ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

В. М. Зюзьков

ТЕОРИЯ АЛГОРИТМОВ

Издательство Томского университета

2005

УДК 512.8, 510.5

Зюзьков В. М.
Теория алгоритмов: Учебное пособие. – Томск: Изд-во Том. Ун-та, 2005. – 101 с.

В пособии содержится материал спецкурса, читаемого автором в последние годы на механико-математическом факультете ТГУ для студентов специализации «Компьютерная математика». Основные разделы: алгоритмы и вычислимые функции, ламбда–исчисление, вычислимое и невычислимое, формальные аксиоматические теории, элементарная арифметика и неполнота, сложность вычислений, *NP*-полнота.

Предназначено для преподавателей математики и компьютерных наук, а также для студентов высших учебных заведений.

Рецензенты:
Кафедра программирования Томского государственного университета, зав. кафедрой д-р техн. наук, профессор А. Ю. Матросова
Зав. кафедрой прикладной математики Томского государственного архитектурно-строительного университета, д-р физ.-мат. наук, профессор М. И. Слободской

Оглавление

Предисловие 5

1 Алгоритмы и вычислимые функции 7

1.1 Понятие алгоритма и неформальная вычислимость 7

1.2 Частично-рекурсивные функции 8

Определения 8

Примеры рекурсивности 9

1.3 Машины Тьюринга 11

2 Вычислимое и невычислимое 15

2.1 Тезис Чёрча 15

2.2 Теорема о рекурсии 16

2.3 Куины 18

2.4 Некоторые алгоритмически неразрешимые проблемы 20

3 Формальные аксиоматические теории 24

3.1 Основные определения 24

Что такое формальная теория? 24

Выводимость 25

Интерпретация 27

Общезначимость и непротиворечивость 28

Полнота, независимость и разрешимость 29

3.2 Примеры формальных теорий 29

Формальная система *MIU* 29

Формальная система *PR* 30

Формальная система *UR* 31

Формальная система *PR*1 31

3.3 Исчисление высказываний 32

Классическое определение исчисления высказываний 32

Полнота, разрешимость и непротиворечивость исчисления высказываний 32

3.4 Теории первого порядка 33

Языки первого порядка 33

Синтаксические свойства истинности теорий с языком первого порядка 34

Определение теории первого порядка 34

Некоторые свойства теорий первого порядка 36

Непротиворечивость, полнота и неразрешимость исчисления предикатов 37

4 Элементарная арифметика и неполнота 38

4.1 Элементарная арифметика 38

Теория элементарной арифметики 38

Арифметические функции и отношения 40

4.2 Неполнота элементарной арифметики 41

Геделева нумерация 41

Лемма о рефлексии 43

Теорема Гёделя о неполноте 44

Нестандартное расширение *EA* 47

Об аксиоматизации 48

4.3 Теорема Гудстейна 50

5 Сложность вычислений *54*

5.1 Асимптотические обозначения **54**

Θ–обозначение 54

O– и Ω-обозначения 55

Сравнение роста функций 56

5.2 Алгоритмы и их сложность **58**

5.3 Сложность задач **60**

6 NP-полнота *64*

6.1 Задачи разрешения и задачи оптимизации **64**

6.2 Формальные языки **65**

6.3 Проверка принадлежности языку и класс NP **67**

6.4 NP-полнота и сводимость **68**

Литература *71*

Предисловие

Использование компьютеров связано с возможностью алгоритмического решения задач и эффективного вычисления функций. Между тем в математике широко используются функции, заданные неэффективными определениями. Это приводит к тому, что некоторые функции поддаются вычислению с помощью алгоритма, скажем на компьютере, как только для этого будет составлена надлежащая программа, тогда как другие функции, заданные неэффективным определением, могут требовать творческого подхода для вычисления своих значений. Столь же часты доказательства разрешимости задач, не сопровождаемые алгоритмами их решения.

Формализация понятия вычислимости позволяет увидеть связь алгоритмов с процессом логического вывода, установить пределы доказуемости в формальных системах, получить результаты о неполноте математики.

В действительности класс задач, доступных классическим средствам, в некотором трудно уточняемом смысле строго шире класса задач, решаемых алгоритмически. В пособии проясняется смысл этого утверждения, и излагаются некоторые математические модели вычислимости. За последние десятилетия стало ясно, что различие между быстро и долго решаемыми задачами не менее философски и практически важно, чем различие между алгоритмически разрешимыми и неразрешимыми, и теория сложности вычислений стала одной из центральных в логике (и вообще в математике). Поэтому также в пособии рассматриваются основные понятия теории сложности алгоритмов.

Особое внимание автор обращает на две темы: ламбда–исчисление и теорема Гёделя о неполноте.

Как изучать теорию алгоритмов? С чего начинать? Нам привлекательна следующая стратегия, описанная в истории, взятой из книги [19].

Кувшин

Знаменитый китайский профессор из известного китайского университета сидел перед новой группой студентов. Прямо перед ним стоял большой стеклянный кувшин, полупрозрачный, легкого зеленоватого оттенка.

Профессор смотрел на студентов, не произнося ни слова. Затем он наклонился вправо. У его правой ноги лежала кучка камней, каждый из которых мог бы поместиться в кулак. Он взял один из камешков и очень осторожно опустил его в кувшин через узкое горлышко. Потом взял следующий и повторил эту процедуру. Он проделывал это до тех пор, пока камни не поднялись до самого горлышка и не заполнили весь кувшин.

Он повернулся к группе и сказал:

– Скажите мне, этот кувшин полон?

Группа согласно зашелестела. Кувшин, без сомнения, был наполнен.

Профессор ничего не сказал и обернулся в левую сторону. Около его левой ноги была насыпана горка мелкой гальки. Он набрал полную горсть и стал аккуратно засыпать гальку через горлышко кувшина. Горсть за горстью, он сыпал гальку в кувшин, а она просыпалась сквозь щели между камнями, пока не дошла до самого верха, и уже невозможно было насыпать даже маленькую толику.

Он повернулся к аудитории и спросил:

– Скажите мне, полон ли кувшин сейчас?

Группа пробормотала, что все выглядит так, как если бы на этот раз кувшин действительно был полон; возможно, полон; наверное.

Профессор ничего не сказал и снова повернулся к правой стороне. Около его ноги была насыпана горка крупного сухого песка. Он набрал горсть песка и начал аккуратно сыпать его через горлышко кувшина. Песок просыпался сквозь камни и гальку, а профессор горсть за горстью сыпал его в кувшин, пока песок не достиг горлышка и стало ясно, что больше насыпать невозможно.

Он повернулся к группе студентов и спросил:

– Кто-нибудь может сказать мне, полон ли сейчас кувшин?

Ответом была тишина.

Профессор снова ничего не сказал и обернулся влево. Около его левой ноги стоял графин с водой. Он взял его в руки и начал осторожно лить воду через горлышко кувшина. Вода стекала на дно, минуя камни, гальку и песок, заполняя свободное пространство, пока не поднялась до самого горлышка.

Он повернулся к группе и спросил:

– Скажите мне, полон ли сейчас кувшин?

В аудитории было тихо, даже тише, чем раньше. Это был тот тип тишины, когда все склоняют свои головы и старательно рассматривают свои ногти или оценивают чистоту своих ботинок. Или делают и то и другое одновременно.

Профессор снова обернулся вправо. На небольшом кусочке голубой бумаги была насыпана небольшая горочка великолепной мелкой соли. Он взял щепотку соли и бережно всыпал её через узкое горлышко кувшина, и она растворилась в воде. Щепотку за щепоткой он всыпал соль в воду, она растворялась, проникая сквозь камни, гальку и песок, пока не стало ясно, что соль больше не может растворяться в воде, так как та перенасыщена ею.

И снова профессор повернулся к группе и спросил:

– Скажите мне, а сейчас кувшин полон?

Один очень смелый студент встал и сказал:

– Нет, профессор, он еще не полон.

– А-ааа! – протянул профессор. – Но он полон.

Затем профессор предложил всем присутствующим обсудить значение этой ситуации. Что она значила? Как мы можем её интерпретировать? Почему профессор сказал это? И несколько минут спустя профессор уже выслушивал их предположения.

Интерпретаций было столько же, сколько студентов в этой аудитории.

Когда профессор выслушал каждого студента, он поздравил их всех, сказав, что удивлен таким обилием интерпретаций. Каждый из присутствующих является уникальным человеком, который живет и смотрит на жизнь через призму собственного, уникального опыта, не схожего ни с чьим другим. Их интерпретации просто отражали их жизненный опыт, особый и уникальный взгляд, через который они понимают мир.

И поэтому ни одна интерпретация не была лучшей или худшей, чем остальные. И он поинтересовался, интересна ли группе его собственная интерпретация? Она, разумеется, не является правильной, не может быть лучше или хуже их предположений. Это просто его интерпретация.

Конечно, всем было очень интересно.

– Что ж, – сказал он, – моя интерпретация проста. Что бы вы ни делали в своей жизни, в каком бы то ни было контексте, будьте уверены, что **сначала вы положили камни**.

В. М. Зюзьков

1 Алгоритмы и вычислимые функции

1.1 Понятие алгоритма и неформальная вычислимость

Под **алгоритмом** понимается способ преобразования представления информации. Слово *algorithm* – произошло от имени аль-Хорезми – автора известного арабского учебника по математике (от его имени произошли также слова «алгебра» и «логарифм»).

Интуитивно говоря, алгоритм – некоторое формальное предписание, действуя согласно которому можно получить решение задачи.

Алгоритмы типичным образом решают не только частные задачи, но и классы задач. Подлежащие решению частные задачи, выделяемые по мере надобности из рассматриваемого класса, определяются с помощью параметров. Параметры играют роль исходных данных для алгоритма.

Основные особенности алгоритма

- *Определенность.* Алгоритм разбивается на отдельные шаги (этапы), каждый из которых должен быть простым и локальным.
- *Ввод.* Алгоритм имеет некоторое (быть может, равное нулю) число входных данных, т.е. величин, заданных ему до начала работы.
- *Вывод.* Алгоритм имеет одну или несколько выходных величин, т. е. величин, имеющих вполне определенное отношение к входным данным.
- *Детерминированность.* После выполнение очередного шага алгоритма, однозначно определено, что делать на следующем шаге.

Обратите внимание, что мы не требуем, чтобы алгоритм заканчивал свою работу для любых входных данных.

Примеры алгоритмов широко известны: изучаемые в школе правила сложения и умножения десятичных чисел или, скажем, алгоритмы сортировки массивов. Для алгоритмически разрешимой задачи всегда имеется много различных способов ее решения, т. е. различных алгоритмов.

