логически вывести еще три дополнительных правила:

1. *Правило объединения*. Если выполняются две зависимости — $\mathbf{X} ->> \mathbf{Y}$ и $\mathbf{X} ->> \mathbf{Z}$, то выполняется и зависимость $\mathbf{X} ->> \mathbf{Y} \cup \mathbf{Z}$.
2. *Правило разности*. Если выполняются две зависимости — $\mathbf{X} ->> \mathbf{Y}$ и $\mathbf{X} ->> \mathbf{Z}$, то выполняется и зависимость $\mathbf{X} ->> \mathbf{Y} - \mathbf{Z}$.
3. *Правило пересечения*. Если выполняются две зависимости — $\mathbf{X} ->> \mathbf{Y}$ и $\mathbf{X} ->> \mathbf{Z}$, то выполняется и зависимость $\mathbf{X} ->> \mathbf{Y} \cap \mathbf{Z}$.

ПРИМЕЧАНИЕ

Дополнительные правила удобно использовать для сокращения длины вывода зависимостей.

В качестве примера, а также для упражнения в обращении с зависимостями, рассмотрим вывод из четырех основных правил правила разности и правила пересечения.

Пусть даны две зависимости $\mathbf{X}\text{–>>}\mathbf{Y}$ и $\mathbf{X}\text{–>>}\mathbf{Z}$, тогда вывод *правила разности* состоит из следующих четырех шагов:

1. По правилу пополнения из $\mathbf{X}\text{–>>}\mathbf{Z}$ следует $\mathbf{X}\bigcup\mathbf{X}\text{–>>}\mathbf{X}\bigcup\mathbf{Z}$, т. е. $\mathbf{X}\text{–>>}\mathbf{X}\bigcup\mathbf{Z}$.
2. По правилу пополнения из $\mathbf{X}\text{–>>}\mathbf{Y}$ следует $\mathbf{X}\bigcup\mathbf{Z}\text{–>>}\mathbf{Y}\bigcup\mathbf{Z}$.
3. По правилу транзитивности из $\mathbf{X}\text{–>>}\mathbf{X}\bigcup\mathbf{Z}$ и $\mathbf{X}\bigcup\mathbf{Z}\text{–>>}\mathbf{Y}\bigcup\mathbf{Z}$ следует $\mathbf{X}\text{–>>}(\mathbf{Y}\bigcup\mathbf{Z})-(\mathbf{X}\bigcup\mathbf{Z})$. Однако $(\mathbf{Y}\bigcup\mathbf{Z})-(\mathbf{X}\bigcup\mathbf{Z})=\mathbf{Y}-\mathbf{Z}-\mathbf{X}$, поэтому выполняется зависимость $\mathbf{X}\text{–>>}\mathbf{Y}-\mathbf{Z}-\mathbf{X}$.
4. Так как $(\mathbf{Y}-\mathbf{Z})\bigcap\mathbf{X}\subseteq\mathbf{X}$, то по правилу пополнения из $\mathbf{X}\text{–>>}\mathbf{Y}-\mathbf{Z}-\mathbf{X}$ следует $\mathbf{X}\bigcup\mathbf{X}\text{–>>}(\mathbf{Y}-\mathbf{Z}-\mathbf{X})\bigcup(\mathbf{Y}-\mathbf{Z})\bigcap\mathbf{X}$. Но $\mathbf{X}\bigcup\mathbf{X}=\mathbf{X}$, а $(\mathbf{Y}-\mathbf{Z}-\mathbf{X})\bigcup(\mathbf{Y}-\mathbf{Z})\bigcap\mathbf{X}=\mathbf{Y}-\mathbf{Z}$, поэтому справедлива зависимость $\mathbf{X}\text{–>>}\mathbf{Y}-\mathbf{Z}$.

Теперь выведем *правило пересечения*. Поскольку $\mathbf{Y}\bigcap\mathbf{Z}=\mathbf{Y}-(\mathbf{Y}-\mathbf{Z})$, правило пересечения получается из правила разности: из $\mathbf{X}\text{–>>}\mathbf{Y}$ и $\mathbf{X}\text{–>>}\mathbf{Y}-\mathbf{Z}$ следует $\mathbf{X}\text{–>>}\mathbf{Y}-\mathbf{Y}-\mathbf{Z}$ или $\mathbf{X}\text{–>>}\mathbf{Y}\bigcap\mathbf{Z}$.

ПРИМЕЧАНИЕ

Аналогичным образом можно обосновать *правило объединения*.

6.3.5. Ключи

При синтезе и анализе отношений (таблиц базы данных) важную роль играет понятие ключа отношения (таблицы). Это понятие — производное от понятия функциональной зависимости, рассмотренного в предыдущем разделе. Можно считать, что ключ является аргументом (левой частью) функциональной зависимости между множествами атрибутов отношения.

Ключ (*key*) — множество из одного или нескольких атрибутов, которое однозначно определяет (идентифицирует) всю запись в отношении. Таким образом, множество атрибутов $\mathbf{X}$ является ключом отношения $R(\mathbf{A})$, если в этом отношении есть функциональная зависимость $\mathbf{X}\text{–>}\mathbf{A}$. И наоборот, если в отношении $R(\mathbf{A})$ имеется функциональная зависимость $\mathbf{X} -> \mathbf{A}$, то множество атрибутов **X** является ключом этого отношения. Очевидно, что главным признаком ключа отношения является уникальность (неповторяемость) его значений.

Из самого определения отношения следует, что в любом отношении всегда найдется ключ. По крайней мере, ключом может быть множество всех атрибутов **A**, т. к. в любом отношении имеется тривиальная функциональная зависимость $\mathbf{A} -> \mathbf{A}$, поскольку каждая запись встречается в отношении всего один раз и, следовательно, однозначно идентифицирует сама себя. Однако при разработке таблиц базы данных обычно интересуются не тривиальными ключами, а состоящими из одного или небольшого количества атрибутов. Действительно, тривиальные зависимости не позволяют декомпозировать отношение на две проекции.

Рассмотрим в качестве примера таблицу, представляющую отношение `Студенты` (`Фамилия`, `Имя`, `Отчество`, `Учебная_группа`). Какие атрибуты в этом отношении могут составить ключ? Предполагается, что каждый студент может находиться только в одной учебной группе. Казалось бы, множество атрибутов {`Фамилия`, `Имя`, `Отчество`} должно однозначно идентифицировать студента, а следовательно, и группу. Это было бы верно, если бы мы исключили возможность существования в одной группе полных тезок (однофамильцев с одинаковыми именами и отчествами) — хотя и маловероятно, но это все же возможно. Таким образом, множество атрибутов {`Фамилия`, `Имя`, `Отчество`} не подходит на роль нетривиального ключа. В подобных ситуациях создают специальный атрибут (столбец), предназначенный играть роль ключа. В нашем примере это может быть `ID_студента` (идентификатор студента). Значения этого атрибута могут быть какими угодно, но обязательно уникальными. В результате получается таблица `Студенты` (`ID_студента`, `Фамилия`, `Имя`, `Отчество`, `Учебная_группа`). А если `ID_студента` является ключом, то в данном отношении имеется функциональная зависимость {`ID_студента`} `->` {`ID_студента`, `Фамилия`, `Имя`, `Отчество`, `Учебная_группа`}, а также зависимость {`ID_студента`} `->` {`Фамилия`, `Имя`, `Отчество`, `Учебная_группа`}. Последнюю зависимость, очевидно, можно использовать для декомпозиции исходного отношения на следующие два: `Студенты_список` (`ID_студента`, `Фамилия`, `Имя`, `Отчество`) и `Группы` (`ID_студента`, `Учебная_группа`). На рис. 6.11 показаны исходное отношение `Студенты` и две его проекции.

Ключ, состоящий из одного атрибута, называют *простым*, а из нескольких — *составным*. Ключи, рассмотренные в данном разделе, также называют *первичными* (*primary key*).

Студенты : таблица

ID_студента	Фамилия	Имя	Отчество	Учебная_группа
001	Иванов	Иван	Иванович	201
002	Петров	Петр	Петрович	201
003	Сидоров	Сидор	Сидорович	201
004	Федоров	Федор	Федорович	202
005	Иванов	Иван	Иванович	202
006	Иванов	Иван	Иванович	201
007	Николаев	Николай	Николаевич	202
008	Степанов	Степан	Степанович	203
009	Михайлов	Михаил	Михайлович	203
010	Иванов	Иван	Иванович	202

Студенты_список : таблица

ID_студента	Фамилия	Имя	Отчество
001	Иванов	Иван	Иванович
002	Петров	Петр	Петрович
003	Сидоров	Сидор	Сидорович
004	Федоров	Федор	Федорович
005	Иванов	Иван	Иванович
006	Иванов	Иван	Иванович
007	Николаев	Николай	Николаевич
008	Степанов	Степан	Степанович
009	Михайлов	Михаил	Михайлович
010	Иванов	Иван	Иванович

Группы : таблица

ID_студента	Учебная_группа
001	201
002	201
003	201
004	202
005	202
006	201
007	202
008	203
009	203
010	202

Рис. 6.11. Декомпозиция отношения `Студенты` с ключом `ID_студента` на две проекции

6.4. Ограничения целостности отношений

До сих пор мы рассматривали отношения, представленные в виде таблиц своими кортежами (записями). Зависимости между атрибутами выражали некоторую целостность отношений, а корректная декомпозиция отношений на основе зависимостей рассматривалась как сохраняющая их целостность. Таким образом, мы изучали реляционные базы данных главным образом как теорию декомпозиции отношений, заданных экстенсионально, т. е. посредством наборов записей. На практике приходится проектировать базу данных как набор из нескольких таблиц, которые изначально пусты и только со временем будут заполняться конкретными данными. Иначе говоря, довольно сложные базы данных проектируются путем разработки множества отдельных таблиц с последующей установкой связей между ними, т. е. путем композиции частей в некое целое. При этом не исключается и декомпозиция таблиц на свои проекции.

Однако каждая таблица в отдельности, а также совокупность таблиц обычно являются неслучайными комбинациями атрибутов, произвольно распределенными между различными таблицами. Как отдельные таблицы, так и вся их совокупность, называемая *базой данных*, обладает некоторой целостностью, которая выражается через различного рода ограничения, накладываемые на значения столбцов и связи между ними. Таким образом, при проектировании базы данных мы заняты не обнаружением зависимостей для использования их при декомпозиции, а наоборот, их заданием с целью объединения таблиц в целостную систему связанных таблиц, содержащих непротиворечивые данные.

Далее мы рассмотрим несколько видов целостности и способы их задания.

6.4.1. Семантическая целостность

Проектируя базу данных, стремятся к тому, чтобы каждая таблица соответствовала некоторому объекту внешнего мира. Существование такого объекта, разумеется, не зависит от базы данных. Если таблица полностью соответствует некоторому объекту, то говорят, что она *обладает семантической целостностью и моделирует (представляет) этот объект*. При этом каждая запись таблицы моделирует некий элемент объекта.

В таблице, обладающей семантической целостностью, должен быть первичный ключ. Первичный ключ (*см. разд. 6.3.5*) — это один столбец или группа столбцов, значения которых должны быть уникальными и определенными.

> ***Примечание***
>
> В языке SQL это ограничение целостности, выражающееся в виде ограничения на значения столбца или группы столбцов, задается ключевыми словами `PRIMARY KEY`.

6.4.2. Доменная целостность

С каждым атрибутом отношения связан домен — множество допустимых значений. При создании таблицы базы данных для каждого столбца кроме имени указывается и тип данных, которые он может содержать. Тип данных ограничивает множество допустимых значений столбца, однако во многих случаях такого ограничения оказывается недостаточно. Например, если в числовом столбце `Возраст` таблицы `Сотрудники` указано значение 1000, то мы не усомнимся, что это ошибочное значение. В той же таблице символьный столбец `Должность` может принимать значения из определенного списка, предусмотренного штатным расписанием организации, а не произвольную комбинацию символов или наименования должностей с орфографическими ошибками. Ограничения на допустимые значения для столбца таблицы предназначены для поддержания доменной целостности.

> **Примечание**
>
> В языке SQL риск нарушить доменную целостность возникает при добавлении и обновлении записей с помощью операторов `INSERT` и `UPDATE` соответственно. Ограничения доменной целостности можно задать при создании таблицы с помощью выражения `CREATE TABLE`, а также предварительно, путем создания домена, применив выражение `CREATE DOMAIN`.

6.4.3. Ссылочная целостность

В хорошо спроектированной базе из нескольких таблиц последние связаны друг с другом. Так, столбец одной таблицы может ссылаться на столбец другой таблицы этой же базы данных. Подобные ссылки представляют собой ограничения ссылочной целостности базы данных и играют важную роль при поддержке ее общей целостности. Вместе с тем, наличие ссылок порождает так называемую *проблему аномалий модификации данных*.

Связи между таблицами обычно несимметричны: одна таблица зависит от другой. Допустим, в базе данных имеются две таблицы:

- `Клиенты` (`Имя_клиента`, `Адрес`, `Телефон`);
- `Продажи` (`ID`, `Товар`, `Количество`, `Цена`, `Стоимость`, `Имя_клиента`).

В таблице `Клиенты` столбец `Имя_клиента` является ключом (`PRIMARY KEY`), т. е. имеет уникальные и определенные значения. В таблице `Продажи` одноименный столбец не является ключом, его значения могут повторяться, т. к. один и тот же клиент мог приобрести несколько товаров. Эта таблица связана с таблицей `Клиенты` по столбцу `Имя_клиента`, т. е. столбец `Имя_клиента` первой таблицы (`Продажи`) ссылается на одноименный столбец второй таблицы (`Клиенты`). Другими словами, данные таблицы находятся в родительско-дочернем отношении: таблица `Клиенты` родительская, `Продажи` — дочерняя. Данная связь между таблицами организуется путем объявления столбца `Имя_клиента` таблицы `Продажи` *внешним ключом* (`FOREIGN KEY`), ссылающимся на первичный ключ в таблице `Клиенты`.

Аномалии модификации могут проявляться разными способами и при различных обстоятельствах, вызывая трудности. Предположим, какой-то клиент перестал вас интересовать (например, он больше не покупает у вас). Если запись о нем удалить из таблицы `Клиенты`, то в дочерней таблице `Продажи` останутся записи, ссылающиеся на отсутствующую запись в родительской таблице. Аналогичная ситуация возникает при попытке добавить в дочернюю таблицу запись, когда в родительскую таблицу еще не было сделано соответствующего добавления.

6.5. Нормализация таблиц

База данных может быть спроектирована хорошо или плохо. Сразу создать хороший проект довольно сложно, и на практике процесс проектирования обычно является итерационным: состав таблиц и их структура модифицируются в несколько этапов, пока не будет получен приемлемый результат.

В плохом проекте часто возникают аномалии модификации данных, устранить которые довольно трудно. Такие аномалии могут возникнуть даже в однотабличной базе данных.

Рассмотрим в качестве примера таблицу `Продажи` (`Клиент`, `Товар`, `Количество`, `Цена`) (рис. 6.12).

Продажи : таблица

Клиент	Товар	Количество	Цена
Иванов	Хлеб	2	24,50р.
Петров	Молоко	3	30,00р.
ОАО "Рога и копыта"	Хвосты	25	120,00р.
ЗАО "111"	Молоко	1	30,00р.
Сидоров	Хлеб	3	24,50р.

Рис. 6.12. Пример таблицы, в которой могут возникнуть аномалии модификации

В таблице, приведенной на рис. 6.12, могут возникнуть аномалии модификации данных. Предположим, было решено удалить из нее запись о клиенте `"ОАО "Рога и копыта""`, поскольку теперь он ничего не приобретает в вашей фирме. Но тогда вы потеряете информацию и цене на товар `"Хвосты"`. Если же вам потребуется добавить запись о каком-нибудь новом товаре, то необходимо будет добавить и сведения о покупателе и количестве этого товара. А если пока такого покупателя нет?

Аномалия модификации, возникшая в рассмотренном примере, обусловлена тем, что данная таблица содержит информацию, относящуюся к различным темам. В ней есть сведения и о том, что приобрели покупатели, и о цене товаров. Лучше разбить ее на две таблицы, посвященные двум различным темам, — `Продажи_клиенты` и `Прайс_лист` (рис. 6.13).

Вообще говоря, любую таблицу, относящуюся к двум или более темам, следует разбить на две или более таблицы. При определенных условиях, о которых говорилось ранее, можно декомпозировать и таблицу, посвященную одной теме. В этом случае каждая из результатных таблиц будет соответствовать какой-то части одной темы. Этот процесс декомпозиции и составляет суть нормализации. Обнаружив в таблице аномалию модификации, мы устраняем ее путем декомпозиции на две или более таблиц, так чтобы они были свободны от аномалий. Однако, производя декомпозицию, мы вынуждены задавать ограничения ссылочной целостности (*см. разд. 6.4.3*).

Продажи_клиенты : таблица

Клиент	Товар	Количество
Иванов	Хлеб	2
Петров	Молоко	3
ОАО "Рога и копыта"	Хвосты	25
ЗАО "111"	Молоко	1
Сидоров	Хлеб	3

Прайс_лист : таблица

Товар	Цена
Хлеб	24,50р.
Молоко	30,00р.
Хвосты	120,00р.

Рис. 6.13. Результат декомпозиции таблицы `Продажи`

Таблицы классифицируются по тем видам аномалий модификации, которым они подвержены. Это так называемые *нормальные формы* таблиц (отношений). В своей статье (1970 г.) И. Кодд определил три источника аномалий и три формы таблиц, свободных от них. В последующие годы он и другие исследователи обнаружили другие виды аномалий и предложили формы таблиц, которые им не подвержены.

Вот список всех специальных нормальных форм:

1. Первая нормальная форма (1НФ)
2. Вторая нормальная форма (2НФ)
3. Третья нормальная форма (3НФ)
4. Нормальная форма Бойса—Кодда (НФБК)
5. Четвертая нормальная форма (4НФ)
6. Пятая нормальная форма (5НФ)

Все нормальные формы вложены друг друга в следующем смысле: таблица в 2НФ является также и таблицей в 1НФ; таблица в 3НФ является таблицей и в 2НФ, и в 1НФ, и т. д. (рис. 6.14).

Каждая из перечисленных нормальных форм могла устранить определенные виды аномалий, и не было гарантии, что с их помощью можно устранить всевозможные аномалии, о которых пока просто не было известно. В 1981 г. Р. Фагин ввел новую нормальную форму, названную *доменно-ключевой* (ДКНФ), и доказал, что таблица в ДКНФ свободна от всех аномалий модификации, и наоборот: таблица, свободная от любых аномалий модификации, находится в ДКНФ. До появления этой важной теоремы теоретики реляционных баз данных должны были продолжать поиск невыявленных видов аномалий и соответствующих им нормальных форм. Теперь же было доказано, что для того, чтобы получить уверенность в отсутствии всех видов аномалий модификации, следует привести таблицу к ДКНФ.

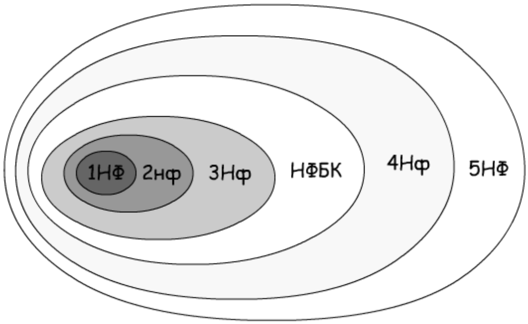


Рис. 6.14. Вложенность нормальных форм таблиц

Далее мы кратко рассмотрим первые три наиболее важные для практики нормальные формы, а также доменно-ключевую нормальную форму.

6.5.1. Первая нормальная форма

Любая таблица, удовлетворяющая определению отношения, находится в 1НФ. Вот основные характеристики таблицы в 1НФ:

- в каждой строке таблицы должны содержаться данные, соответствующие некоторому объекту или его части;
- в каждом столбце таблицы должны находиться данные, соответствующие одному из атрибутов отношения;
- в каждой ячейке таблицы должно находиться только единственное значение;
- у каждого столбца таблицы должно быть уникальное имя;
- все строки (записи) в таблице должны быть различными;
- порядок расположения столбцов и строк в таблице не имеет значения.

Таблица (отношение) в 1НФ свободна от некоторых аномалий, но все же подвержена многим другим. Например, таблица, показанная на рис. 6.12, находится в 1НФ, но, как уже было отмечено, подвержена аномалиям удаления и добавления записей.

6.5.2. Вторая нормальная форма

Каждая таблица в 1НФ должна иметь первичный ключ. Он может состоять из одного или более столбцов (атрибутов). В последнем случае ключ называется составным. Чтобы таблица была в 2НФ, все ее неключевые столбцы должны однозначно определяться всем ключом, т. е. всеми его компонентами, а не некоторыми из них.

Рассмотрим пример отношения `Секции` (`Имя`, `Секция`, `Плата`), показанный на рис. 6.15.

Секции : таблица

Имя	Секция	Плата
Иванов	Футбол	100
Иванов	Волейбол	120
Петров	Лыжи	170
Сидоров	Шахматы	200
Сидоров	Лыжи	170
Федоров	Лыжи	170
Федоров	Волейбол	120

Рис. 6.15. Отношение `Секции` (`Имя`, `Секция`, `Плата`)

Ключом в данном отношении является {`Имя`, `Секция`}, но оно содержит функциональную зависимость `Секция -> Плата`. Аргумент (левая часть) этой зависимости является лишь частью составного ключа. Отношение `Секции` имеет аномалии удаления и добавления. Так, если мы захотим удалить записи с именем `"Иванов"`, то потеряем информацию о стоимости футбольной секции. Мы не сможем добавить запись о новой секции, пока в нее кто-нибудь не запишется. Данных аномалий можно было бы избежать, если бы атрибут `Плата` зависел от всего ключа (однозначно определялся всем ключом).

Отношение `Секции` (`Имя`, `Секция`, `Плата`) в 1НФ можно разбить на два отношения в 2НФ:

- `Секция_члены` (`Имя`, `Секция`);
- `Секция_плата` (`Секция`, `Плата`).

6.5.3. Третья нормальная форма

В отношениях могут быть так называемые транзитивные зависимости, являющиеся источником аномалий модификации данных, против которых 2НФ бессильна. Транзитивная зависимость имеет место тогда, когда один атрибут однозначно определяет второй, второй однозначно определяет третий и т. д.

Рассмотрим в качестве примера отношение `Гости` (`ID_гостя`, `Тип_номера`, `Плата`), представляющее сведения о проживающих в гостинице. Ключом в этом отношении является `ID_гостя`, `Плата` однозначно определяется атрибутом `Тип_номера` (например, люкс, полулюкс и т. д.), т. е. имеется функциональная зависимость `Тип_номера -> Плата`. Поскольку каждый гость проживает только в одном номере определенного типа, в отношении есть и функциональная зависимость `ID_гостя -> Тип_номера`. Таким образом, возникает транзитивная (опосредованная) зависимость `ID_гостя -> Плата`. Так как ключ состоит из единственного атрибута `ID_гостя`, отношение находится в 2НФ.

В рассматриваемом отношении существует аномалия удаления. Удалив запись, мы потеряем не только информацию о каком-то госте (где он проживает), но и сведения о том, сколько стоит номер соответствующего типа.

Чтобы устранить указанную аномалию, следует декомпозировать исходное отношение `Гости` (`ID_гостя`, `Тип_номера`, `Плата`) на два:

- `Проживание` (`ID_гостя`, `Тип_номера`);
- `Тип_платы` (`Тип_номера`, `Плата`).

Эти отношения будут находиться в 2НФ и не содержать транзитивных зависимостей.

Таким образом, отношение находится в 3НФ, если оно находится в 2НФ и не содержит транзитивных зависимостей.

6.5.4. Доменно-ключевая нормальная форма

Если таблица находится в 3НФ, то шансов для возникновения аномалий модификации данных остается довольно мало, но они есть. Чтобы исключить все виды возможных аномалий, таблица должна находиться в доменно-ключевой нормальной форме (ДКНФ).

Понятие ДКНФ довольно просто: отношение находится в ДКНФ, если каждое ограничение, накладываемое на него, является логическим следствием определения доменов и ключей. Термин *ограничение* (*constraint*) здесь намеренно трактуется широко. Р. Фагин определяет ограничение как любое правило, регулирующее возможные статические значения атрибутов, достаточно точное, чтобы можно было проверить его выполнимость. Правила редактирования, ограничения взаимосвязей и структуры отношений, функциональные и многозначные зависимости являются примерами таких ограничений. Отсюда исключаются ограничения, связанные с изменением данных (ограничения, зависящие от времени). Другими словами, отношение находится в ДКНФ, если выполнение ограничений на домены и ключи влечет за собой выполнение всех ограничений.