Примеры «почти» алгоритмов: медицинский и кулинарный рецепты. Кстати, почему такие рецепты во многих случаях нельзя рассматривать как алгоритмы?

Данное здесь определение алгоритма не является, конечно, строгим, но оно интуитивно кажется вполне определенным. К сожалению, для решения некоторых задач не существует алгоритма. Установление таких фактов требует введения строгого понятия алгоритма.

Мы будем рассматривать алгоритмы, имеющие дело только с натуральными числами. Можно доказать, что это не является потерей общности, так как объекты другой природы можно закодировать натуральными числами. Для пользователей компьютеров такое утверждение должно быть очевидным.

Пусть $\boldsymbol{N}$ обозначает множество натуральных чисел {0, 1, 2, ...}. Объекты, которые мы будем рассматривать, будут функциями с областью определения $D_f \subseteq \boldsymbol{N}^k$ (k – целое положительное число) и с областью значений $R_f \subseteq \boldsymbol{N}$. Такие функции будем называть *k-местными частичными*. Слово «частичная» должно напомнить о том, что функция определена на подмножестве $\boldsymbol{N}^k$ (конечно, в частном случае может быть $D_f = \boldsymbol{N}^k$, тогда функция называется *всюду определенной*).

Вычислимые функции

Назовем k-местную функцию f: $\boldsymbol{N}^k \rightarrow \boldsymbol{N}$ *эффективно вычислимой* (или просто *вычислимой*), если существует алгоритм A, её вычисляющий, то есть такой алгоритм A, что:

1. Если на вход алгоритма A поступил вектор $\boldsymbol{x} = \langle x_1, x_2, ..., x_k \rangle$ из D_f, то вычисление должно закончиться после конечного числа шагов и выдать $f(\boldsymbol{x})$.

2. Если на вход алгоритма A поступил вектор $\boldsymbol{x}$, не принадлежащий области определения D_f, то алгоритм A никогда не заканчивается.

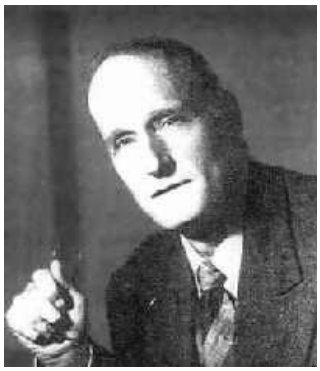
Несколько замечаний по поводу этого определения:

1. Понятие вычислимости определяется здесь для частичных функций (областью определения которых является некоторое подмножество натурального ряда). Например, нигде не определенная функция вычислима, в качестве A надо взять программу, которая всегда зацикливается.
2. Можно было бы изменить определение, сказав так: «если $f(\boldsymbol{x})$ не определено, то либо алгоритм А не останавливается, либо останавливается, но ничего не печатает на выходе». На самом деле от этого ничего бы не изменилось (вместо того, чтобы останавливаться, ничего не напечатав, алгоритм может зацикливаться).
3. Входами и выходами алгоритмов могут быть не только натуральные числа, но и двоичные строки (слова в алфавите {0, 1}), конечные последовательности слов и вообще любые, как говорят, «конструктивные объекты».
4. Множество эффективно вычисляемых функций мы не отождествляем с множеством «практически вычисляемых» функций, так как не накладываем на первое множество никаких ограничений, связанных с современными вычислительными машинами. Хотя каждое входное натуральное число должно быть конечным, тем не менее не предполагается верхняя граница размера этого числа, так, например, количество цифр числа может быть больше числа электронов во Вселенной. Точно так же нет никакой верхней границы на число шагов, которые может сделать алгоритм для конкретных $\boldsymbol{x}$ из области определения.

Рассматривая теорию алгоритмов, мы можем ссылаться на программистский опыт, говоря об алгоритмах, программах, интерпретаторах и т. д. Это позволяет нам игнорировать детали построения тех или иных алгоритмов под тем предлогом, что читатель их легко восстановит (или хотя бы поверит). Но в некоторых случаях этого недостаточно, поэтому мы собираемся дать строгое определение нового множества функций, которое в некотором смысле будет совпадать с множеством вычислимых функций. Мы дадим три различные формализации понятия вычислимой функции.

1.2 Частично-рекурсивные функции

Определения



Стивен Коул Клини

Этот подход к формализации понятия алгоритма принадлежит Гёделю и Клини (1936).

Основная идея Гёделя состояла в том, чтобы получить все вычислимые функции из существенно ограниченного множества базисных функций с помощью простейших алгоритмических средств.

Множество **исходных** функций таково:

- постоянная функция $0(\boldsymbol{x}) = 0$;
- одноместная функция следования $s(x) = x+1$;
- функция проекции pr_i, $1 \le i \le k$, $pr_i(\boldsymbol{x}) = x_i$.

Нетривиальные вычислительные функции можно получать с помощью композиции (суперпозиции) уже имеющихся функций. Этот способ явно алгоритмический.

- **Оператор суперпозиции**. Говорят, что k-местная функция $f(\boldsymbol{x})$ получена с помощью суперпозиции из m-местной функции $\varphi(y_1, y_2, ..., y_m)$ и k-местных функций $g_1(\boldsymbol{x})$, $g_2(\boldsymbol{x})$, ..., $g_m(\boldsymbol{x})$, если $f(\boldsymbol{x}) = \varphi(g_1(\boldsymbol{x}), g_2(\boldsymbol{x}), ..., g_m(\boldsymbol{x}))$.

Второй (несколько более сложный) способ действует так.

- **Примитивная рекурсия**. При $n \geq 0$ из n-местной функции f и $(n+2)$-местной функции g строится $(n+1)$-местная функция h по следующей схеме:
  $h(\boldsymbol{x}, 0) = f(\boldsymbol{x})$,
  $h(\boldsymbol{x}, y+1) = g(\boldsymbol{x}, y, h(\boldsymbol{x}, y))$.

При $n=0$ получаем (a – константа):
$h(0) = a$;
$h(y+1) = g(y, h(y))$.

Два упомянутых способа позволяют задать только всюду определенные функции. Частично-определенные функции порождаются с помощью третьего гёделева механизма.

- **Оператор минимизации**. Эта операция ставит в соответствие частичной функции f: $\boldsymbol{N}^{k+1} \to \boldsymbol{N}$ частичную функцию h: $\boldsymbol{N}^k \to \boldsymbol{N}$, которая определяется так ($\boldsymbol{x} = \langle x_1, \ldots, x_k \rangle$):
  1) область определения $D_h = \{\boldsymbol{x} \mid$ существует $x_{k+1} \geq 0, f(\boldsymbol{x}, x_{k+1}) = 0$ и $\langle \boldsymbol{x}, y \rangle \in D_f$ для всех $y \leq x_{k+1}\}$;
  2) $h(\boldsymbol{x})$ = наименьшее значение y, при котором $f(\boldsymbol{x}, y) = 0$.

Оператор минимизации обозначается так $h(\boldsymbol{x}) = \mu y[f(\boldsymbol{x},y) = 0]$.

Очевидно, что даже если f всюду определено, но нигде не обращается в 0, то $\mu y[f(\boldsymbol{x}, y) = 0]$ нигде не определено. Естественный путь вычисления $h(\boldsymbol{x})$ состоит в подсчете значения $f(\boldsymbol{x}, y)$ последовательно для $y = 0,1, 2,...$ до тех пор, пока не найдется y, обращающее $f(\boldsymbol{x}, y)$ в 0. Этот алгоритм не остановится, если $f(\boldsymbol{x}, y)$ нигде не обращается в 0.

Все функции, которые можно получить из базисных функций за конечное число шагов только с помощью трех указанных механизмов, называются **частично-рекурсивными**. Если функция получается всюду определенная, то тогда она называется **общерекурсивной**. Если функция получена без механизма минимизации, то в этом случае она называется **примитивно-рекурсивной**.

Любую примитивно-рекурсивную функцию можно вычислить с помощью цикла в форме *for*, так как верхнюю границу для числа повторений можно указать заранее. Оператор минимизации позволяет описать функции, которые нельзя вычислить за заранее ограниченное число итераций, для вычисления их значений требуется цикл в форме *while*.

Можно легко показать [17, с.136], что введение фиктивных переменных, а также перестановка и отождествление переменных не выводят за пределы класса примитивно-рекурсивных функций и класса частично-рекурсивных функций. Это проще всего объяснить на примерах.

Введение фиктивных переменных. Если $g(x_1, x_3)$ – примитивно-рекурсивная функция и $f(x_1, x_2, x_3) = g(x_1, x_3)$, то $f(x_1, x_2, x_3)$ – примитивно-рекурсивная функция.

Перестановка переменн*ых*. Если $g(x_1, x_2)$ - примитивно-рекурсивная функция и $f(x_2, x_1) = g(x_1, x_2)$, то f есть также примитивно-рекурсивная функция.

Отождествление переменных. Если $g(x_1, x_2, x_3)$ – примитивно-рекурсивная функция и $f(x_1, x_2) = g(x_1, x_2, x_1)$, то $f(x_1, x_2)$ есть также примитивно-рекурсивная функция.

Примеры рекурсивности

Рассмотрим примеры частично-рекурсивных функций. Все эти примеры и много других можно найти в [15, 17].

- Сложение двух чисел
  sum: $\langle x,y \rangle \to x+y$.
  Эта функция является общерекурсивной в силу примитивной рекурсии
  $sum(x,0) = pr_1(x) = x$,
  $sum(x,y+1) = s(sum(x,y)) = sum(x,y)+1$.

- Умножение двух чисел

prod: $<x,y> \to x\ y$.
Используем примитивную рекурсию
$prod(x,0) = 0(x) = 0$,
$prod(x,y+1) = sum(prod(x,y),x)$.

- Усеченное вычитание 1
  $\delta(x) = x-1$, если $x>0$,
  $\delta(0) = 0$.
  Эта функция примитивно-рекурсивна, действительно,
  $\delta(0) = 0 = 0(x)$,
  $\delta y+1) = y = pr_2(<x,y>)$.

- Усеченная разность
  $x \div y = x - y$, если $x \geq y$,
  $x \div y = 0$, если $x < y$.
  Эта функция примитивно-рекурсивна, действительно,
  $x \div 0 = x$,
  $x \div (y+1) = \delta(\ x \div y)$.