Однако в настоящее время алгоритм преобразования отношения в ДКНФ неизвестен. Неизвестно также, какие отношения в принципе могут быть приведены к ДКНФ. Поиск и создание отношений в ДКНФ сейчас является искусством, а не наукой. В литературе обычно приводятся только примеры отношений в ДКНФ, которые мы здесь рассматривать не будем.

6.5.5. Денормализация

Чтобы исключить как можно больше аномалий модификации данных, старайтесь как можно больше нормализовать таблицы базы данных. Лучше, если вы доведете их до ДКНФ, хотя на практике это редко происходит. Чаще ограничиваются второй или третьей нормальными формами. Занимаясь нормализацией, вы увеличиваете количество таблиц в базе данных и при определенном их числе эффективность работы может оказаться слишком низкой. Кроме того, формулировать SQL-запросы к базе данных тем легче, чем меньше в ней таблиц. Так что на любом этапе своего развития база данных может быть в какой-то степени денормализованной.

Глава 7

Классификация

Ученый должен систематизировать; наука строится из фактов, как дом из кирпичей; но простое собрание фактов столь же мало является наукой, как куча кирпичей — домом.

А. Пуанкаре

…Значение какого-нибудь тела, следовательно, также и название его, не зависит уже от его состава, а обусловлено скорее его положением в том ряду, к которому оно принадлежит.

Ф. Энгельс

7.1. Что такое классификация?

Классификация, как странно это ни показалось бы, является неотъемлемой частью нашей повседневной жизни. Без классифицирования было бы невозможным существование людей и животных в безгранично многообразной окружающей среде. Если бы они не были способны собирать сходные раздражители в группы, для которых нужна положительная или отрицательная реакция, то они были бы слишком плохо приспособлены, чтобы выжить. Трудно найти область деятельности, где бы не применялось классифицирование — группировка различимых индивидуальных объектов в классы и соотнесение каждого индивидуума с содержащим его классом. Мы настолько часто и почти везде занимаемся классифицированием, что редко задумываемся о законах, которым оно подчинено. Впрочем, глубоко задумываться об обычном — дело ученых, в частности, математиков. Обычные люди больше задумываются о чем-то особенном, исключительном.

Чтобы классифицировать, надо иметь *классификационную схему*, т. е. систему классов, к которым любой предмет из данной совокупности мог быть отнесен.

Таким образом, классификационная схема позволяет увидеть с некоторой точки зрения устройство предметной области и, далее, описать любой ее элемент, выразив его название через имена тех классов, которым он принадлежит. Например, мы наблюдаем конкретного осла, которого кличут Иа. Этот ослик принадлежит, допустим, классу млекопитающих и подклассу непарнокопытных (в отличие от парнокопытных, у которых копыта раздвоены). Возможно, с некоторой точки зрения, этой информации нам достаточно для его классифицирования. Но тогда мы вынуждены согласиться, что не только ослики, но и лошади, а также мулы принадлежат тому же классу. Ладно, допустим и это. А вдруг туда же попадут кентавры и другие синтетические чудовища с нераздвоенными копытами? Разумеется, мы попытаемся придумать еще какой-нибудь отличительный признак, чтобы в нужный нам класс не попало что-нибудь постороннее. Например, мы можем просто сказать, что данный объект это не какой-нибудь ослик, а именно Иа. Иначе говоря, мы присваиваем ему индивидуальное имя, чтобы он не смешался с другими индивидуумами в общем с ним классе. Что ж, и такое бывает. Наряду с обобщением индивидуальностей, образованием общего класса, мы иногда поступаем и наоборот, выделяя из сонма, расплывчатого и маловыразительного сообщества, нечто особенное — индивидуумов или, по крайней мере, виды.

Каким образом происходит соотнесение элементов предметной области с классификационной схемой? Ответ очевиден — по признакам, т. е. параметрам объекта, доступным непосредственному или опосредованному наблюдению. Непосредственно наблюдаемые признаки это те, которые мы воспринимаем нашими органами чувств — зрением, обонянием, осязанием. Однако мы можем усилить наши ощущения, используя измерительные приборы, такие как микроскоп, телескоп, амперметр, манометр и т. п., переходя к косвенным ощущениям. При этом возникает ряд проблем. Например, не все из наблюдаемых признаков обязательно оказываются существенными для того, чтобы подлежащий классифицированию объект мог быть соотнесен с классификационной схемой. И наоборот, может случиться, что реально наблюдаемых признаков будет недостаточно для этой цели. В первом случае возникает задача поиска существенных признаков, во втором — недостающих. Итак, классифицирование это анализ признаков с тем, чтобы отнести наблюдаемый объект к классу классификационной схемы.

Рассматривая классифицирование, мы предполагаем, что классификационная схема (набор классов) уже задана. Но как ее построить? Если учесть характер классифицирования путем анализа признаков, то естественно предположить, что вид классификационной схемы зависит от того, какие признаки положены в основу… чего? Классификационной схемы или классифицирования?

Допустим сначала, что рассматриваются признаки, положенные в основу классификационной схемы. Тогда они определяют классы объектов: объекты, обладающие общими признаками, принадлежат одному классу, а различные классы отличаются друг от друга различиями в признаках. Рассуждая таким образом, мы как бы описываем уже имеющуюся классификационную схему, причем так, чтобы не возникло особых трудностей при классифицировании. К сожалению, мы при этом ни на шаг не приближаемся к пониманию процесса построения самой классификационной схемы. На практике в большинстве случаев создание хорошей в некотором смысле классификационной схемы является творческим актом, хотя после его совершения почти всегда находятся мотивы и объяснения происхождения его конечного результата, как это бывает с художественными произведениями — кинофильмами, спектаклями и романами.

Теперь допустим, что рассматриваются признаки, положенные в основу классифицирования. Но и этот случай, как и предыдущий, не выводит нас на верный путь. Действительно, алгоритм классифицирования нужен для того, чтобы применять классификационную схему для описания объектов предметной области. Если в качестве отправного пункта взять такой алгоритм, то тогда следует признать, что построение соответствующей ему классификационной схемы необходимо лишь для того, чтобы узнать, что можно делать с его помощью. Но если мы этого не знали раньше, то каким образом мог появиться такой алгоритм классифицирования? Получается порочный круг. Выход из него возможен, если допустить, что классификационная схема определяется не признаками, а отношениями между объектами, которые мы собираемся классифицировать. А роль признаков при этом сводится лишь к тому, чтобы с их помощью можно было осуществить соотнесение объектов с классами классификационной схемы, т. е. выполнить классифицирование.

Как уже отмечалось, классификационная схема отражает устройство предметной области. А какого вида устройство она отражает и для чего эта информация может быть использована на практике?

Из опыта известно, что хорошие классификационные схемы получаются при условии, что предметная область достаточно хорошо изучена и, наоборот, при слабой ее изученности классификационные схемы, как правило, оказываются плохими. Отсюда можно заключить, что классификационная деятельность бессмысленна на этапе изучения предметной области, поскольку хорошая классификационная схема может получиться только по его окончании. Но, с другой стороны, когда изучение завершено, классификационная деятельность оказывается бесплодной, не дающей ничего нового. Данную точку зрения наиболее отчетливо выразил А. Раппопорт, сказав: "Я прекрасно осознаю различие между решением важной задачи и формалистическими бесплодными упражнениями в систематизации… Формалист-систематизатор бесплоден потому, что он вовсе не решает никакой задачи. Он уже знает все ответы, так что его деятельность сводится лишь к тому, что этими ответами он морочит голову людям, перед которыми стоят подлинно сложные задачи".

Однако мы не будем столь категоричны, чтобы осуждать классификационную деятельность в целом. Попытаемся разобраться в сути дела и выйти из порочного круга предыдущих рассуждений. В мире известны выдающиеся классификационные схемы, как естественные, так и искусственные.

Типичным примером так называемой искусственной классификации является систематизация растений, разработанная шведским натуралистом Карлом Линнеем. Ее искусственность обычно объясняют тем, что выбор несущественных признаков (количество и способ прикрепления тычинок к цветку) привел к тому, что родственные растения (например, злаки) попали в разные классы, а совершенно несходные растения (например, дуб и один вид осоки) — в один и тот же класс.

В данном рассуждении, во-первых, с очевидностью предполагается известным некоторое отношение (например, родства) между классифицируемыми объектами. Однако оно остается в тени, пока не возникает вопрос об оценке естественности данной систематизационной схемы. Во-вторых, схема Линнея явным образом отражает не отношение родства между растениями, а отношение между ними и выбранными тем или иным образом признаками.

Если систематизатор сразу выбрал, как говорят, существенные признаки и это произошло не случайно, то, видимо, он уже имел заранее классификационную схему, соответствующую тому важному отношению, которое используется в критерии оценки существенности признаков или, другими словами, естественности разрабатываемой систематизационной схемы. Впрочем, эта критериальная или, иначе, априорная классификационная схема могла и не осознаваться систематизатором, группирующим объекты по выбранным признакам.

Линней выбрал признаки, исходя из удобства классифицирования растений. Во-первых, эти признаки хорошо наблюдаемы. Во-вторых, они обеспечивают группировку растений на относительно небольшое количество классов (всего 24 класса). Так что у него получилось весьма компактное описание огромного растительного мира.

Мысль о компактном представлении знаний в виде классификационных схем довольно долго затмевала собой их прогностическое назначение, т. е. способность предсказывать поведение объектов или наступление тех или иных событий. Рассмотрим в качестве примера знаменитую классификацию химических элементов Д. И. Менделеева, которую он представил в виде таблицы, открыв так называемый периодический закон. В чем заключается ее значение? На момент своего создания таблица не была целиком заполнена, но в пустые клетки могли попасть элементы не с произвольными свойствами. Свойства еще не открытого элемента зависели от свойств элементов, уже занявших свое место в соседних клетках. Таким образом, классификационная схема Менделеева позволила предсказать существование новых химических элементов. В настоящее время эта таблица практически заполнена, но утратила ли она свое прогностическое значение? Нет, поскольку эта таблица позволяет предсказывать взаимодействие между элементами. Например, из нее следует, что щелочные элементы взаимодействуют с галогенами, а инертные газы не вступают в реакцию с редкоземельными металлами, и т. д.

Но каким образом Менделеев мог построить свою классификационную схему? Он мог сделать это на основе наблюдения различных химических реакций, а также протоколов (описаний) химических экспериментов. Его заслуга состояла в удачном выборе признака (атомный вес) на основе той гипотезы, что информация об отношениях между объектами должна быть как-то запечатлена в самих объектах посредством их признаков (наблюдаемых параметров).

Принятие указанной гипотезы в качестве рабочей имеет следующие основания. Допустим, мы хотим предсказывать реакции или отношения, в которые могут вступить объекты из некоторого множества. Предположим далее, что у нас есть протокол всех реакций, наблюденных ранее. Тогда мы можем поступить двумя способами:

1. Просто осуществить реакцию и полученный результат считать прогнозом.
2. Обратиться к протокольным записям и найти в них описание интересующей нас реакции.

В первом случае необходимы затраты на эксперимент, а во втором — на поиск нужных данных в протоколе ранее проведенных экспериментов. Причем в обоих случаях требуется, чтобы время упреждения прогноза было не меньше некоторой заданной величины. Другое дело, если бы результаты будущей реакции были запечатлены некоторым образом в самих объектах. Тогда не нужно было бы проводить дорогостоящие эксперименты и поиски в "бездонных" протоколах. Так возникает вопрос о признаках. Однако основным мотивом его появления было уменьшение трудностей классифицирования, в то время как мотивом самого классифицирования была необходимость предсказания результатов взаимодействия объектов. Для того чтобы классифицирование обеспечивало предсказание, соответствующая ему классификационная схема должна содержать информацию о поведении объектов.

Рассмотрим еще один пример. Представим себе исследователя-первопроходца, который впервые попал в уссурийскую тайгу и повстречал там тигра. Он удивлен, поскольку знает, что тигры обычно встречаются в "жарких странах". Итак, в отношении "тигры—регионы" появились новые данные. В сущности, сформирован новый класс. Вопрос только в том, несут ли тигры на себе сведения о своем месте обитания? Именно поэтому у исследователя возникает потребность пристально рассматривать встречающиеся особи в надежде найти необходимые признаки. Наконец, такие признаки найдены. Если мы теперь придем в зоопарк и спросим, почему этот тигр — уссурийский, нам скажут, например, что он более крупный и шерсть у него гуще, и т. п.

Итак, получается, что сначала тигр был "уссурийским", потому что он был уссурийским. Кроме того, он имел особые признаки именно потому, что жил в уссурийской тайге. А после создания алгоритма классифицирования он стал уссурийским по другим причинам и не перестает быть им даже в зоопарке.

Подобная логика довольно часто проявляется в сознательной деятельности людей. Для краткости такую эволюцию причин и следствий назовем *законом инверсии импликации*. В реальной деятельности людей инверсия импликации нередко идет еще дальше. Так, мы обычно считаем, что если обеспечить объекту некоторые признаки (чтобы он попадал в нужный класс), то он станет входить в нужные отношения. Интересно, что подобные действия нередко приводят к успеху. Закон инверсии импликации не является каким-то дефектом мыслительной деятельности. Напротив, это мощное средство "мышления по аналогии", позволяющее прогнозировать поведение объектов, с которыми классификатор ранее не встречался.

Далее мы попробуем построить аксиоматическую теорию классификации или хотя бы ее набросок, чтобы получить возможность исследовать классификационную деятельность математическими средствами.

7.2. Аксиомы классификации и основные следствия

Классификатор каким-то образом формирует классы, состоящие из объектов некоторой совокупности. В результате у него получается классификационная схема — упорядоченное по отношению включения множество всех классов. В таком понимании классификационную схему обычно называют *таксономией* (греч. taxis — расположение по порядку, nomos — закон), ее элементы (классы объектов) называют *таксонами*, а множество всех классифицируемых объектов — *таксономическим универсумом*.

Нас интересует операция, с помощью которой классификатор творит таксоны. На данном этапе нам не важен алгоритм действий, выполнение которых приводит к осуществлению этой операции. Пока важны только ее свойства, которые мы зафиксируем в виде некоторых аксиом. Хорошо, если бы таких свойств оказалось немного и в то же время они обеспечивали развитие достаточно содержательной теории.

Начнем с обозначений и неформальных рассуждений, которые и положим в основу будущих аксиом. Обозначим через T_0 таксономический универсум, т. е. множество всех объектов, подлежащих классифицированию, а через θ — операцию образования таксонов, которую будем называть *операцией таксонообразования*. Мы считаем, что классификатор формирует таксоны, беря за основу некоторые, вообще говоря произвольные, выборки X объектов из таксономического универсума T_0 (т. е. $X \subseteq T_0$) и применяя к ним операцию таксонообразования $\theta : X \to \theta(X)$. Результат $\theta(X)$ применения этой операции к множеству X является подмножеством таксономического универсума, так что $\theta(X) \subseteq T_0$. Множество $\theta(X)$ назовем таксоном, а любое его подмножество $Y \subseteq \theta(X)$ — множеством примеров данного таксона. При этом множество X, на базе которого построен таксон $\theta(X)$, естественно назвать базовым множеством примеров.

Проще говоря, классификатор начинает рассматривать некую выборку объектов как множество примеров будущего таксона и затем каким-то образом на ее основе формирует таксон. На практике исходные совокупности объектов выбираются, видимо, не случайно, но мы не будем накладывать на их выбор никаких ограничений. Так что выборка примеров может быть любым подмножеством таксономического универсума.

Теперь дадим определение операции таксонообразования, постулировав для нее всего два свойства.

Определение 7.1

Операцией таксонообразования называется операция $\theta : X \to \theta(X)$, определенная на подмножествах таксономического универсума T_0, обладающая следующими свойствами:

$$X \subseteq \theta(X); \tag{A1}$$

$$\text{если } Y \subseteq \theta(X), \text{ то } \theta(Y) \subseteq \theta(Y). \tag{A2}$$

Множество всех таксонов называется *таксономией*.

С одной стороны, интуитивно ясно, что множество примеров, на базе которого мы строим некоторое понятие (таксон), является частью объема этого понятия (свойство А1). С другой стороны, множество примеров, на котором построено некоторое понятие, не может быть базовым для более широкого или несравнимого по объему понятия (свойство А2). Так, если мы знаем, что какое-то множество яблок входит в объем понятия "фрукты", то было бы неестественно полагать, что на базе этого множества можно образовать более широкое понятие "плоды" или несравнимое понятие "овощи".

Операцию таксонообразования естественно интерпретировать как операцию обобщения выборочного множества примеров синтезируемого таксона. Так, классификатор выделяет в обозреваемом множестве примеров некоторую общность и добавляет к нему другие объекты, обладающие той же общностью, что и исходные. При этом из части таксона указанным выше образом можно построить только подтаксон (в смысле нестрогого включения $\subseteq$).

Теперь рассмотрим важнейшие следствия из аксиом A1 и A2.

Из свойства A2 и транзитивности отношения включения следует

Теорема 7.1. Если $X_1 \subseteq \theta(X_2)$ и $X_2 \subseteq \theta(X_3)$, то $X_1 \subseteq \theta(X_3)$.

Из данной теоремы, свойства A1 и транзитивности отношения включения вытекает следующая теорема.

Теорема 7.2. $\theta(X) = \theta(\theta(X))$ — из таксона нельзя построить другой таксон путем многократного применения операции таксонообразования. Иначе говоря, таксон устойчив относительно операции таксонообразования.

Из свойств A1 и A2 и транзитивности включения вытекает еще одна теорема.

Теорема 7.3. Если $X \subseteq Y$, то $\theta(X) \subseteq \theta(Y)$ — операция таксонообразования сохраняет порядок включения (является изотонной).

Нетрудно доказать и справедливость следующего утверждения.

Теорема 7.4. $\theta(X) \bigcap \theta(Y) = \theta(\theta(X) \bigcap \theta(Y))$ — пересечение таксонов является таксоном.

Некоторые свойства операции таксонообразования показаны на рис. 7.1.

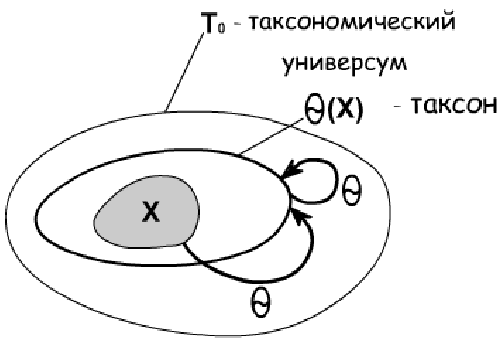


Рис. 7.1. Как работает операция таксонообразования

В 1974 г., примерно за 10 лет до предложенного мной определения таксономии (см. после определения 7.1), Ю. А. Шрейдер предложил такое ее аксиоматическое определение:

1. Множество всех таксонов конечно.
2. Таксономический универсум является таксоном.
3. Пересечение таксонов есть таксон.

Как соотносится наше определение 7.1 с данным? Нетрудно заметить, что для выполнения первой аксиомы нам достаточно предположить конечность таксономического универсума или же из определения Шрейдера удалить первое требование. Вторая аксиома непосредственно следует из свойства A1 ($T_0 \subseteq \theta(T_0)$) и того, что $\theta(T_0) \subseteq T_0$. Третья аксиома выводится из A1 и A2 как теорема 7.4.

Оказывается, некоторые системы свойств операции таксоноообразования равносильны друг другу и, следовательно, мы могли бы положить в основу нашей теории любую из этих систем.

Обозначим свойства операции θ, сформулированные в теоремах 7.1—7.3, как T1, T2 и T3 соответственно.

Теорема 7.5. Системы свойств <A1,A2>, <A1, T1, T3> и <A1, T2, T3> равносильны.

Доказательство. Достаточно показать, что из второй и третьей систем следует свойство A2.

Из второй системы (свойства A1, T1) вытекает T2. С учетом T2 путем подстановки в T2 равенства $Y = \theta(X_1)$ получаем свойство A2. Но T2 уже содержится в третьей системе свойств, следовательно, и из нее выводится A2.

В математике операция, определяемая третьей системой свойств <A1, T2, T3>, известна как *операция замыкания множеств*. Отсюда с учетом теоремы 7.5 можно заключить, что операция таксонообразования, определяемая свойствами A1 и A2, является операцией замыкания на подмножествах таксономического универсума. При этом таксонами являются множества, равные своим замыканиям, т. е. такие $X \subseteq T_0$, что $X = \theta(X)$.

Операция замыкания широко используется во многих разделах алгебры и общей топологии. Так что перед нами открывается возможность изучения классификации уже готовыми математическими средствами. В частности уже сейчас мы можем ответить на вопрос — что же представляет собой таксономия как множество всех таксонов в структурном отношении? Прежде всего, это частично упорядоченное по отношению включения $\subseteq$ множество, в котором есть наибольший и наименьший элементы. Наибольшим элементом является таксономический универсум, а наименьшим — пересечение всех таксонов. Такие структуры называются *полными решетками*, в которых операция выделения наибольшей нижней грани является теоретико-множественным пересечением.

Вот список всех основных свойств операции таксонообразования (постулируемых и выводимых), которые мы рассмотрели в данном разделе (для краткости вместо союза "если …, то …" мы использовали стрелку =>):

- аксиомы:
  - $X \subseteq \theta(X)$ — экстенсивность;
  - $Y \subseteq \theta(X)$ => $\theta(Y) \subseteq \theta(Y)$;
- следствия:
  - $\theta(X) = \theta(\theta(X))$ — идемпотентность (теорема 7.2);
  - $X \subseteq Y$ => $\theta(X) \subseteq \theta(Y)$ — изотонность (теорема 7.3);
  - $\theta(X) \bigcap \theta(Y) = \theta(\theta(X) \bigcap \theta(Y))$ — пересечение таксонов есть таксон (теорема 7.4);
  - $X_1 \subseteq \theta(X_2)$ и $X_2 \subseteq \theta(X_3)$ => $X_1 \subseteq \theta(X_3)$;
  - $\theta(T_0) = T_0$ — таксономический универсум есть таксон.

7.3. Реляционный подход к классификации

Как конкретно может работать операция таксонообразования, обладающая рассмотренными выше свойствами? В частности, каким образом расширяется базовое множество примеров (свойство экстенсивности A1)? Естественно думать, что осуществляя таксонообразование, мы рассматриваем объекты базового множества примеров в некоторых отношениях между ними и/или с объектами других множеств. Если допустить данную гипотезу, то исходной информацией для синтеза классификационной схемы (таксономии) являются отношения между объектами, и операция таксонообразования должна быть как-то выражена через них.

В *главе 5*, рассматривая бинарные отношения, мы использовали понятия классов и ко-классов, порожденных операциями Δ и ∇. Эти операции определялись довольно просто как сечения 2-го рода (определение 5.9) отношения $R \subseteq A \times B$ через подмножества $X \subseteq A$ и $Y \subseteq B$. Там же мы выяснили, что операции $X \to X^{\Delta\nabla}$ и $Y \to Y^{\nabla\Delta}$ обладают всеми свойствами операции замыкания:

- $X \subseteq X^{\Delta\nabla}$
- $X^{\Delta\nabla} = (X^{\Delta\nabla})^{\Delta\nabla}$
- если $X' \subseteq X$, то $X'^{\Delta\nabla} \subseteq X^{\Delta\nabla}$

Аналогичные свойства имеет двойственная операция $\nabla\Delta$.

По теореме 7.5 любая операция замыкания равносильна операции таксонообразования, заданной определением 7.1. Таким образом, классы и ко-классы произвольного бинарного отношения, порожденные операциями $\Delta\nabla$ и $\nabla\Delta$, являются таксонами и ко-таксонами. Иначе говоря, операции $\Delta\nabla$ и $\nabla\Delta$ представляют собой интерпретации аксиоматически определенной операции таксонообразования θ, а примеры таксономий — это структуры множеств классов и ко-классов, приведенные в *главе 5*. Особенностью классификационной схемы, построенной на основе отношения, является то, что она представляется в виде двух взаимно однозначно связанных антиизоморфных структур — таксономии и ко-таксономии. Однако если рассматривается отношение на одном и том же множестве объектов, то таксономию и ко-таксономию можно совместить, объединяя множества таксонов и ко-таксонов, а также добавляя всевозможные их попарные пересечения.

Примечание

Тот факт, что операция $\Delta\nabla$ удовлетворяет аксиомам определения 7.1, означает, что данная система аксиом непротиворечива.