- Модуль разности
  $|x-y| = x-y$, если $x \geq y$,
  $|x-y| = y-x$, если $x<y$.
  Эта функция примитивно-рекурсивна в силу суперпозиции
  $|x-y| = (x \div y)+(y \div x)$.

- Факториал
  Действительно,
  $0! = 1$,
  $(y+1)! = prod(y!,y+1)$.

- $\min(x,y)$ – наименьшее из чисел x и y
  В силу суперпозиции: $\min(x,y) = x \div (x \div y)$.
- Знак числа
  $sg(x) = 0$, если $x = 0$,
  $sg(x) = 1$, если $x>1$.
  В силу рекурсии
  $sg(0) = 0$,
  $sg(y+1) = 1$.

- $rm(x, y)$ – остаток от деления y на x, если $x \neq 0$, и y, если $x=0$.
  В силу рекурсии и суперпозиции
  $rm(x,0) = 0$,
  $rm(x,y+1) = prod(s(rm(x,y)),sg(|x-s(rm(x,y))|))$.

Используя функции, для которых уже установлено, что они являются частично-рекурсивными, мы получаем все новые и новые частично-рекурсивные функции. Существуют критерии, которые позволяют установить частичную рекурсивность сразу для обширных классов функций (см., например, [17, с. 135–151]).

Используя минимизацию (μ-оператор) можно получать частично-определенные функции из всюду определенных функций. Например, полагая что $f(x,y)$ есть частично-

рекурсивная функция $|x-y^2|$, мы обнаруживаем, что $g(x) = \mu y[f(x,y)=0]$ – не всюду определенная функция:

$g(x) = \sqrt{x}$, если x есть точный квадрат, и неопределенна в противном случае.

Таким образом, тривиально используя μ-оператор вместе с суперпозицией и рекурсией, можно построить больше функций, исходя из основных, чем только с помощью суперпозиции и рекурсии (так как эти операции порождают из всюду определенных функций – всюду определенные). Существуют, однако, и общерекурсивные (всюду определенные) функции, для построения которых нельзя обойтись без минимизации.

Приведем пример функции, не являющейся примитивно рекурсивной, хотя и вычислимой в интуитивном смысле.

Определим последовательность одноместных функций F_n: $\boldsymbol{N} \to \boldsymbol{N}$, $n \in \boldsymbol{N}$, следующим образом:

$$F_0(x) = x+1,$$
$$F_{n+1}(x) = \underbrace{F_n(F_n(\ldots F_n}_{x+1 \text{ раз}}(1)\ldots))$$

Поэтому $F_1(x) = x+2$, $F_2(x) = 2x+3 \approx 2x$, $F_3(x) \approx 2^x$,

$F_4(x) \approx 2^{2^{\cdot^{\cdot^{2}}}}$ (башня из $x+1$ двойки) и т. д.

Имеем следующие свойства:

1) для каждого n функция $x \to F_n(x)$ является примитивно-рекурсивной;
2) $F_n(x) > 0$,
3) $F_n(x+1) > F_n(x)$,
4) $F_n(x) > x$,
5) $F_{n+1}(x) \geq F_n(x+1)$,
6) для каждой k–местной примитивно-рекурсивной функции $f(x_1,x_2,\ldots,x_k)$ существует такое n, что F_n мажорирует f, т.е. $f(x_1,x_2,\ldots,x_k) \leq F_n(\max(x_1,x_2,\ldots,x_k))$ для всех $x_1,x_2,\ldots,x_k$,
7) функция $A(n, x) = F_n(x)$ не является примитивно-рекурсивной [11, с. 53] (эта функция известна как **функция Аккермана**).

Функцию Аккермана можно определить и в традиционной записи:

$f(0, y) = y+1$,

$f(x+1,0) = f(x,1)$,

$f(x+1,y+1) = f(x, f(x+1,y))$.

Позднее мы приведем доводы в пользу правдоподобности того, что понятие частично-рекурсивной функции есть точный математический эквивалент интуитивной идеи (эффективно) вычислимой функции.

1.3 Машины Тьюринга

Рассмотрим еще один способ определения вычислимых функций, следуя в изложении [24, с. 12–14]. Формулировка, выраженная в терминах воображаемой вычислительной машины, была дана английским математиком Аланом Тьюрингом в 1936 г. Главная трудность при нахождении этого определения была в том, что Тьюринг искал его до создания реальных цифровых вычислительных машин. Познание шло от

абстрактного к конкретному: фон Нейман был знаком с работой Тьюринга, и сам Тьюринг позднее сыграл вдохновляющую роль в развитии вычислительных машин.

На неформальном уровне мы можем описывать машину Тьюринга как некий черный ящик с лентой. Лента разбита на ячейки, и каждая ячейка может содержать пустой символ 0 либо непустой символ 1. Лента потенциально бесконечна в обе стороны в том смысле, что мы никогда не придем к ее концу, но в любое время лишь конечное число ячеек может быть непустым. В начале лента содержит числа входа, в конце – число–выход. В промежуточное время лента используется как пространство памяти для вычисления.

Если мы откроем черный ящик, то обнаружим, что он устроен очень просто. В любой момент времени он может обозревать лишь одну ячейку памяти. Устройство содержит конечный список инструкций (или **состояний**) q_0, q_1,..., q_n. Каждая инструкция может указать два возможных направления действий; одного нужно придерживаться, если на обозреваемой ячейке ленты находится 0, а другого – если там находится 1. В любом случае следующее действие может состоять из таких трех типов элементарных шагов:

Алан Тьюринг

- символ (возможно, такой же, как старый) пишется на обозреваемой ячейке ленты, при этом предыдущий символ стирается;
- лента сдвигается на одну ячейку влево или вправо;
- указывается следующая инструкция.

Таким образом, список инструкций определяет некоторую функцию перехода, которая по данной инструкции и обозреваемому символу указывает три компоненты того, что нужно делать. Мы можем формализовать эти идеи, взяв в качестве машины Тьюринга эту функцию перехода.

Машина Тьюринга

Машина Тьюринга – это функция M такая, что для некоторого натурального числа n область определения этой функции есть подмножество множества $\{0, 1,..., n\}\times\{0, 1\}$, а область значений есть подмножество множества $\{0, 1\}\times\{Л, П\}\times\{0, 1, ..., n\}$.

Например, пусть $M(3,1) = <0, Л, 2>$. Подразумеваемый смысл этого состоит в том, что как только машина дойдет до инструкции q_3, а на обозреваемой ячейке написан символ 1, она должна стереть 1 (оставляя на ячейке 0), передвинуть ленту так, чтобы обозреваемой ячейкой стала левая соседняя ячейка от той, которая обозревалась, и перейти к следующей инструкции q_2. Если $M(3,1)$ не определено, тогда как только машина дойдет до инструкции q_3, а на обозреваемой ячейке написан символ 1, то машина останавливается. (Это единственный путь остановки вычисления.)

Такая подразумеваемая интерпретация не включена в формальное определение машины Тьюринга, но она мотивирует и подсказывает формулировки всех следующих определений. В частности, можно определить, что означает для машины M передвижение (за один шаг) от одной конфигурации до другой. Нам не нужно здесь давать формальных определений, так как они являются простыми переводами наших неформальных идей.

Входные и выходные данные – это строчки из 1, разделенные 0. Пусть $<n>$ будет строчкой из 1 длины n+1. Тогда

$$<n_1> 0 <n_2> 0 ... 0 <n_k>$$

получено комбинацией k строчек из 1, каждая отделена от другой 0.

Наконец, мы можем определить вычислимость.

Вычислимость по Тьюрингу

Пусть $D_f \subseteq \boldsymbol{N}^k$ – область определения k-местной функции f: $D_f \to \boldsymbol{N}$. Функция f называется **вычислимой по Тьюрингу**, если существует машина Тьюринга M такая, что как только M начинает с инструкции q_0, обозревая самой левый символ строки

$$<n_1> 0 <n_2> 0 \ldots 0 <n_k>,$$

(вся остальная часть ленты пуста), тогда:

- если $f(n_1, n_2,..., n_k)$ определено, то M, в конце концов, остановится, обозревая самый левый символ строки $<f(n_1, n_2,..., n_k)>$, при этом часть, находящаяся справа от этой строчки, пустая;
- если $f(n_1, n_2,..., n_k)$ не определено, то M никогда не останавливается.

Заметим, что имеется бесконечное множество машин Тьюринга, для каждой вычислимой функции своя. Более того, для любой вычислимой функции имеется бесконечное множество машин Тьюринга, вычисляющих эту функцию.

Пример. Построим машину Тьюринга, вычисляющую сумму n_1+n_2.

Зададим функцию M следующим образом:

$M(0, 1) = <1, П, 0>;$
$M(0, 0) = <1, П, 1>;$
$M(1, 1) = <1, П, 1>;$
$M(1, 0) = <0, Л, 2>;$
$M(2, 1) = <0, Л, 3>;$
$M(3, 1) = <0, Л, 4>;$
$M(4, 1) = <1, Л. 4>;$
$M(4, 0) = <0, П, 5>.$

Посмотрим, как происходит сложение 1+1. В текущей строке символов обозреваемый символ выделен.

номер инструкции	текущая строка символов	комментарий
0	0***1***10110	прохождение через первое слагаемое
0	01***1***0110	
0	011***0***110	заполнение промежутка
1	0111***1***10	прохождение через второе слагаемое
1	01111***1***0	
1	011111***0***	конец второго слагаемого
2	01111***1***0	стирание 1
3	0111***1***00	стирание второй 1
4	011***1***000	движение назад
4	01***1***1000	
4	0***1***11000	
4	***0***111000	остановка
5	0***1***11000	

Мы должны заметить, что многие детали нашего определения машины Тьюринга до некоторой степени произвольны. Если бы было более одной ленты, то класс вычислимых функций остался бы неизменным, хотя некоторые функции могли бы быть вычислены более быстро. Аналогично, мы могли бы допускать больше символов, чем 0 и 1, или же у нас могла бы быть лента, бесконечная только в одну сторону от начальной точки, вместо имеющейся бесконечной в обоих направлениях. Ни одно из этих изменений не затрагивает класса вычислимых функций. Что действительно существенно в этом

определении – это разрешение произвольно большого количества материала для запоминающего устройства и произвольно длинных вычислений.