7.4. Классифицирование

Как мы уже говорили, классифицирование — это соотнесение объектов с классами классификационной схемы (с таксонами таксономии). В общем случае объект может принадлежать одновременно нескольким таксонам. По крайней мере он принадлежит наибольшему из них — таксономическому универсуму. К какому из нескольких таксонов следует отнести данный объект?

Естественно считать, что коль скоро классификатор может построить всю таксономию, он тем более успешно может решить более скромную задачу классифицирования, причем теми же средствами, которые использовались при синтезе таксономии. Поэтому в основу правила классифицирования разумно положить принцип, согласно которому всякий объект $x \in T_0$ из таксономического универсума следует отнести к таксону $\theta(\{x\})$, т. е. к таксону, который может быть построен на базе одноэлементного множества $\{x\}$ с помощью операции таксонообразования θ. Разумеется, эта же операция использовалась при построении всей таксономии, а теперь она применяется для классифицирования. Нетрудно доказать, что $\theta(\{x\})$ является наименьшим таксоном, содержащим объект x.

Множество всех таксонов вида $\theta(\{x\})$ образует покрытие, но, в общем случае, не разбиение таксономического универсума. Иначе говоря, различные таксоны вида $\theta(\{x\})$ могут пересекаться друг с другом. Однако это пересечение не произвольно. Пусть имеются два таксона $\theta(\{x\})$ и $\theta(\{y\})$, тогда возможны только следующие два взаимно исключающих случая:

- $\theta(\{x\})\bigcap\theta(\{y\})=Ш$ — наименьшие для объектов x и y таксоны не пересекаются;
- либо $\theta(\{x\})\subseteq\theta(\{y\})$, либо $\theta(\{y\})\subseteq\theta(\{x\})$ — наименьшие для объектов x и y таксоны либо вложены друг в друга, либо совпадают.

Обозначим через Tax множество всех таксонов данной таксономии и рассмотрим естественное соответствие $\varphi: T_0 \to Tax,$ которое сопоставляет любому объекту $x \in T_0$ множество всех тех таксонов из $Tax,$ в которые входит данный объект. Тогда мы можем ввести определение *таксономического вида объектов* — понятия, широко используемого в классификационной деятельности.

Определение 7.2

Два объекта x, y принадлежат одному и тому же таксономическому виду, если $\varphi(x) = \varphi(y)$.

Другими словами, два объекта принадлежат одному и тому же виду, если они принадлежат одним и тем же таксонам. Очевидно, что отношение "принадлежать одному виду" рефлексивно, симметрично и транзитивно. Следовательно, оно является эквивалентностью, а виды есть классы этой эквивалентности. Особенностью видов в рамках рассматриваемой теории классификации является то, что вид не обязательно является таксоном. Если таксономия построена по бинарному отношению, то виды — это классы ядра (ко-ядра) данного бинарного отношения.

На рис. 7.2 в качестве примера показано бинарное отношение "поедания" на множестве рыб {щука, окунь, карась, плотва}, которое показывает, кто кого поедает. Хищники щука и окунь, бывает, поедают себе подобных, а карась и плотва — нет. Для данного отношения можно построить таксономию и ко-таксономию, как множества классов и ко-классов. Поскольку бинарное отношение задано на одном множестве, таксономию и ко-таксономию можно объединить, добавив попарные пересечения классов и ко-классов. На рис. 7.2 показана объединенная таксономия, в которой таксон {окунь} является пересечением класса {щука, окунь} и ко-класса {окунь, карась, плотва}. Стрелки между таксонами указывают взаимно однозначное соответствие Δ между классами и ко-классами.

Таксоны {щука} и {окунь} являются одновременно и видами. Карась и плотва образуют еще один вид, не являющийся таксоном; этот вид включается как подмножество в таксон {окунь, карась, плотва}. Таким образом, алгоритм классифицирования для данной таксономии должен решать задачу отнесения рыб к одному из трех непересекающихся таксономических видов, а в качестве окончательного результата указывать наименьший таксон, содержащий этот вид.

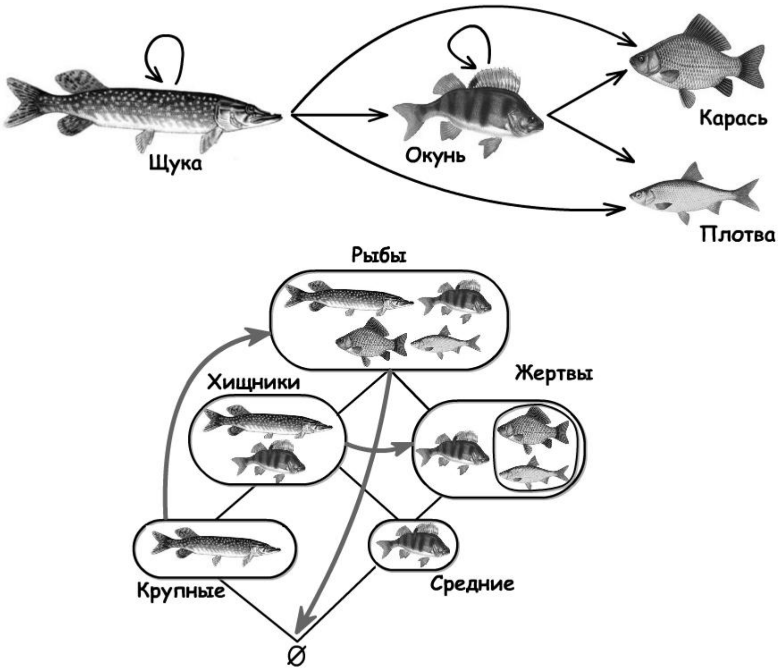


Рис. 7.2. Отношение "поедания" на множестве рыб и соответствующая ему таксономия

Таксоны в рассматриваемом случае легко интерпретируются как "хищники", "жертвы", "крупные хищники" и "средние хищники". Все рыбы таксономического универсума и пустое множество Ø также образуют таксоны. Прогностическая функция таксономии отображается стрелками между таксонами. Например, пусть мы встретили щуку. Алгоритм классифицирования отнес ее к таксону крупных хищников, стрелка из которого ведет к таксону всех рыб. Следовательно, щука поедает всех рыб. Отсюда мы делаем прогноз о том, что и повстречавшаяся нам щука может съесть любую рыбу из таксономического универсума. С окунем мы поступаем следующим образом. Он классифицируется в таксон, не связанный непосредственно с другими таксонами стрелками. Однако он является подтаксоном двух таксонов — хищников и жертв, между которыми имеется стрелка. Следовательно, окунь может выступить в роли хищника, а также стать жертвой. Анализ содержимого таксонов показывает, что он может съесть карася, плотву, а также своего сородича и, вместе с тем, может быть проглочен щукой или другим окунем. Кроме того мы видим, что общей жертвы одновременно для всех рыб нет, т. к. соответствующий таксон пуст.

7.5. Именование таксонов

Предположим, что в нашем распоряжении имеется уже готовая таксономия, представленная для наглядности в виде системы вложенных друг в друга ящиков, заполненных объектами. Тогда мы могли бы, используя подходящий алфавит или словарь, составить надписи на этикетках и прикрепить их к соответствующим ящикам. Если бы каждый объект, подлежащий классифицированию, имел при себе экземпляр одной из данных этикеток, то классифицирование было бы более чем простой задачей. Вся проблема сводилась бы к тому, чтобы быстро найти нужный ящик. Так, например, не представляет особого труда классифицирование книг в системе универсальной десятичной классификации (УДК) или в международной системе, использующей индексы ISBN, если на самих книгах уже напечатан их индекс. Но как быть с книгами, которым индекс еще не присвоен? Поставим данный вопрос шире: если объект не "переносит" с собой наименований таксонов, которым он принадлежит, то что тогда он "переносит" и каким образом это переносимое связано с именами таксонов?

Итак, таксономия представляет собой конструкцию, элементы которой (таксоны) могут быть поименованы. Имена представляются нам как синтетические конструкции, составленные из атомарных элементов — символов некоторого алфавита или же слов некоторого словаря. Далее, мы предполагаем, что сами объекты устроены в некотором смысле подобно именам, которыми они могут быть обозначены. На первый взгляд, это чересчур сильное допущение. Однако оно основано на естественной мысли о том, что понятие об объекте должно как-то воплощаться в нем, а сами объекты должны быть устроены так, чтобы, наблюдая их, можно было вывести (абстрагировать) соответствующее им понятие.

Исходя из указанного выше предположения, примем в качестве рабочей гипотезы, что объекты можно представить себе в виде наборов признаков, через которые они являются в наших ощущениях. Коль скоро имена таксонов и входящих в них объектов состоят из частей, то при определении соответствия между частями имен и признаками в принципе возможно выделение групп признаков, которые можно было бы назвать частями объектов. Например, пусть перед нами две книги. Одна из них уже имеет индекс УДК, а второй его только еще предстоит присвоить. Разумно присвоить второй книге индекс первой, если в их содержании найдутся похожие тексты, ключевые слова и т. п. Так часто и поступают на практике. Положение не изменится, если мы будем знать ключевые слова, соответствующие разрядам индексов УДК и не заимствовать индексы у других книг, а конструировать их по ключевым словам, найденным в данной книге.

Во многих случаях классификационная схема формируется сначала как структура имен классов, а не самих классов (заданных объемом или/и свойствами). При этом элементы имен интерпретируются как признаки, которые можно либо непосредственно наблюдать в объектах, либо получать путем обобщения непосредственно наблюдаемых признаков. В таких случаях обычно стремятся построить минимальное множество так называемых существенных признаков, чему, однако, мешает опасение, что одним и тем же именем будут названы существенно различные объекты. По данному поводу уместно вспомнить легенду о классификационной деятельности Платона, когда он пытался дать краткое, но емкое определение человека *(см. разд. 2.3.1)*.

Выбор удачных (существенных) признаков имеет большое значение при создании структуры имен классов (таксонов). При этом необходимость обеспечения процесса классифицирования реальных объектов требует вычленять в них те части, которые соответствуют элементам имен таксонов.

В 1970-х годах наш соотечественник С. В. Мейен ввел в теорию классификации термин *"мерономия"* (греч. meros — часть; nomos — закон) для обозначения классификационной деятельности, связанной с вычленением частей, образующих структуру (архетип) объекта. В дальнейшем концепция мерономии была развита в работах С. В. Мейена и Ю. А. Шрейдера. Возможно, более удачным для обозначения структуры имен таксонов был бы термин "номономия" (лат. nomen — имя и греч. nomos — закон). Однако поскольку структура имен таксонов тесно связана со структурой признаков объектов (их реально наблюдаемых частей), назовем ее мерономией.

Мерономию (структуру имен таксонов) для заданной таксономии можно построить по следующим правилам. Пусть C — словарь, тогда:

1. Каждому таксону, не являющемуся пересечением других таксонов, сопоставляется слово из C так, чтобы всем таксонам были сопоставлены различные слова.
2. Каждому таксону сопоставляется множество слов из C, которые на первом этапе были сопоставлены таксонам, включающим в себя данный таксон.

На рис. 7.3 показаны одновременно таксономия и соответствующая ей мерономия, построенная по указанным правилам. Имена таксонов представлены в виде множеств слов из словаря $C = \{c_0, c_1, c_2, c_3\}$.

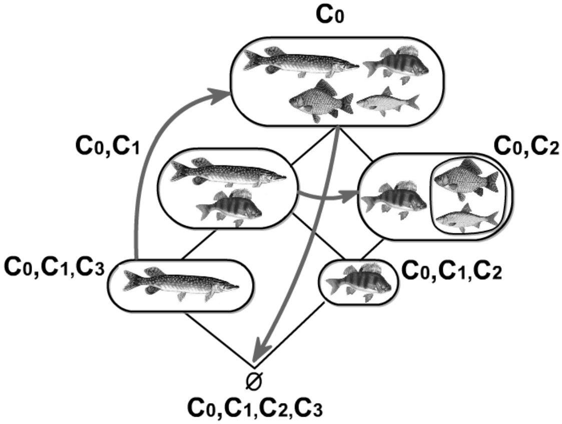


Рис. 7.3. Merономия — структура имен таксонов

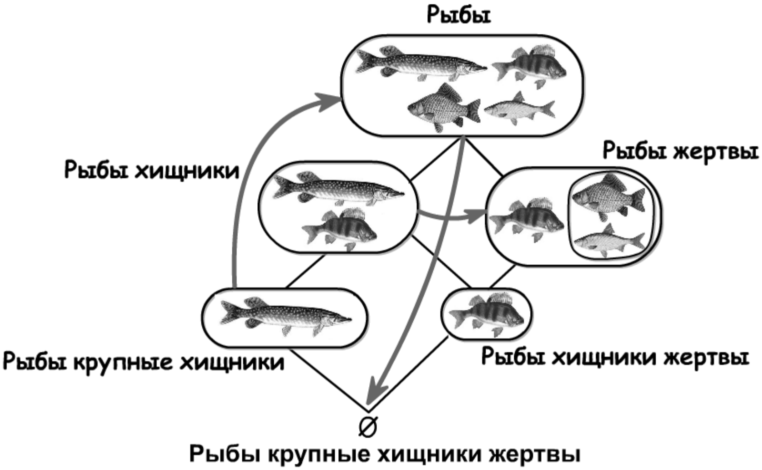


Рис. 7.4. Мерономия с составляющими имен таксонов, взятыми из русского словаря

В данном примере вместо символов $c_0, c_1, ...$ можно было бы использовать, например, слова "рыбы", "хищники", "жертвы" и "крупные", тогда мерономия приняла бы менее абстрактный вид, показанный на рис. 7.4.

Мерономия отражает языковой аспект классификации. Если воспользоваться этой аналогией, то имя таксона, из которого выходит стрелка, можно интерпретировать как группу подлежащего (существительные и прилагательные), а имя таксона, в который входит стрелка, — как группу сказуемого. При этом само соответствие между таксонами (стрелка) интерпретируется как собственно сказуемое (глагол). Так, в мерономии на рис. 7.4 можно увидеть такие предложения, как, например, "хищные рыбы поедают рыб-жертв" ($\{щука, окунь\} ® \{окунь, карась, плотва\}$) или "рыбы, являющиеся хищными жертвами, поедают рыб-жертв и поедаются хищными рыбами" (это, очевидно относится к таксонам $\{окунь\}, \{окунь, карась, плотва\}, \{щука, окунь\}$).

Мерономия устроена таким образом, что имя каждого таксона может быть определено по классическому правилу: через ближайший род и видовое отличие, т. е. имя данного таксона определяется через имя надтаксона и дополнительные символы, выделяющие его из множества всех подтаксонов. Так например, таксон $\{щука\}$ определяется через ближайший род "рыбы хищники" и видовое отличие "крупные".

В ряде случаев, например, при классификации животных, стремясь сблизить между собой классификационную схему и классифицирование, стараются составлять имена таксонов из имен признаков, которые можно реально наблюдать в самих объектах. Отсюда появляются такие названия таксонов, как "хордовые", "парнокопытные" или "хлористый натрий" (вместо обиходного "поваренная соль").

Представим себе работника, который по роду своей службы имеет дело с объектами, входящими в ограниченное число таксонов некоторой достаточно большой таксономии. Хорошо знакомый со своей предметной областью, он обыкновенно не будет использовать имена старших (по включению) таксонов и, как часто бывает на практике, может даже заменить эти длинные мерономические имена более короткими именами нарицательными. Правда, возможно, при этом ему придется расширить используемый словарь. Например, таксон "хордовые черепные млекопитающие парнокопытные" обычно называют просто "коровы".

Теперь предположим, что кругозор нашего работника стал расширяться при неизменной емкости хранилищ информации. Очевидно, что прежний способ кодирования (именования) таксонов с некоторого момента станет непригодным. Тогда работник попросту будет вынужден группировать прежние имена таксонов, а сгруппировав их, заменять новыми. Пусть эти имена будут длинными, зато используемый словарь можно сократить. Однако, обобщая таким образом исходную информацию, необходимо побеспокоиться о том, чтобы сохранить на достаточном уровне связь между именами новых таксонов и объектами, которые этими именами обозначаются. При этом желательно также, а во многих случаях и необходимо, чтобы образуемые таксоны были обязаны своим существованием не только стремлению классификатора сэкономить свои ресурсы, необходимые для представления знаний, но и объективным отношениям между объектами.

Рассмотрим мерономию в структурном отношении. Нетрудно заметить, что частично упорядоченное по включению множество всех имен таксонов (т. е. мерономия) антиизоморфно таксономии:

1. Отображение (обозначим его через μ) множества всех таксонов в множество их имен, построенное по рассмотренному выше правилу, взаимно однозначно.
2. Для всех таксонов T_i, T_j включение $T_i \subseteq T_j$ выполняется тогда и только тогда, когда $\mu(T_j) \subseteq \mu(T_i)$.

Как уже неоднократно отмечалось, таксономия есть полная решетка, в которой операция умножения таксонов является их теоретико-множественным пересечением. Однако операция сложения двух таксонов в общем случае не может быть столь же просто выражена через теоретико-множественные операции. Эта операция сопоставляет двум таксонам наименьший таксон, содержащий их объединение. В частном случае, когда таксономия построена на основе бинарного отношения, операция сложения (обозначим ее через +) таксонов может быть выражена так:

$$T_1 + T_2 = (T_1 \bigcup T_2)^{\Delta\nabla}.$$

С учетом антиизоморфизма таксономии и мерономии операцию сложения таксонов можно выразить через пересечение их имен и, наоборот, сложение имен можно выразить через пересечение соответствующих таксонов:

$$T_1 + T_2 = T_3 <=> \mu(T_1) \bigcap \mu(T_2) = \mu(T_3),$$

$$T_1 \bigcap T_2 = T_3 <=> \mu(T_1) + \mu(T_2) = \mu(T_3).$$

Здесь символом <=> обозначен союз "тогда и только тогда".

Таким образом, складывая два таксона, мы пересекаем их имена, и, наоборот; пересекая два таксона, мы складываем их имена, и наоборот.

Рассмотренные выше таксономия и мерономия являются примером того частного случая, когда операция сложения совпадает с теоретико-множественным объединением.

7.6. Увеличим нашу немощь, чтобы расширить границы господства

Существуют операции, которые довольно естественно интерпретируются как классификационные, но не являются операциями замыкания, и существуют семейства множеств, которые естественно интерпретируются как классификационные схемы, хотя и не являются таксономиями в смысле определения 7.1 *(см. разд. 7.2)*. Попробуем ослабить требования к операции таксонообразования, чтобы и другие интересные операции порождения классов удовлетворяли ее новому определению. Таким образом, мы хотим сделать некое обобщение операции таксонообразования, сохранив ее существенные свойства.

7.6.1. Определение слабой операции таксонообразования

Как и раньше, мы исходим из того, что классификатор создает таксоны на базе пробных подмножеств X таксономического универсума T_0, используя операцию θ. В общем случае не любое подмножество $X \subseteq T_0$ является таксоном. Будем считать, что таксонами являются только подмножества, устойчивые относительно операции θ, т. е. такие $X \subseteq T_0$, что $X = \theta(X)$. Другими словами, таксоны суть неподвижные точки отображения $\theta : 2^{T_0} \to 2^{T_0}$, где 2^{T_0} — множество всех подмножеств таксономического универсума T_0. Устойчивость таксонов относительно операции таксонообразования является одним из фундаментальных свойств любой естественной таксономии. В общем случае не для всякого отображения существуют неподвижные точки. Так что нам придется принять допущения, гарантирующие их существование.

Для классификационной деятельности довольно типичен итерационный процесс синтеза отдельных таксонов и таксономии в целом, который, однако, должен завершиться за конечное время. Поэтому допустим, что операция θ позволяет для любого исходного множества $X \subseteq T_0$ построить таксон методом последовательного приближения за конечное число шагов. Кроме того, потребуем, чтобы из произвольного подмножества таксона можно было построить подтаксон данного таксона в смысле теоретико-множественного включения.

Введем следующие обозначения:

$\theta^0(X) = X$,

$\theta^n(X) = \theta(\theta^{n-1}(X))$ при $n \ge 1$.

Очевидно, что $\theta^n(\theta^k(X)) = \theta^k(\theta^n(X)) = \theta^{n+k}(X)$ для любых $n \ge 0, k \ge 0$.

Определение 7.3

Пусть дана операция $\theta : X \to \theta(X)$, определенная для любого $X \subseteq T_0$, $\theta(X) \subseteq T_0$, и $n \ge 0$ — наименьшее целое число, для которого при любых $X, Y \subseteq T_0$ выполняются следующие соотношения:

$\theta^n(X) = \theta^{n+1}(X)$; (B1)

$Y \subseteq \theta^n(X) \Rightarrow \theta^n(Y) \subseteq \theta^n(X)$. (B2)

Тогда операцию θ будем называть *n*-слабой (1/*n*-сильной) операцией таксонообразования, а множества $X = \theta(X)$ — таксонами.

Очевидно, что если θ — *n*-слабая операция таксонообразования, то $\theta^n(X)=\theta^{n+k}(X)$ для любого $X \subseteq T_0$ и $k=0,1,2,...$. При этом множества вида $\theta^n(X)$ и только они являются таксонами.

По свойству В2 для любых $X, Y \subseteq T_0$ имеет место включение:

$$\theta^n(\theta^n(X) \bigcap \theta^n(Y)) \subseteq \theta^n(X) \bigcap \theta^n(Y).$$

Это аналог теоремы 7.4 для обычной операции таксонообразования.

7.6.2. Примеры

Теперь рассмотрим несколько операций преобразования множеств на предмет соответствия их определению *n*-слабой операции таксонообразования.

Тождественное преобразование

Тождественное преобразование сопоставляет любому множеству это же самое множество, т. е. ничего с ним не делает. Очевидно, что оно является простейшим примером 0-слабой операции таксонообразования. Если классификатор может рассматривать каждое подмножество таксономического универсума в качестве таксона, то он обладает самой сильной (∞-сильной) операцией таксонообразования.

Операция замыкания

Операция замыкания по теореме 7.5 равносильна операции таксонообразования, отвечающей определению 7.1. Поскольку она нам понадобится при рассмотрении последующих примеров, обозначим ее через C и еще раз приведем ее определение:

$$X \subseteq C(X);$$

$$Y \subseteq C(X) \Rightarrow C(Y) \subseteq C(X).$$

Данная операция является 1-слабой операцией таксонообразования.

Операция открывания

Операция открывания или взятия внутренности множества $O: X \to O(X)$ определяется следующими свойствами:

$$O(X) \subseteq X; \qquad \text{(O1)}$$

$$O(X) \subseteq Y \Rightarrow O(X) \subseteq O(Y). \qquad \text{(O2)}$$

Множества X, такие что $X = O(X)$, называются *открытыми*, при этом множество $O(X)$ называется *внутренностью множества* X.

Операция открывания противоположна (двойственна) операции замыкания. С точки зрения классификации и в отличие от операции замыкания ее можно интерпретировать как уточнение. Так, рассматривая какое-нибудь множество объектов, классификатор обнаруживает некоторые свойства, общие для объектов "ядра" (внутренности) данного множества. При этом он удаляет из исходного множества остальные объекты, как исключения или аномальные отклонения от обнаруженной закономерности.

Теорема 7.6. Операция $O: X \to O(X)$ на множестве T_0 тогда и только тогда является операцией открывания, когда для любых $X, Y \subseteq T_0$ выполняются следующие соотношения:

$O(X) \subseteq X$;

$O(X) = O(O(X))$; (O3)

$X \subseteq Y => O(X) \subseteq O(Y)$. (O4)

Доказательство аналогично доказательству теоремы 7.5.

Теорема 7.7. Операция открывания является 1-слабой операцией таксонообразования.

Доказательство. Свойство O3 есть частный случай свойства B1 определения 7.3 при $n = 1$: $O^1(X) = O^2(X)$. По свойству O4 имеем:

$$Y \subseteq O(X) => O(Y) \subseteq O(O(X)),$$

откуда с учетом свойства O3 получаем частный случай свойства B2:

$$Y \subseteq O(X) => O(Y) \subseteq O(X).$$

Известно, что если C — операция замыкания на множестве T_0, то операция $X \to \overline{C(X)}$, где черта означает дополнение до T_0, является операцией открывания. Так, рассматривая некоторое выборочное множество X примеров синтезируемого таксона, классификатор переходит к анализу противоположного (контрарного) множества примеров X, на базе которого с помощью некоторой операции замыкания C строит контрарный предтаксон $C(\overline{X})$ (это еще не окончательный результат, поэтому мы и говорим о предтаксоне, а не о таксоне). Затем классификатор удаляет из исходного множества X все элементы, которые принадлежат контрарному предтаксону $C(\overline{X})$, получая в результате таксон $X \setminus C(\overline{X}) = \overline{C(X)}$, являющийся внутренностью множества X. Рассмотренный способ построения таксонов моделирует метод рассуждений от противного.