2 Вычислимое и невычислимое

2.1 Тезис Чёрча

За последние 70 лет было предложено много различных математических уточнений интуитивного понятия алгоритма. три из этих подхода мы разобрали. Перечислим некоторые другие альтернативные способы, которые предлагались следующими авторами:

- Гёдель–Эбран–Клини. Общерекурсивные функции, определенные с помощью исчисления рекурсивных уравнений [17, с. 261–278].
- Пост. Функции, определяемые каноническими дедуктивными системами [11, с. 66–72].
- Марков. Функции, задаваемые некоторыми алгоритмами (известные под названием нормальными алгоритмами) над конечным алфавитом [17, с.228–250].
- Шепердсон–Стерджис. МНР–вычислимые функции [11].

Между этими подходами (в том числе, и три выше рассмотренных) имеются большие различия; каждый из них имеет свои преимущества для соответствующего описания вычислимости. Следующий замечательный результат получен усилиями многих исследователей.

Теорема 5 (Основной результат). [11, с. 57]. Каждое из вышеупомянутых уточнений эффективной вычислимости приводит к одному и тому же классу вычислимых функций.

Вопрос: насколько хорошо неформальное и интуитивное понятие эффективно вычислимой функции отражено в различных формальных описаниях?

Чёрч, Тьюринг и Марков каждый в соответствии со своим подходом выдвинули утверждение (тезис) о том, что класс определенных ими функций совпадает с неформально определенным классом вычислимых функций. В силу основного результата все эти утверждения логически эквивалентны. Название *тезис Чёрча* теперь применяется к этим и аналогичным им утверждениям.

Тезис Чёрча

Интуитивно и неформально определенный класс эффективно вычислимых функций совпадает с классом частично-рекурсивных функций.

Сразу же заметим, что этот тезис не является теоремой, а скорее утверждение, которое принимается на веру, причем вера подкрепляется следующими аргументами [11, с. 75–76].

- Фундаментальный результат: многие независимые инварианты уточнения интуитивного понятия вычислимости привели к одному и тому же классу функций.
- Обширное семейство эффективно вычислимых функций принадлежит этому классу. Конкретные функции, рассмотренные в 1.2, образуют исходную часть этого семейства, которую можно расширять до бесконечности методами из 1.2 или более мощными и сложными методами.
- Никто еще не нашел функцию, которую можно было признать вычислимой в неформальном смысле, но которую нельзя было бы построить, используя один из формальных методов.

Приведем парадокс, показывающий насколько мы подходим к опасной границе между непротиворечивым и противоречивым в обсуждении вычислимых функций [12, с. 184].

Легко доказать, что множество всюду определенных вычислимых функций f: $\boldsymbol{N}\to\boldsymbol{N}$ является перечислимым, т. е. их можно перенумеровать в виде последовательности $f_1, f_2, f_3, \ldots$.

Определим теперь новую функцию g формулой

$$g(n) = f_n(n)+1.$$

Она не входит в нашу последовательность, поскольку при n=1 она отличается от f_1, при n=2 – от f_2 и т. д. Следовательно, она не вычислима.

С другой стороны, ясно, что она вычислима, так как $f_n(n)$ вычислима, а прибавив 1 к $f_n(n)$, мы получим $g(n)$.

Этот парадокс можно объяснить. На самом деле мы здесь используем две формальные системы. В рамках одной системы (скажем, элементарной арифметики *EA*), мы описываем вычислимые функции f, а то, что g является вычислимой функцией мы получаем в рамках другой формальной системы, в которой уже используется возможность упорядочить f. Вторая формальная система является надсистемой или метасистемой относительно первой.

2.2 Теорема о рекурсии

После изучения эффективно вычислимых операций на множестве *чисел*, естественно рассмотреть вопрос о том, существует ли понятие подобного рода и для операций над *функциями*. Существенная разница между функциями и числами (если их рассматривать как множества основных объектов) состоит в том, что первые являются, как правило, бесконечными объектами, а вторые – конечными.

Для простоты изложения мы будем рассматривать частичные функции только одного аргумента, но все определения и полученные результаты обобщаются и на случай функций от нескольких аргументов.

Обозначим через $\boldsymbol{F}$ класс всех частичных функций из $\boldsymbol{N}$ в $\boldsymbol{N}$. Словом **оператор** мы будем обозначать функцию Ф: $\boldsymbol{F}\to\boldsymbol{F}$, причем будем рассматривать лишь такие операторы Ф, область определения которых совпадает со всем классом $\boldsymbol{F}$.

Основная проблема при попытке дать определение понятию вычислимого (или эффективного) оператора Ф: $\boldsymbol{F}\to\boldsymbol{F}$ состоит в том, что как вводимая функция f, так и выводимая функция Ф(f) являются, скорее всего, бесконечными объектами и, следовательно, не могут быть заданы за конечное время. В то же время, согласно нашему интуитивному представлению об алгоритмических процессах, мы должны в некотором смысле уметь проводить их за конечное время.

Чтобы разобраться, как можно преодолеть эту трудность, рассмотрим следующие операторы из $\boldsymbol{F}$ в $\boldsymbol{F}$:

1) $\Phi_1(f) = 2f$,
2) $\Phi_2(f) = g$, где $g(x) = \sum_{y\le x} f(y)$.

Эти операторы, несомненно, очень естественны и записаны в явном виде. На интуитивном уровне их вполне можно было бы считать вычислимыми. Попытаемся сейчас разобраться, почему. Пусть $f\in\boldsymbol{F}$. Рассмотрим случай $g_1= \Phi_1(f)$. Заметим, что любое конкретное значение $g_1(x)$ (если оно определено) может быть вычислено за конечное время по одному единственному значению $f(x)$ функции f. В случае $g_2= \Phi_2(f)$ для вычисления значения $g_2(x)$ (если оно определено) необходимо знать *конечное* число значений $f(0), f(1), \ldots, f(x)$. Тем самым в обоих случаях любое определенное значение выводимой функции ($\Phi_1(f)$ или $\Phi_2(f)$) может быть эффективно вычислено за *конечное* время при использовании лишь *конечной* информации о вводимой функции f. В этом и состоит суть приведенного ниже определения рекурсивного оператора.

Для того чтобы дать определению точный смысл, условимся о некоторых технических деталях. Если f, g – функции, то будем говорить, что $\boldsymbol{g}$ **продолжает** $\boldsymbol{f}$, если $D_f \subseteq D_g$ и $f(x) = g(x)$ для всех $x\in D_f$. Это отношение функций f, g записывается как $f\subseteq g$.

Под «конечной частью» функции f понимается некоторая конечная функция θ, которая продолжается до f. (Функция θ называется *конечной*, если её область определения есть конечное множество.)

Из приведенных выше соображений ясно, что в определении вычислимого оператора должны участвовать эффективные вычисления с конечными функциями. Придадим этому точный смысл, закодировав каждую конечную функцию θ натуральным числом $g(\theta)$. Удобный для наших целей кодирующий алгоритм определяется следующим образом.

Пусть p_k (k = 0, 1, 2,…) обозначает последовательность простых чисел 2, 3, 5, 7, 11, … Для данного $\theta \in \boldsymbol{F}$ положим

$$g(\theta) = \prod_k p_k^{\theta(k)+1}, \text{ где } k \in D_\theta \neq \varnothing,$$

$$g(\theta) = 0, \text{ если } D_\theta = \varnothing.$$

Имеется простой алгоритм, с помощью которого можно для любого числа z определить, верно ли $z = g(\theta)$ для некоторой конечной функции θ, и, если ответ на вопрос положителен, выяснить, принадлежит ли некоторое заданное x множеству D_θ, и если да, то вычислить значение $\theta(x)$.

Сформулируем теперь наше определение.

Рекурсивный оператор

Оператор Φ: $\boldsymbol{F} \to \boldsymbol{F}$ называется **рекурсивным** (или **вычислимым**), если существует вычислимая функция $\phi(z, x)$, такая, что для всех $f \in \boldsymbol{F}$ и $x, y \in \boldsymbol{N}$

$\Phi(f)(x) = y$ тогда и только тогда, когда существует конечная функция $\theta \subseteq f$, такая, что $\phi(g(\theta), x) = y$.

Число $g(\theta)$ – кодовый номер функции θ.

Пример. Оператор $\Phi(f) = 2f$ является рекурсивным оператором. Чтобы убедиться в этом, положим

$$\phi(z, x) = \begin{cases} 2\theta(x), \text{ если } z = g(\theta) \text{ и } x \in D_\theta, \\ \text{не определена в остальных случаях.} \end{cases}$$

В силу тезиса Чёрча функция ϕ вычислима. Далее, для любых f, x, y мы имеем $\Phi(f)(x) = y \Leftrightarrow x \in D_f$ и $y = 2f(x) \Leftrightarrow$ существует $\theta \subseteq f$, для которой $x \in D_\theta$ и $y = 2\theta(x) \Leftrightarrow$ существует $\theta \subseteq f$, для которой $\phi(g(\theta), x) = y$. Отсюда Φ – рекурсивный оператор.

Важное свойство рекурсивных операторов состоит в том, что они **непрерывны** и **монотонны** в следующем смысле.

Непрерывный оператор

Оператор Φ: $\boldsymbol{F} \to \boldsymbol{F}$ называется непрерывным, если для любой функции $f \in \boldsymbol{F}$ и любых x, y: $\Phi(f)(x) = y$ тогда и только тогда, когда существует конечная функция $\theta \subseteq f$, такая, что $\Phi(\theta)(x) = y$.

Монотонный оператор

Оператор Φ: $\boldsymbol{F} \to \boldsymbol{F}$ называется монотонным, если для любых функций $f, h \in \boldsymbol{F}$ таких, что $f \subseteq h$, выполняется условие $\Phi(f) \subseteq \Phi(h)$.

Теорема 6. Любой рекурсивный оператор является непрерывным и монотонным. Доказательство см., например, [11, с. 194–195].

Неформально использование слова «непрерывный» для описания свойства оператора можно объяснить следующим образом. Можно показать, что если функция h находится «близко» от функции f (в том смысле, что они совпадают на конечном множестве D_θ, где $\theta \subseteq f$), то для непрерывного оператора Ф функция $\Phi(h)$ расположена «близко» от функции $\Phi(f)$.

Первая теорема о рекурсии Клини представляет собой теорему о неподвижной точке для рекурсивных операторов, и ее зачастую называют также теоремой о неподвижной точке (из теории рекурсии).

Теорема 7 (теорема о неподвижной точке).[1] Пусть Ф: $\boldsymbol{F \to F}$ – рекурсивный оператор. Тогда существует вычислимая функция f, которая является наименьшей неподвижной точкой для Ф, т.е

1) $\Phi(f) = f$,
2) если $\Phi(h) = h$, то $f \subseteq h$.