Операция определения границы

Пусть C — операция замыкания на множестве T_0, тогда операцию B, такую что для любого $X \subseteq T_0$ выполняется $B(X) = C(X) \bigcap C(\overline{X})$, будем называть *операцией определения границы*, а множество $B(X)$ — границей множества X.

Легко видеть, что

$$B(X) = B(\overline{X});$$

$$X \bigcup B(X) = C(X).$$

Если обозначить $C(\overline{X})$ через $O(X)$, то

$$O(X) = X \setminus B(X);$$

$$O(X) \bigcap B(X) = \text{Ш};$$

$$C(B(X)) = B(X).$$

Можно также показать, что операция B является 2-слабой операцией таксонообразования:

$$B^2(X) = B^3(X);$$

$$Y \subseteq B^2(X) => B^2(Y) \subseteq B^2(X).$$

Если множество X открыто, то оно не имеет общих границ ни с $B(X)$, ни с $B^2(X)$. Таким образом, при классифицировании объектов из X в данном случае не обязательно имеет место включение $X \subseteq B^2(X)$. Иначе говоря, результатом классифицирования объекта может быть таксон, который его не содержит. Например, в основе классифицирования удаленных от наблюдателя морских судов по их силуэтам лежит операция выявления границы (контура) между множеством точек судна и фоном.

Комбинация операций замыкания и открывания

Можно представить себе классификационную деятельность как итерационный процесс перемежающихся уточнений и обобщений.

Пусть на одном и том же таксономическом универсуме в качестве обобщения используется некоторая операция замыкания C, а в качестве уточнения — двойственная ей операция открывания O. Обозначим через OC и CO такие операции на таксономическом универсуме T_0, что для любого $X \subseteq T_0$ имеют место соотношения:

$$OC(X) = O(C(X));$$

$$CO(X) = C(O(X)).$$

Примечание

Если использовать топологическую терминологию, то $OC(X)$ — открытая область, а $CO(X)$ — замкнутая область.

Непосредственно из определений операций C и O следует еще одна теорема.

Теорема 7.8. Операции OC и CO являются 1-слабыми операциями таксонообразования.

7.6.3. Особенности классифицирования

Как и в случае обычной операции таксонообразования, в основу правила классифицирования естественно положить принцип, согласно которому всякий объект $x \in T_0$ относится к таксону $\theta^n(\{x\})$, т. е. к таксону, который может быть построен на базе одноэлементного множества $\{x\}$ с помощью *n*-слабой операции таксонообразования θ.

Рассмотрим таксономию Tax, построенную с помощью некоторой *n*-слабой операции таксонообразования θ на таксономическом универсуме T_0. Пусть $\bigcap T(x)$ — пересечение всех таксонов $T(x)$, содержащих объект x. Тогда для любого $x \in T_0$ по свойству B2 *(см. разд. 7.6.1)* имеет место включение $\theta^n(\{x\}) \subseteq \bigcap T(x)$. При этом, однако, в общем случае оказываются неверными такие традиционные в теории классификации высказывания, что

1. Для всякого $x \in T_0$ найдется таксон $T \in Tax$, такой что $x \in T$.
2. Для всякого $x \in T_0$ имеет место отношение принадлежности $x \in \theta^n(\{x\})$.

Традиционно считается, что функция классифицирования (отнесения к таксонам) $f : T_0 \to Tax$ объектов $x \in T_0$ к таксонам $T \in Tax$ должна обладать следующими свойствами:

1. $f(x)$ определено для каждого $x \in T_0$.
2. f однозначна.
3. $f(x) = T \Rightarrow x \in T$.

При этом не всегда удается обеспечить выполнение последних двух свойств одновременно. В таких случаях обычно жертвуют вторым свойством, сохраняя третье: результатом отнесения объекта к таксономии должны быть таксоны, содержащие данный объект в смысле теоретико-множественного отношения принадлежности $\in$. Этот принцип считается сам собой разумеющимся, поэтому обычно его не оговаривают. Однако мы не будем настаивать на его обязательном выполнении, оставляя в силе первые два свойства функции классифицирования. Вместо третьего свойства у нас для любого $x \in T_0$ имеет место следующее соотношение:

$$x \in T \Rightarrow f(x) \subseteq T.$$

Иначе говоря, функция классифицирования должна быть всюду определенной и однозначной, но при этом объект может относиться и к не содержащему его таксону. Мы уже встречались с такой ситацией. Когда рассматривали в качестве 2-слабой операции таксонообразования операцию определения границы множества. Вместе с тем, отмеченное обстоятельство представляется довольно естественным в условиях нечеткой информации о классифицируемых объектах. Рассмотрим это подробнее.

Классифицируемые нечеткие объекты можно представить в виде подмножеств таксономического универсума, а не в виде его элементов. Очевидно, одноэлементные подмножества можно рассматривать как четкие объекты. Например, пусть таксономический универсум состоит из точек некоторого n-мерного пространства, координаты точек измеряются с погрешностями, а измерительные приборы время от времени могут выходить из строя. Тогда классификатор, имея набор координат, измеренных исправными в данный момент приборами, и принимая во внимание предельно допустимые погрешности, воспринимает не точку, а некоторое множество точек ("облако" или размытую точку). Аналогичная ситуация возникает, когда классификатор вполне уверен, что видит единственный предмет, но, жалуясь на плохие условия наблюдения, может лишь сказать, что наблюдаемый им предмет — один из некоторого списка (множества) возможных объектов. Данный список и представляет собой нечеткий объект, наблюдаемый классификатором. Результат классифицирования такого объекта зависит от свойств конкретной операции таксонообразования. Это может быть операция уточнения, и тогда результатом классифицирования является подмножество данного множества. Классификатора может интересовать граница нечеткого объекта, и тогда результат классифицирования может не пересекаться с множеством, представляющим данный нечеткий объект. Если используется операция обобщения, то результат классифицирования должен включать наблюдаемое множество таксономического универсума. Таким образом, рассматривая в качестве объектов классифицирования произвольные подмножества таксономического универсума, мы можем допустить различные соотношения между объектом классифицирования и таксоном, к которому он относится, а не только отношение принадлежности элемента множеству.

7.7. Классификация по нечетким отношениям

В прикладной математике есть раздел, посвященный так называемым нечетким множествам. Идея нечеткого множества возникла в результате попытки определения подмножества X некоторого обычного множества A, когда мы не можем с полной уверенностью сказать, принадлежит ли данный элемент из множества A его подмножеству X .

Обратите внимание: постановка такого вопроса противоречит классической идее множества, которая предполагает, что, формируя некоторое множество, мы вполне определенно знаем относительно каждого элемента, принадлежит ли он этому множеству или нет. Ведь обычное множество — это вполне определенная совокупность элементов, относительно каждого из которых однозначно можно сказать, принадлежит ли он этому множеству, или нет. Сама крамольная мысль о неоднозначности принадлежности элемента множеству сразу же выводит нас за рамки классической теории множеств. Вместе с тем, в середине XX века была создана некая математическая конструкция для оперирования множествами, которая допускала возможность указать, что тот или иной элемент принадлежит данному множеству с некоторой степенью неопределенности, выражаемой действительным числом в пределах от 0 до 1. Мотивация была простой: на практике довольно часто даже профессионал в предметной области не может с полной уверенностью констатировать, что данный индивидуум принадлежит данному классу; он может лишь указать некоторое число, выражающее степень его уверенности в том, что это так. Иногда опрометчиво говорят, что степень уверенности соответствует вероятности. Однако в такой интерпретации не следует заходить слишком далеко. Теория вероятности, хотя и возникла из идей статистической обработки эмпирического материала наблюдений, является математической теорией, не зависящей от субъективных ощущений. Она занимается изучением объективного. В противном случае она вышла бы за рамки научного.

Практика требует привнесения нечеткости, недоопределенности и недопонимания в математику, которая для того и существует, чтобы устранить неопределенность, неоднозначность и нечеткость. Разве это не парадокс? Попробуем разобраться в этом на примере создания классов некоторого вида, а именно таксонов для заданного отношения.

Прежде чем отправиться в путь, напомню, что интерпретацией операции образования таксонов по заданному обычному бинарному отношению является рассмотренная ранее операция $\Delta\nabla$ *(см. разд. 7.3)*. С обычными отношениями вам должно быть все ясно, но как быть с нечеткими отношениями?

7.7.1. Что такое нечеткие множества?

Нечеткое множество — это конструкция, определяющая неким специальным образом подмножество обычного множества A, которое мы назовем для удобства универсумом. Универсум A — обычное множество, состоящее из вполне определенных и различимых между собой элементов. Именно этими элементами в их полной совокупности и определяется множество A. Добавлю (точнее, напомню), что определить множество A означает предоставить список всех элементов, принадлежащих этому множеству, или указать правило (функцию), которое для любого элемента однозначно показывает, принадлежит данный элемент данному множеству A или не принадлежит.

Итак, универсум для нас сейчас — это обычное множество. Теперь мы перейдем к определению нечетких множеств. Нечеткое множество — это подмножество X универсума A, которое представляется как перечень всех элементов множества A с указанием степени принадлежности данного элемента множеству X. Пусть, например, $A=\{a,b,c,d\}$ — обычное множество, состоящее из четырех элементов a,b,c,d; определим подмножество X, указав для каждого элемента множества A степень его принадлежности подмножеству X: $X=\{a/0{,}1, b/0{,}5, c/1, d/0\}$. В данном примере элемент c с полной уверенностью принадлежит подмножеству X; элемент d — с полной уверенностью не принадлежит подмножеству X; элемент a принадлежит множеству X с уверенностью 0,1, а элемент b — с уверенностью 0,5. Степени уверенности, как предполагается, присваиваются элементам универсума при формировании его подмножеств экспертами — специалистами в конкретной предметной области.

Мы будем считать, что при определении подмножества X фиксированного универсума A используется некоторая функция принадлежности $\mu_X(x)$, определенная для всех элементов множества A и принимающая значения из интервала $0 \le \mu_X(x) \le 1$. Таким образом, нечеткое множество X можно представить как множество пар вида $(x,\mu_X(x))$, где $x \in A, 0 \le \mu_X(x) \le 1$. Нечеткое множество X, представленное таким образом, будем обозначать как $\tilde{X}$.

Нечеткое множество $\tilde{X} \subseteq A$ будем называть *четким*, если для всех $x \in A$ функция принадлежности $\mu_X(x)$ равна либо 1, либо 0.

Очевидно, что если $\tilde{X}$ — четкое множество, то оно может быть представлено в форме обычного множества, а обычное множество X может быть представлено в форме четкого множества (функция принадлежности для каждого его элемента имеет значение 1). Когда четкое множество $\tilde{X}$ может быть представлено как обычное множество X и наоборот, будем говорить, что $\tilde{X}$ и X *соответствуют* друг другу.

Пусть, например, $\tilde{X}=\{a/0, b/1, c/1, d/0\}$ — нечеткое множество. В действительности оно является четким, поскольку значения функции принадлежности для всех его элементов равны 0 либо 1. Этому множеству соответствует обычное множество $X=\{b,c\}$.

Допустим теперь, что дано обычное множество $X=\{a,b,c\}$, являющееся подмножеством универсума $A=\{a,b,c,d,e\}$. Этому обычному подмножеству соответствует нечеткое подмножество $\tilde{X}=\{a/1, b/1, c/1, d/0, e/0\}$.

Из анализа значений функции принадлежности мы заключаем в данном случае, что нечеткое множество $\tilde{X}$ является четким.

Итак, мы имеем дело с обычными и нечеткими множествами. Последние рассматриваются как некое обобщение обычных множеств. При этом нечеткое множество называется *четким*, если для любого его элемента функция принадлежности принимает только одно из двух значений — 0 или 1; в данном частном случае мы имеем возможность взаимно однозначно представить четкое множество как обычное.