Отсюда если функция f всюду определена, то она является единственной неподвижной точкой Ф.

Доказательство см., например, [11, с. 203–204].

Теорема вызывает недоумение: пусть оператор Ф задан так: функция $f(x)$ отображается в функцию $f(x)+1$. Это отображение конечно можно задать с помощью алгоритма (т.е. оператор рекурсивный), но разве у этого оператора есть неподвижная точка? Да, это функция, которая нигде не определена.

Наш программистский опыт говорит о том, что определения, даваемые в терминах примитивной рекурсии, имеют смысл. В случае более сложных рекурсивных определений (см., например, функцию Аккермана) это не так очевидно – вполне можно представить себе, что функций, удовлетворяющих тому или иному определению, вообще не существует. Именно здесь оказывается полезной теорема 7. Рекурсивные определения весьма общего типа могут быть представлены уравнением вида

$$f = \Phi(f),$$

где Ф – некоторый рекурсивный оператор. Теорема о неподвижной точке показывает, что такое определение имеет смысл: всегда существует даже вычислимая функция, которая ему удовлетворяет. Поскольку в математике требуется, чтобы определения описывали различные понятия однозначно, можно сказать, что данное рекурсивное определение определяет наименьшую неподвижную точку оператора Ф. Таким образом, согласно теореме 7, класс вычислимых функций замкнут относительно рекурсивного определения очень общего типа. Таким образом, понятно, почему эту теорему называют теоремой о рекурсии (точнее первой теоремой о рекурсии, есть также вторая теорема о рекурсии).

2.3 Куины

Куином называется программа, которая на своем выходе генерирует собственный исходный текст. Этот термин предложил Д. Хофштадтер в честь американского философа и математика *Willard van Orman Quine* (1908 – 2000).

Докажем, что куины существуют для любого универсального языка программирования. Действительно, теорему Клини о рекурсии (или о неподвижной точке) можно сформулировать в следующем виде: «Нельзя найти алгоритма, преобразующего программы, который бы по каждой программе давал бы другую (не эквивалентную ей)».

Пусть теперь дано преобразование программ:

F: $p \to$ {программа, которая на любом входе печатает p).

[1] В ламбда-исчислении этой теореме соответствует утверждение о существование комбинатора неподвижной точки.

В силу теоремы Клини о рекурсии существует программа p такая, что $F(p) = p$, т. е. существует программа, печатающая (на любом входе) свой собственный текст.

Написать куин – хорошая и популярная задача для любителей программирования. Неформальную инструкцию на русском языке можно представить следующим образом:

напечатать два раза, второй раз в кавычках, такой текст: «напечатать два раза, второй раз в кавычках, такой текст:»

В наиболее чистом виде явление самовоспроизведения можно выразить в лямбда–исчислении. Известный комбинатор $(\lambda x.\ x\ x)\ (\lambda x.\ x\ x)$ редуцируется к самому себе

$$(\lambda x.\ x\ x)\ (\lambda x.\ x\ x) \rightarrow (\lambda x.\ x\ x)\ (\lambda x.\ x\ x).$$

Этому комбинатору соответствует в Лиспе форма – ламбда-выражение, примененное к такому же ламбда-выражению с формой quote, т. е. это лямбда-вызов:

((lambda (x) (list x (list ‘quote x))) ‘(lambda (x) (list x (list ‘quote x))))

Форма является одновременно автоаппликативной! Она содержит изображение самой себя, но не целиком, поскольку такое изображение было бы бесконечно длинным. Изображение получается ограниченным благодаря механизму блокировки и ламбда-механизму, а также благодаря автоаппликации.

Достаточно просто написать куин на языке Haskell (автор Jon Fairbairn: Jon.Fairbairn@cl.cam.ac.uk):

```
main = putStr (quine q)
quine s = s ++ show s
q = "main = putStr (quine q)\nquine s = s ++ show s\nq = "
```

Таким образом, куин на языке Haskell полностью повторяет неформальную инструкцию на русском языке.

Написание куинов на процедурных языках требует некоторых хитростей. Вот пример на Паскале:

```
var a:array[1..7] of string;
    i:integer;
begin
a[1]:='var a:array[1..7] of string;';
a[2]:='    i:integer;';
a[3]:='begin';
a[4]:='for i:=1 to 3 do writeln(a[i]);';
a[5]:='for i:=1 to 7 do writeln(#97#91,i,#93#58#61#39,a[i],#39#59);';
a[6]:='for i:=4 to 7 do writeln(a[i]);';
a[7]:='end.';
for i:=1 to 3 do writeln(a[i]);
for i:=1 to 7 do writeln(#97#91,i,#93#58#61#39,a[i],#39#59);
for i:=4 to 7 do writeln(a[i]);
end.
```

Чтобы понять эту программу полезно иметь в виду соответствие между символами и кодами:

a	[	]	:	=	;	'
97	91	93	58	61	59	39

Предложенный текст куина на Паскале имеет существенный недостаток: он использует значения кодов ASCII – но коды не являются частью самого языка Паскаль. Следующий куин исправляет этот недостаток (автор Dan Hoey: hoey@aic.nrl.navy.mil):

```
program s;const bbb='program s;const bbb';a='a';b='b';bb=');writeln(';
aa='''';ab='=''';ba=''';';
aaa='begin writeln(bbb,ab,bbb,ba,a,ab,a,ba,b,ab,b,ba,b,b,ab,bb,ba';
aba='a,a,ab,aa,aa,ba,a,b,ab,ab,aa,ba,b,a,ab,aa,ba,ba';
abb='a,a,a,ab,aaa,ba);writeln(a,b,a,ab,aba,ba);writeln(a,b,b,ab,abb,ba';
baa='b,a,a,ab,baa,ba);writeln(b,a,b,ab,bab,ba);writeln(aaa,bb';
bab='aba,bb);writeln(abb);writeln(bb,baa);writeln(bb,bab)end.';
begin writeln(bbb,ab,bbb,ba,a,ab,a,ba,b,ab,b,ba,b,b,ab,bb,ba);writeln(
a,a,ab,aa,aa,ba,a,b,ab,ab,aa,ba,b,a,ab,aa,ba,ba);writeln(
a,a,a,ab,aaa,ba);writeln(a,b,a,ab,aba,ba);writeln(a,b,b,ab,abb,ba
);writeln(b,a,a,ab,baa,ba);writeln(b,a,b,ab,bab,ba);writeln(aaa,bb
);writeln(aba,bb);writeln(abb);writeln(bb,baa);writeln(bb,bab)end.
```

Как видим, вопреки общепринятому мнению, куин на процедурном языке не обязан содержать итерацию или рекурсию.

Множество куинов на различных языках можно найти на сайте:
Куины: www.nyx.net/~gthompso/quine.htm

2.4 Некоторые алгоритмически неразрешимые проблемы

> ... Найти задачу – не меньшая радость, чем отыскать решение.
>
> Томас де Куинси

> – Это же проблема Бен Бецалеля. Калиостро же доказал, что она не имеет решения... Как же искать решения, когда его нет? Бессмыслица какая-то...
> – Бессмыслица – искать решение, если оно и так есть. Речь идет о том, как поступать с задачей, которая решения не имеет.
>
> А. и Б. Стругацкие.
> Понедельник начинается в субботу

Решение вопроса о том, обладают ли натуральные числа данным свойством, является часто встречающейся задачей математики. Поскольку свойства чисел можно выразить с помощью подходящего предиката, то решение задачи сводится к выяснению того, является ли данный предикат **разрешимым** или нет (т.е. существует ли алгоритм, который позволил бы распознать, является ли предикат истинным или ложным; см. раздел 5.1). Задачи с произвольными универсумами во многих случаях можно переформулировать в виде задач с натуральными числами, если использовать подходящее кодирование.

В контексте разрешимости предикаты часто называются **проблемами**.

Имея точное определение вычислимости, удалось доказать, что некоторые проблемы неразрешимы.

Теорема 8 (теорема Черча о неразрешимости логики предикатов. (см. раздел 4). Не существует алгоритма, который для любой формулы логики предикатов устанавливает, общезначима она или нет.

Теорема 9. Проблема остановки неразрешима. Не существует никакого общего алгоритма, позволяющего установить, остановится ли некоторая конкретная программа

(на любом языке программирования), запущенная после введения в неё некоторого конкретного набора данных. Смысл этого утверждения для теоретического программирования очевиден: не существует совершенно общего метода проверки программ на наличие в них бесконечных циклов.

Приведем два варианта доказательства. Первый вариант является эскизом доказательства [20, стр. 123-125], и он следует первоначальному доказательству А. Тьюринга.

Второй вариант – строгое доказательство, сделанное в рамках ламбда–исчисления.

Доказательство 1. Не теряя общности, мы можем рассматривать только программы, вычисляющие одноместные функции от натуральных чисел. Поскольку вычислимых одноместных функций – счетное число, то пусть последовательность c_1, c_2, c_3,… – содержит все такие функции. Доказательство проведем от противного.

Пусть имеется такая программа (двуместная функция) $A(q,n)$, что

1. Если завершается $A(q,n)$, то не завершается $C_q(n)$.

(Т.е. мы считаем, что алгоритм A проверяет, остановится или нет вычисление функции.)
Тогда, в частности, для $q=n$

2. Если завершается $A(n,n)$, то не завершается $C_n(n)$.

Функция $A(n,n)$ зависит от одного параметра, следовательно, существует такое k, что

3. $A(n,n) = C_k(n)$.

Тогда

4. $A(k,k) = C_k(k)$ (следует из (3) при $n=k$).

Имеем из (2) при $n=k$

5. Если завершается $A(k,k)$, то не завершается $C_k(k)$.

Подставляя (4) в (5), получаем

6. Если завершается $C_k(k)$, то не завершается $C_k(k)$.

Однозначное следствие:
$C_k(k)$ не завершается, а $A(k,k)$ не может это установить (поскольку $A(k,k) = C_k(k)$ и, следовательно, не останавливается).

Доказательство 2. Поскольку вычислимые функции – это в точности те, которые ламбда–представимы (теорема Клини), то проблему остановки можно сформулировать в терминах ламбда–исчисления: «Существует ли алгоритм, с помощью которого можно установить, имеет ли нормальную форму данный ламбда–терм?».

Доказывать будем от противного. Предположим, что искомый алгоритм существует: если терм t имеет нормальную форму, то алгоритм выдает 0, а если нет, то алгоритм выдает 1. В силу тезиса Черча такой алгоритм представляется в виде ламбда–определимой функции, следовательно, существует комбинатор H такой, что

$$Ht = \begin{cases} <0>, \text{åñëè òåðì } t \text{ èìååò íîðìàëüíóþ ôîðìó}, \\ <1>, \text{â ïðîòèâíîì ñëó÷àå} \end{cases}$$

Комбинаторы <0> и <1> – произвольные нумералы, кодирующие соответственно 0 и 1. Обозначим через ***zero?*** , как обычно, комбинатор – тест на ноль:

$$\textbf{\textit{zero?}}\ <0> \rightarrow \textbf{\textit{true,}}$$
$$\textbf{\textit{zero?}}\ <1> \rightarrow \textbf{\textit{false.}}$$

Как обычно, ***true*** ≡ ***K***, ***false*** ≡ ***KI***.

Мы используем комбинатор H, чтобы построить комбинатор D такой, что

$$Dt = \begin{cases} K(tt), \text{åñëè òåðì } (tt) \text{ èìååò íîðìàëüíóþ ôîðìó}, \\ S, \text{â ïðîòèâíîì ñëó÷àå} \end{cases}$$

Он строится очевидным образом:

$$D \equiv \lambda t.\ \textbf{\textit{zero?}}\ (H(tt))(\boldsymbol{K}(tt))\boldsymbol{S}.$$

Действительно, пусть терм (*tt*) имеет нормальную форму, тогда

Dt
$\equiv (\lambda t.\ \textbf{\textit{zero?}}\ (H(tt))(\boldsymbol{K}(tt))\boldsymbol{S})\ t$
$= \textbf{\textit{zero?}}\ (H(tt))(\boldsymbol{K}(tt))\boldsymbol{S}$
$= \textbf{\textit{zero?}}\ \langle 0\rangle\ (\boldsymbol{K}(tt))\boldsymbol{S}$
$= \textbf{\textit{true}}\ (\boldsymbol{K}(tt))\boldsymbol{S}$
$\equiv \boldsymbol{K}\ (\boldsymbol{K}(tt))\boldsymbol{S}$
$= (\boldsymbol{K}(tt)$

Если же терм (*tt*) не имеет нормальную форму, то

Dt
$\equiv (\lambda t.\ \textbf{\textit{zero?}}\ (H(tt))(\boldsymbol{K}(tt))\boldsymbol{S})\ t$
$= \textbf{\textit{zero?}}\ (H(tt))(\boldsymbol{K}(tt))\boldsymbol{S}$
$= \textbf{\textit{zero?}}\ \langle 1\rangle\ (\boldsymbol{K}(tt))\boldsymbol{S}$
$= \textbf{\textit{false}}\ (\boldsymbol{K}(tt))\boldsymbol{S}$
$\equiv (\boldsymbol{KI}\ )(\boldsymbol{K}(tt))\boldsymbol{S}$
$= \boldsymbol{IS}$
$= \boldsymbol{S}$

Для любого *t* терм *Dt* имеет нормальную форму (поскольку он равен либо $\boldsymbol{K}(tt)$, имеющего нормальную форму, либо равен $\boldsymbol{S}$), поэтому терм *DD* также имеет нормальную форму (скажем некий комбинатор α). Согласно определению D, комбинатор DD равен K(DD), т.е. $D = \alpha$ и $D = K\alpha \rightarrow \lambda y.\alpha$. Но комбинаторы α и $\lambda y.\alpha$ оба находятся в нормальной форме и не совпадают. Получили противоречие.

Используя аналогичные идеи получены и следующие результаты о неразрешимости. Не существует никакого общего алгоритма, позволяющего установить, вычисляет ли некоторая конкретная программа (на любом языке программирования) постоянную нулевую функцию [11, с. 110]. То же самое справедливо и для любой другой конкретной вычислимой функции. И как следствие, можно утверждать, что вопрос о том, вычисляют ли две данные программы одну и ту же одноместную функцию, также неразрешим. Тем самым, мы получаем, что в области тестирования компьютерных программ, мы имеем принципиальные ограничения.

Диофантовы уравнения [11, с. 114]. Пусть $p(x_1, x_2, ..., x_n)$ – многочлен от переменных $x_1, x_2, ..., x_n$ с целыми коэффициентами. Тогда уравнение

$p(x_1, x_2, ..., x_n) = 0,$

для которого мы ищем только целые решения, называется **диофантовым уравнением**. Первым диофантовые уравнения систематизировал и изучил греческий математик Диофант в третьем веке нашей эры. Ниже приводится пример системы диофантовых уравнений:

$$\begin{cases} 6w + 2x^2 - y^3 = 0, \\ 5xy - z^2 + 6 = 0, \\ w^2 - w + 2x - y + z - 4 = 0. \end{cases}$$

Вот еще один пример:

$$\begin{cases} 6w + 2x^2 - y^3 = 0, \\ 5xy - z^2 + 6 = 0, \\ w^2 - w + 2x - y + z - 3 = 0. \end{cases}$$

Решением первой системы является, в частности, следующее:

$$w=1,\ x=1,\ y=2,\ z=4,$$

тогда как вторая система вообще не имеет решения. В самом деле, судя по первому уравнению, число y должно быть четным, судя по второму уравнению, число z также должно быть четным, однако это противоречит третьему уравнению, причем при любом w, поскольку значение разности $w^2–w$ – это всегда четное число, а число 3 нечетно.

Десятая проблема Гильберта, сформулированная в 1900 году, состоит в том, чтобы установить, существует ли алгоритм, с помощью которого можно было бы проверить, имеет ли данное диофантово уравнение решение. В 1970 году советский математик Ю. Матиясевич доказал, что такого алгоритма не существует. Доступное доказательство этого можно найти в [15].

Занимательному изложению вопросов вычислимости, вплоть до получения доказательства теоремы Гёделя, посвящена книга Р. Смаллиана [23].

...эмпирические системы утрачивают свою актуальность, математические же – никогда. Их бессмертие – в их «пустоте».

Станислав Лем,
Сумма технологии

3 Формальные аксиоматические теории

Формальная теория представляет собой множество чисто абстрактных объектов (не связанных с внешним миром), в которой представлены правила оперирования множеством символов в чисто синтаксической трактовке без учета смыслового содержания (или семантики).

Исторически понятие формальной теории было разработано в период интенсивных исследований в области оснований математики для формализации собственно логики и теории доказательства. Сейчас этот аппарат широко используется при создании специальных исчислений для решения конкретных прикладных задач. Одним из таких приложений является логическое программирование.

Формальную теорию иногда называют **аксиоматикой** или **формальной аксиоматической теорией**. Родоначальником аксиоматических теорией можно считать «Начала» Евклида.

3.1 Основные определения

Что такое формальная теория?

Формальная теория ***T*** считается определенной, если:

- задано некоторое счетное множество ***A*** символов – символов теории ***T***; конечные последовательности символов теории ***T*** называются **выражениями** теории ***T*** (множество выражений обозначают через ***A****);
- имеется подмножество $F \subset A^*$ выражений теории ***T***, называемых **формулами** теории ***T***;
- выделено некоторое множество $B \subset F$ формул, называемых **аксиомами** теории ***T***;
- имеется конечное множество $\{R_1, R_2, \ldots, R_m\}$ отношений между формулами, называемых **правилами вывода**. Правила вывода позволяют получать из некоторого конечного множества формул другое множество формул.

Множество символов ***A*** – алфавит теории – может быть конечным или бесконечным. Обычно для образования символов используют конечное множество букв, к которым, если нужно, приписывают в качестве индексов натуральные числа.

Множество формул ***F*** обычно задается индуктивным определением, например, с помощью формальной грамматики. Как правило, это множество бесконечно. Множества ***A*** и ***F*** в совокупности определяют **язык формальной теории**.

Множество аксиом ***B*** может быть конечным или бесконечным. Если множество аксиом бесконечно, то, как правило, оно задается с помощью конечного множества **схем аксиом** и правил порождения конкретных аксиом из схемы аксиом. Обычно аксиомы делятся на два вида: **логические** аксиомы (общие для целого класса формальных теорий) и **нелогические** (или **собственные**) аксиомы (определяющие специфику и содержание конкретной теории).

Обычно для формальной теории имеется алгоритм, позволяющий по данному выражению определить, является ли оно формулой. Точно так же чаще всего существует алгоритм, выясняющий, является ли данная формула теории ***T*** аксиомой; в таком случае ***T*** называется **эффективно аксиоматизированной** теорией.

Выводимость

Пусть $A_1, A_2,\ldots, A_n, A$ – формулы теории ***T***. Если существует такое правило вывода R, что $< A_1, A_2,\ldots, A_n, A> \in R$, то говорят, что формула A **непосредственно выводима** из формул $A_1, A_2,\ldots, A_n$ по правилу вывода R. Обычно этот факт записывают следующим образом:

$\frac{A_1, A_2, ..., A_n}{A} R$,

где формулы $A_1, A_2,..., A_n$ называются **посылками**, а формула A – **заключением**.

Выводом формулы A из множества формул Γ в теории ***T*** называется такая последовательность $F_1, F_2,..., F_k$, что $A = F_k$, а любая формула F_i ($i < k$) является либо аксиомой, либо $F_i \in \Gamma$, либо непосредственно выводима из ранее полученных формул $F_{j_1},..., F_{j_n}$ ($j_1,..., j_n < i$). Если в теории ***T*** существует вывод формулы A из множества формул Γ, то это записывается следующим образом:

$\Gamma \vdash_{\boldsymbol{T}} A$,

где формулы из Γ называются **гипотезами** вывода. Если теория ***T*** подразумевается, то её обозначение обычно опускают.

Если множество Γ конечно: $\Gamma = \{B_1, B_2,..., B_n\}$, то вместо $\{B_1, B_2,..., B_n\} \vdash A$ пишут $B_1, B_2,..., B_n \vdash A$. Если Γ есть пустое множество ∅, то A называют **теоремой** (или **доказуемой** формулой) и в этом случае используют сокращенную запись $\vdash A$ («A есть теорема»).

Отметим, что в соотношении {теоремы} ⊂ {формулы} ⊂ {выражения} включение множеств является строгим.

Приведем несколько простых свойств понятия выводимости из посылок.

1. Если $\Gamma \subseteq \Sigma$ и $\Gamma \vdash A$, то $\Sigma \vdash A$.

Это свойство выражает тот факт, что если A выводимо из множества гипотез Γ, то оно остается выводимым, если мы добавим к Γ новые гипотезы.

2. $\Gamma \vdash A$ тогда и только тогда, когда в Γ существует конечное подмножество Σ, для которого $\Sigma \vdash A$.