Далее теория нечетких множеств развивалась в направлении придумывания таких функций принадлежности, чтобы, с одной стороны, они не очень сильно ограничивали экспертов в формировании своих оценок, а с другой стороны, не сжигали мосты, т. е. оставляли возможность интерпретации в теории обычных множеств. Так, если эксперты сформировали каким-то образом нечеткие множества $\tilde{X}$ и $\tilde{Y}$, то они должны были безоговорочно согласиться с результатами применения к этим множествам теоретико-множественных операций, таких как дополнение, объединение и пересечение. Эти операции в теории нечетких множеств могут определяться так или иначе, но в зависимости от ранее определенных функций принадлежности элементов к множествам.

Ниже приведены наиболее типичные определения отношений и операций между нечеткими множествами.

Для любых нечетких множеств $\tilde{X}, \tilde{Y} \subseteq A$ отношение включения $\tilde{X} \subseteq \tilde{Y}$ выполняется тогда и только тогда, когда для всех $x \in A$ имеет место неравенство $\mu_X(x) \leq \mu_Y(x)$.

Для дополнения, пересечения и объединения нечетких множеств функция принадлежности обычно определяется следующим образом:

$$\mu_{\overline{X}}(x) = 1 - \mu_X(x);$$

$$\mu_{X \cap Y}(x) = \min(\mu_X(x), \mu_Y(x));$$

$$\mu_{X \cup Y}(x) = \max(\mu_X(x), \mu_Y(x)).$$

Здесь $\min, \max$ — операции выбора соответственно минимального и максимального из двух значений.

Пусть, например, $\tilde{X} = \{a/0.1, b/0.5, c/1, d/0\}$,

$$\tilde{Y} = \{a/0{,}3,\ b/0{,}8,\ c/0{,}9,\ d/0\}, \text{ тогда}$$

$$\overline{\tilde{X}} = \{a/0{,}9,\, b/0{,}5,\, c/0,\, d/1\};$$

$$\overline{\tilde{Y}} = \{a/0{,}7,\, b/0{,}2,\, c/0{,}1,\, d/1\};$$

$$\tilde{X} \bigcap \tilde{Y} = \{a/0{,}1,\, b/0{,}5,\, c/0{,}9,\, d/0\};$$

$$\tilde{X} \bigcup \tilde{Y} = \{a/0{,}3,\, b/0{,}8,\, c/1,\, d/0\}.$$

7.7.2. Классы и ко-классы нечетких бинарных отношений

Обычное бинарное отношение R между обычными множествами A и B определяется как подмножество $R \subseteq A \times B$. Нечеткое отношение $\tilde{R}$ между множествами A и B определяется посредством функции принадлежности $\mu_R(x, y)$, которая для любых $x \in A, y \in B$ принимает значение от 0 до 1, указывающее степень нашей уверенности в том, что элемент x находится в отношении $\tilde{R}$ с элементом y или, иначе говоря, что пара (x, y) принадлежит подмножеству $\tilde{R} \subseteq A \times B$.

Теперь определим операцию построения классов и ко-классов нечеткого бинарного отношения. Данная задача сводится к заданию надлежащей функции принадлежности элементов к множествам вида $\tilde{X}^{\Delta}$ и $\tilde{Y}^{\nabla}$.

Пусть $\tilde{X} \subseteq A$ — некоторое нечеткое подмножество, а $\tilde{R} \subseteq A \times B$ — нечеткое отношение. Тогда для каждого элемента $y \in B$ определим функцию принадлежности множеству $\tilde{X}^{\Delta}$ следующим образом:

$$\mu_{X^{\Delta}}(y) = \min_{x \in A} \mu_{x^{\Delta}}(y),$$

где функция $\mu_{x^{\Delta}}(y)$ определяется так:

$$\mu_{x^{\Delta}}(y) = 1, \text{ если } \mu_R(x, y)/\mu_X(x) > 1;$$

$$\mu_{x^{\Delta}}(y) = \mu_R(x, y)/\mu_X(x) \text{ — в противном случае.}$$

Аналогичным образом определяются $\mu_{Y^{\nabla}}(x) = \min\limits_{y \in B} \mu_{y^{\nabla}}(x)$ и $\mu_{y^{\nabla}}(x)$.

Можно показать, что операция $\tilde{X} \to \tilde{X}^{\Delta\nabla}$ является операцией замыкания, причем для любых четкого множества $\tilde{X}$ и обычного множества X, четкого отношения $\tilde{R}$ и обычного отношения R, находящихся в соответствии, четкое множество $\tilde{X}^{\Delta\nabla}$ и обычное множество $X^{\Delta\nabla}$ также соответствуют друг другу. Поэтому мы можем считать, что рассмотренное определение операции $\tilde{X} \to \tilde{X}^{\Delta\nabla}$ является естественным расширением операции $X \to X^{\Delta\nabla}$ на случай нечетких множеств и отношений.

Заключение

> Пояснительные выражения объясняют темные мысли.
>
> *Козьма Прутков*

Дело сделано, книга говорит сама за себя. Я надеюсь, что читатель понял, что мне хотелось донести до него. И все же я считаю нелишним рассказать о соображениях, побудивших меня написать данную книгу.

Широко распространено мнение, что математика слишком суха и абстрактна, слишком далека от красочной жизни, чтобы было интересно ею заниматься. Так думает большинство юристов, журналистов, философов и других, кто считает свою профессию творческой. Мне иногда кажется, что, подчеркивая слово "творческая", они имеют в виду, в первую очередь, нерегламентированность объективными законами, неподверженность четкому определению и формальной логике и только во вторую очередь — необъяснимость спонтанного созидательного акта, подобного изобретению, открытию посредством наития, осененности свыше и т. п. Недаром у нас под творческой интеллигенцией понимают, как правило, артистов и писателей-беллетристов. Математик не тот, который быстро производит всякие там вычисления в уме, а тот, кто может распознать в текущей жизненной ситуации некую общую задачу, решение которой, возможно, уже известно, и требуется только разыскать его в книгах и справочниках. В том же случае, когда решения нет, математик, в отличие от других, это понимает и осознает, что либо его следует искать путем исследований и вспомогательных построений, либо отложить на время или навсегда (если, конечно, удастся доказать, что данная задача принципиально не имеет решения).

Психологи, психоаналитики и психотерапевты занимаются в основном сферой чувств (эмоций). Хотя мышление — не чуждая им область, но не она главный предмет их внимания. А как же иначе, если даже элементарная математическая логика никогда надлежащим образом ими не изучалась... Их познания в ней ограничиваются лишь элементарными импликациями вида "если …, то …", да и то из-за необходимости как-то различать причины и следствия. Без этой способности вообще невозможно слыть мыслящим человеком. А где искать материал по человеческому мышлению? Конечно же не в историях болезней пациентов неврологических клиник. Вполне достаточно истории математики и математиков, чтобы выйти в широчайшее поле для исследования особенностей развития человеческого интеллекта. К сожалению, логическому и математическому образованию специалистов-гуманитариев в этих важных для практики областях традиционно уделяется слишком мало внимания и времени. Мне приходилось наблюдать в разговорах и даже в книгах на медицинские темы беспомощные попытки авторов разъяснить, что является причиной, а что — видимым следствием (синдромом) болезни. Беспомощные — только потому, что авторы не имели элементарных навыков анализа систем с обратными связями. Неудивительно, что, занявшись описанием кольца, они в конце концов путали, где его начало, а где — конец. Суть данной проблемы осознана давным-давно. Еще в конце 1940-х годов английский врач и биолог (заметьте — не математик или инженер) Уильям Росс Эшби, вслед за основоположником кибернетики Норбертом Винером, написал замечательную популярную книгу "Введение в кибернетику". Винер, занимаясь частными вопросами изучения технических систем с обратными связями (автоматов), имел смелость провозгласить тезис (спекулятивный и крамольный) о том, что более сложные системы, такие как живые, построены на этом же принципе. Как он посмел, не зная профессионально биологии и медицины? Но тут, к счастью и своевременно, появилась книга профессионального биолога Эшби. Она была замечательна не только тем, что популяризировала философские идеи о существовании объективных законов управления, связи и информации, но и тем, что автор рассматривал эти законы по отношению к святая святых — живому. Прецедент состоялся, но практически ему не суждено было повториться, к сожалению, до последнего времени. Однако следует заметить, что общефилософские замечания Винера и Эшби не прошли даром. Спустя лишь несколько лет, в середине 1950-х, появилась теория конечных автоматов, которая легла в основу разработки не только элементарных переключательных устройств типа светофоров на улицах, но также интерпретаторов и компиляторов языков программирования и компьютерных систем перевода текстов с одного языка на другой. И действительно убеждаешься в том, что "Вначале было слово…".

В начале 1980-х небольшая группа энтузиастов, в которую входил и ваш покорный слуга, по воле случая, а может, и необходимости, была вовлечена в увлекательную работу по автоматизации процессов управления космическими аппаратами (искусственными спутниками Земли) для обеспечения связи. Мы могли, что называется, взять быка за рога и попытаться решить непосредственно конкретные технические задачи, которые ставились перед нами.

Но нас интересовало нечто большее. Конкретная задача была лишь поводом для занятий более общими и, как нам тогда казалось, важными вопросами. Наше базовое вузовское математическое образование ограничивалось главами математического анализа и теории случайных процессов, практически не включая в себя алгебру и дискретную математику. Однако насущные задачи требовали познаний в логике, теории графов, теории оптимизации и т. д. Именно тогда нам, к счастью, попалась изумительная книга Ю. А. Шрейдера "Равенство, сходство, порядок" (кстати, ее автор, кандидат физико-математических наук, немного позднее стал доктором философии). Эта книга не столько дала нам конкретные и нужные знания, сколько вооружила методикой исследования в необычной для нас области. Множества и отношения — как можно конструктивно применять все это в практической области, а не только заимствовать специфические терминологию и символику, как это часто бывает в современных диссертациях на технические темы? Первое, что удалось сделать, — переизложить теорию отношений с точки зрения классов. Данный подход придумал мой товарищ Олег Поляков, а Ю. А. Шрейдер тогда признался, что в свое время упустил из виду так называемые связи Галуа, хотя был так близок к ним. Это замечание окрылило нас, и мы двинулись дальше в сторону разработки математической теории классификации и симбиоза результатов теории конечных автоматов с реляционными базами данных. Это было нужно, по нашему мнению, для создания моделей сложных технических устройств, чтобы применять их для оценки технического состояния, испытаний и диагностики неисправностей реальных технических систем.

Я не теряю надежды, что, может быть, и для кого-нибудь из вас, уважаемые читатели, данная книга сыграет роль, подобную той, которую для меня в свое время сыграла книга "Равенство, сходство, порядок".

Литература

Общая, знаменитая и популярная

1. Пуанкаре А. О науке. — М.: Наука, 1983. — 560 с.
2. Клайн М. Математика. Утрата определенности. — М.: Мир, 1984. — 434 с.
3. Стройк Д. Краткий очерк истории математики. — М.: Наука, 1989. — 400 с.
4. Реньи А. Трилогия о математике. — М.: Мир, 1980. — 376 с.
5. Белнап Н., Стил Т. Логика вопросов и ответов. — М.: Прогресс, 1981. — 288 с.

Более специальная, но достаточно популярная

1. Вейль Г. Математическое мышление. — М.: Наука, 1983. — 560 с.
2. Шрейдер Ю. А. Равенство, сходство, порядок. — М.: Наука, 1971. — 254 с.
3. Шрейдер Ю. А. Алгебра классификации. — Научно-техническая информация. Сер. 2, 1974, №9, С. 1—5.
4. Шрейдер Ю. А., Шаров А. А. Системы и модели. — М.: Радио и связь, 1982. — 152 с.
5. Мейен С. В., Шрейдер Ю. А. Методологические аспекты теории классификации. — Вопросы философии, 1976, №12, С. 67—79.
6. Дунаев В. В. Об одной модели классификации. — Научно-техническая информация. Сер. 2, 1990, №3, С. 22—27.

7. Дунаев В. В., Поляков О. М. Методологические аспекты реляционной теории классификации. — Научно-техническая информация. Сер. 2, 1987, №4, С. 21—27.
8. Дунаев В. В. Базы данных. Язык SQL. — СПб.: БХВ-Петербург, 2006. — 288 с.
9. Поляков О. М., Дунаев В. В. Классификационные схемы: синтез через отношения. — Научно-техническая информация. Сер. 2, 1985, №6, С.15—21.

Очень специальная

1. Биркгоф Г. Теория решеток. — М.: Наука, 1984. — 567с.
2. Келли Д. Общая топология. — М.: Наука, 1981. — 431 с.

Предметный указатель

П

Р

С

Т

У

Ф, Х

Ц

Ч

Ш

Э

Я