Часть «тогда» утверждения 2 вытекает из утверждения 1. Часть «только тогда» этого утверждения очевидна, поскольку всякий вывод A из Γ использует лишь конечное число гипотез из Γ.

3. Если $\Sigma \vdash A$ и $\Gamma \vdash B$ для любого B из множества Σ, то $\Gamma \vdash A$.

Смысл этого утверждения прост: если A выводимо из Σ и любая формула из Σ выводима из Γ, то A выводима из Γ.

Понятие формальности можно определить в терминах теории алгоритмов: теорию ***T*** можно считать формальной, если построен алгоритм (механически применяемая процедура вычисления) для проверки правильности рассуждений с точки зрения принципов теории ***T***. Это значит, что если некто предлагает математический текст, являющийся, по его мнению, доказательством некоторой теоремы в теории ***T***, то, механически применяя алгоритм, мы можем проверить, действительно ли предложенный текст соответствует стандартам правильности, принятым в ***T***. Таким образом, стандарт правильности рассуждений для теории Т определен настолько точно, что проверку его соблюдения можно передать вычислительной машине (следует помнить, что речь идет о **проверке правильности** готовых доказательств, а не об их поиске!). Если проверку правильности доказательств в какой-либо теории нельзя передать вычислительной машине, и она доступна в полной мере только человеку, значит, еще не все принципы теории аксиоматизированы (то, что мы не умеем передать машине, остается в нашей интуиции и «оттуда» регулирует наши рассуждения).

В качестве несерьезного примера формальной теории можно рассматривать игру в шахматы – назовем это теорией ***Ch***. Формулами в ***Ch*** будем считать **позиции**

(всевозможные расположения фигур на доске вместе с указанием «ход белых» или «ход черных»). Тогда аксиомой теории ***Ch*** естественно считать **начальную позицию**, а правилами вывода – **правила игры**, которые определяют, какие ходы допустимы в каждой позиции. Правила позволяют получать из одних формул другие. В частности, отправляясь от нашей единственной аксиомы, мы можем получать теоремы ***Ch***. Общая характеристика теорем ***Ch*** состоит, очевидно, в том, что это – всевозможные позиции, которые могут получиться, если передвигать фигуры, соблюдая правила.

В чем выражается формальность теории ***Ch***? Если некто предлагает нам «математический текст» и утверждает, что это – доказательство теоремы *A* в теории ***Ch***, то ясно, что речь идет о непроверенной записи шахматной партии, законченной (или отложенной) в позиции *A*. Проверка не является, однако, проблемой: правила игры сформулированы настолько точно, что можно составить программу для вычислительной машины, которая будет осуществлять такие проверки. (Еще раз напомним, что речь идет о проверке правильности записи шахматной партии, а не о проверке того, можно ли заданную позицию получить, играя по правилам, – эта задача намного сложнее!)

Несколько серьезнее другой пример формальной теории. Пусть {*a*, *b*} есть алфавит теории ***L***. Формулами в теории ***L*** являются всевозможные цепочки, составленные из букв *a*, *b*, например *a*, *aa*, *aba*, *abaab*. Единственной аксиомой ***L*** является цепочка *a,* наконец, в ***L*** имеется два правила вывода:

$$\frac{X}{Xb} \quad и \quad \frac{X}{aXa}.$$

Такая запись означает, что в теории ***L*** из цепочки *X* непосредственно выводятся *Xb* и *aXa*. Примером теоремы ***L*** является цепочка *aababb*; вывод для нее есть

a, *ab*, *aaba*, *aabab*, *aababb.*

Формальные теории являются не просто игрой ума, а всегда представляют собой модель какой-то реальности (либо конкретной, либо математической). Вначале математик изучает реальность, конструируя некоторое абстрактное представление о ней, т. е. некоторую формальную теорию. Затем он доказывает теоремы этой формальной теории. Вся польза и удобство формальных теорий как раз и заключается в их абстрагировании от конкретной реальности. Благодаря этому одна и та же формальная теория может служить моделью многочисленных конкретных ситуаций. Наконец, он возвращается к исходной точке всего построения и дает интерпретацию теорем, полученных при формализации.

При изучении формальных теорий нужно различать теоремы формальной теории и теоремы *о* формальной теории, или **метатеоремы**. Это различие не всегда явно формализуется, но всегда является существенным.

Множество теорем формальной теории является точно определенным объектом (обычно бесконечным) и поэтому можно доказывать утверждения, относящиеся ко всем теоремам одновременно. Например, в теории ***Ch*** множество всех теорем оказывается, правда, конечным (хотя конечность эта с практической точки зрения ближе к бесконечности). Легко доказать следующее утверждение, относящееся ко *всем* теоремам ***Ch***: ни в одной теореме белые не имеют 10 ферзей. В самом деле, достаточно заметить, что в аксиоме ***Ch*** белые имеют одного ферзя и восемь пешек и что по правилам игры белым ферзем может стать только белая пешка. Остальное решает арифметика: 1+8<10. Таким образом, мы подметили в системе аксиом и правил вывода теории ***Ch*** особенности, которые делают справедливым наше общее утверждение о теоремах ***Ch***.

Аналогичные возможности имеем в случае теории ***L***. Можно доказать, например, следующее утверждение, относящееся ко всем теоремам ***L***: если *X* – теорема, то *aaX* – тоже теорема.

Рассмотрим, как соотносятся неформальные доказательства и логический вывод [27]. Логический вывод напоминает процесс мышления, но при этом мы не должны равнять его правила с правилами человеческой мысли. Доказательство – это нечто неформальное; иными словами – это продукт нормального мышления, записанный на человеческом языке и предназначенный для человеческого потребления. В доказательствах могут использоваться всевозможные сложные мыслительные приемы, и хотя интуитивно они могут казаться верными, можно усомниться в том, возможно ли доказать их логически. Именно поэтому мы нуждаемся в формализации. Вывод – это искусственное соответствие доказательства: его назначение – достичь той же цели, на этот раз с помощью логической структуры, методы которой не только ясно выражены, но и очень просты.

Обычно формальный вывод бывает крайне длинен по сравнению с соответствующей «естественной» мыслью. Это, конечно, плохо – но это та цена, которую приходится платить за упрощение каждого шага. Часто бывает, что вывод и доказательство «просты» в дополнении друг к другу. Доказательство просто в том смысле, что каждый шаг «кажется правильным», даже если мы и не знаем точно, почему; логический вывод прост, потому что каждый из мириада её шагов так прост, что к нему невозможно придраться и, поскольку весь вывод состоит из таких шагов, мы предполагаем, что он безошибочен. Каждый тип простоты, однако, привносит свой тип сложности. В случае доказательств, это сложность системы, на которую они опираются – а именно, человеческого языка; в случае логических выводов, это их астрономическая длина, делающая их почти невозможными для понимания.

Таким образом, мы считаем логический вывод частью общего метода для составления искусственных структур, подобных доказательствам. Однако он лишен гибкости или всеобщности, поскольку предназначен только для работы с математическими понятиями, которые, в свою очередь, жестко определены.

В качестве средства общения, открытия, фиксации материала никакой формальный язык не способен конкурировать со смесью национального математического арго и формул, привычной для каждого работающего математика.

Однако в силу своей жесткой нормализованности формальные тексты могут сами служить объектом математического исследования. Метатеоремы вызывают значительный интерес (и сильные эмоции), будучи интерпретированы расширительно (как теоремы о математике). Но именно возможность таких и еще более расширительных толкований определяет общефилософское и общегуманитарное значение математической логики.

Интерпретация

Формулы теории имеют смысл только тогда, когда имеется какая-нибудь интерпретация входящих в них символов.

Если ***T*** – теория с языком первого порядка, то можно рассматривать различные интерпретации этой теории. Чтобы определить **интерпретацию,** мы должны задать:

- универсум M, называемый **областью интерпретации**;
- соответствие, относящее:
    - каждому n-местному предикату некоторое n-местное отношение в M,
    - каждой функциональной букве, требующей n аргументов, – некоторую функцию $M^n \to M$ и
    - каждой константе – некоторый элемент из M;
- предметные переменные мыслятся пробегающими область интерпретации M;
- логическим связкам придается их обычный смысл.

Для заданной интерпретации всякая формула P без свободных переменных представляет собой высказывание, которое истинно или ложно. Если высказывание является истинным, то говорят, что формула P **выполняется** в данной интерпретации.

Всякая формула со свободными переменными выражает некоторое отношение на области интерпретации; это отношение может быть выполнено (истинно) для одних значений переменных из области интерпретации и не выполнено (ложно) для других.

Формализация задается не только синтаксисом и семантикой формального языка (эти компоненты как раз чаще всего берутся традиционными, из хорошо известного крайне ограниченного набора), но и множеством утверждений, которые считаются истинными. Именно эта формулировка базисных свойств, аксиом, описывающих некоторую предметную область, обычно рассматривается как математическое описание объектов. Таким образом, практически нас интересуют не все интерпретации данной теории, а лишь те из них, на которых выполнены аксиомы.

Модели

- Интерпретация называется **моделью множества формул** Г, если все формулы этого множества выполняются в данной интерпретации.
- **Моделью теории** называется такая интерпретация, в которой истинны все теоремы теории.

Общезначимость и непротиворечивость

- Формула P называется **общезначимой**, если она истинна в каждой интерпретации (обозначается $\models P$).
- Формула P называется **противоречием**, если формула P ложна во всякой интерпретации.
- Формула P называется **логическим следствием** множества формул Г, если P выполняется в любой модели Г.
- Формула B является **логическим следствием** формулы A (обозначение: $A \Rightarrow B$), если формула B выполнена в любой интерпретации, в которой выполнена формула A.
- Формулы A и B **логически эквивалентны** (обозначение: $A=B$), если они являются логическим следствием друг друга.
- Формальная теория ***T*** называется **семантически непротиворечивой**, если ни одна ее теорема не является противоречием.

Формальная теория пригодна для описания тех предметных областей, которые являются ею моделями.

Справедлива **метатеорема**:

Модель для формальной теории ***T*** существует тогда и только тогда, когда ***T*** семантически непротиворечива.

Обсудим подход к разрешению противоречий, принятый в математике [27]. Формальные системы, использующие логический вывод, не могут содержать противоречий – противоречия мгновенно заражают всю систему, подобно глобальному раку. Это не похоже на человеческую мысль. Если вы обнаружите в своих рассуждениях противоречие, вряд ли это разрушит все здание вашего мышления. Вместо этого вы, скорее всего, попытаетесь найти те идеи или методы рассуждения, которые явились причиной противоречия. Иными словами, вы попытаетесь, насколько это возможно, выйти из ваших внутренних систем, приведших к противоречию, и попробуете их исправить. Маловероятно, что вы поднимите руки вверх и воскликнете: «Это показывает, что теперь я верю во все, что угодно!» – разве что в шутку.

В действительности, противоречия – это важный источник прогресса во всех областях жизни, и математика не является исключением. В прошлом, когда в математике обнаруживалось противоречие, математики тут же пытались найти виновную в этом систему, выйти из таковой, проанализировать её и, если возможно, залатать прореху.

Вместо того чтобы делать математику слабее, нахождение и «починка» противоречивых систем только усиливали её.

Полнота, независимость и разрешимость

Пусть универсум M, рассматриваемый с соответствующей интерпретацией, является моделью формальной теории ***T***.

- Формальная теория ***T*** называется ***полной*** (относительно данной интерпретации), если каждому истинному высказыванию об объектах M соответствует теорема теории ***T*** .
- Если для предметной области M существует формальная полная непротиворечивая теория ***T***, то M называется **аксиоматизируемой** (или **формализуемой**).
- Система аксиом (или аксиоматизация) формально непротиворечивой теории ***T*** называется **независимой**, если никакая из аксиом не выводима из остальных по правилам вывода теории ***T***.

Одним из первых вопросов, которые возникают при задании формальной теории, является вопрос о том, возможно ли, рассматривая какую-нибудь формулу формальной теории, определить, является ли она доказуемой или нет. Другими словами, речь идет о том, чтобы определить, является ли данная формула теоремой или *не-теоремой* и как это доказать.

- Формальная теория ***T*** называется **разрешимой**, если существует алгоритм, который для любой формулы теории определяет, является ли это формула теоремой теории.
- Формальная теория ***T*** называется **полуразрешимой**, если существует алгоритм, который для любой формулы P теории выдает ответ «да», если P является теоремой теории, и выдает «нет» или, может быть, не выдает никакого ответа, если P не является теоремой (то есть алгоритм применим не ко всем формулам).

3.2 Примеры формальных теорий

Приведем несколько примеров формальных теорий из книги Хофштадтера [27].

Формальная система *MIU*

Даглас Хофштадтер

Алфавит: *M, I, U*.
Формулы = {*M, I, U*}*.
Аксиома: *MI*.
Правила вывода:

1) $xI \to xIU$ (продукция);
2) $Mx \to Mxx$ (продукция);
3) $III \to U$ (правило переписывания);
4) $UU \to \varnothing$ (правило переписывания, $\varnothing$ обозначает пустую последовательность).

- Продукция – правило, применяемое к формулам, рассматриваемым как единое целое.
- Правило переписывания – правило, которое может применяться к любой подформуле формулы.

Приведем типичный вывод в этой теории:

MI	Аксиома
MII	Правило 2
MIIII	Правило 2
MUI	Правило 3
MUIU	Правило 1

MUIUUIU Правило 2
MUIIU Правило 4

Любые утверждения о свойствах этой теории являются метатеоремами.
Предлагается читателю задача: «Найдите вывод *MU* или докажите, что он невозможен».

Формальная система *PR*

Алфавит: {*P, R*, –}.
Выражения – элементы {*P, R*, –}*.
Формулы – строки (последовательности) вида *xPyRz*, где *x*, *y* и *z* – строчки, состоящие только из тире.
Схема аксиом:
xP–Rx– является аксиомой, когда x состоит только из тире (каждое из двух вхождений *x* замещает одинаковое число тире).
Правило вывода (схема):
Пусть *x, y* и *z* – строчки, состоящие только из тире. Пусть *xPyRz* является теоремой. Тогда *xPy–Rz–* также будет теоремой.

В системе *PR* используются только удлиняющие правила, т. е. количество символов в формуле в результате применения правила вывода увеличивается.

Задачи. 1. Докажите: Формальная система, имеющая только удлиняющие правила (и не имеющая укорачивающих правил), имеет разрешающий алгоритм.
2. Найдите разрешающий алгоритм для теорем *PR*.

Прежде, чем читать дальше, попробуйте решить данные задачи.

Разрешающий алгоритм для формальных систем, имеющих только удлиняющие правила, можно сформулировать в следующем виде.

Пусть F – проверяемая формула и M – множество всех аксиом теории.

While $\forall x \in M((\text{длина}(x) \le \text{длина}(F)) \ \&\ F \notin M$ **do**
Begin
Для каждой формулы $x \in M$ выполняем следующее:
Begin
Пусть R_x обозначает множество всех формул, получаемых из x однократным применением правил вывода.
Полагаем M равным $(M \cup R_x) \setminus \{x\}$.
End
End
If $F \in M$ **then** F – теорема **else** F – не теорема.

Выберем следующую интерпретацию системы *PR* (одна из возможных).

- Универсум – множество целых положительных чисел.
- Последовательность, состоящая из *n* тире, интерпретируется как число *n*.
- *P* интерпретируется как символ +.
- *R* интерпретируется как символ =.

Нетрудно убедиться, что указанная интерпретация теореме *xPyRz* ставит в соответствие истинное утверждение о целых положительных числах «$x+y=z$» и поэтому данная интерпретация является моделью системы *PR*.

Теперь мы можем использовать более простой разрешающий алгоритм для теории *PR*: формула *xPyRz* является теоремой тогда и только тогда, когда $x+y=z$ – истина.

В модели теоремы и истины совпадают – то есть, между теоремами и фрагментами реального мира существует изоморфизм.

Грубо говоря: «изоморфизм» есть преобразование, сохраняющее информацию. Слово «изоморфизм» приложимо к тем случаям, когда две сложные структуры могут быть отображены одна в другую таким образом, что каждой части одной структуры соответствует какая-то часть другой структуры («соответствие» здесь означает, что эти части выполняют в своих структурах сходные функции).

При данной интерпретации есть изоморфизм между системой *PR* и сложением.

Формальная система *UR*

Алфавит: {*U, R*, –}.
Выражения – элементы {*U*, *R*, –}*.
Формулы – строки вида *xUyRz*, где *x*, *y* и *z* – строчки, состоящие только из тире.
Схема аксиом:
xU–Rx является аксиомой, когда *x* состоит только из тире (каждое из двух вхождений x замещает одинаковое число тире).
Правило вывода (схема):
Пусть x, y и z – строчки, состоящие только из тире. Пусть *xUyRz* является теоремой. Тогда *xUy–Rzx* также будет теоремой.

Задача. Найдите разрешающий алгоритм для теорем *UR*.

Выберем следующую интерпретацию системы *UR*.

- Универсум – множество целых положительных чисел.
- Последовательность, состоящая из n тире, интерпретируется как число n.
- P интерпретируется как символ ×.
- R интерпретируется как символ =.

Нетрудно убедиться, что указанная интерпретация теореме *xUyRz* ставит в соответствие истинное утверждение о целых положительных числах «$x \times y = z$» и поэтому данная интерпретация является моделью системы *UR*.

Формальная система *PR*1

Алфавит: {*P, R*, –}.
Выражения – элементы {*P, R*, –}*.
Формулы – строки вида *xPyRz*, где *x, y* и *z* – строчки, состоящие только из тире.
Схемы аксиом:

1. *xP–Rx–* является аксиомой, когда *x* состоит только из тире (каждое из двух вхождений *x* замещает одинаковое число тире).
2. *xP–Rx* является аксиомой, когда *x* состоит только из тире (каждое из двух вхождений *x* замещает одинаковое число тире).

Правило вывода (схема):
Пусть *x, y* и *z* – строчки, состоящие только из тире. Пусть *xPyRz* является теоремой. Тогда *xPy–Rz–* также будет теоремой.

Задача. Найдите разрешающий алгоритм для теорем *PR1*.

Рассмотрим различные интерпретации системы *PR1*.

1. Выберем интерпретацию системы *PR1* такую же, как для *PR*.

- Универсум – множество целых положительных чисел.
- Последовательность, состоящая из *n* тире, интерпретируется как число *n*.
- *P* интерпретируется как символ +.
- *R* интерпретируется как символ =.

Указанная интерпретация теореме *xPyRz* ставит в соответствие утверждение о целых положительных числах «$x+y=z$». Но эти утверждения могут быть и ложными, поэтому данная интерпретация не является моделью системы *PR*1.

2. Вторая интерпретация системы *PR*1 отличается от первой только тем, как интерпретируется символ *R*.

- *R* интерпретируется как «равняется или больше на 1».

Указанная интерпретация теореме *xPyRz* ставит в соответствие истинное утверждение о целых положительных числах «$x+y=z+1 \vee x+y=z$» и поэтому данная интерпретация является моделью системы *PR*1. Более того, любое истинное утверждение «$x+y=z+1 \vee x+y=z$» описывается в виде теоремы *xPyRz* теории *PR*1.

3. В последней интерпретации символ R понимается снова по-другому.

- R интерпретируется как символ «≥».

Указанная интерпретация теореме $xPyRz$ ставит в соответствие истинное утверждение о целых положительных числах «$x+y\geq z$», и поэтому данная интерпретация является моделью системы $PR1$. Но мы сейчас имеем существенное отличие от предыдущей интерпретации: не все истинные утверждения вида $x+y\geq z$ являются теоремами в $PR1$. Так, например, формула —P–R– имеет истинную интерпретацию 2+1 ≥ 1, но это не теорема.

3.3 Исчисление высказываний

Классическое определение исчисления высказываний

Исчислением высказываний называется формальная теория с языком логики высказываний, со схемами аксиом

$A1)$ $A\supset(B\supset A)$;

$A2)$ $(A\supset(B\supset C))\supset((A\supset B)\supset(A\supset C))$;

$A3)$ $(\neg B\supset\neg A)\supset((\neg B\supset A)\supset B)$

и правилом вывода MP (*Modus Ponens* – обычно переводится как **правило отделения**):

$$\frac{A,\ A\supset B}{B}\ MP$$

Здесь A, B и C – *любые* формулы. Таким образом, множество аксиом исчисления высказываний бесконечно, хотя задано тремя схемами аксиом. Множество правил вывода также бесконечно, хотя оно задано только одной схемой.

Пример. Для любой формулы A построим вывод формулы $A\supset A$, т. е.

$A\supset A$ – теорема.

Подставляем в схему аксиом $A2$ вместо B формулу $A\supset A$ и вместо C формулу A, получаем аксиому

$(A\supset((A\supset A)\supset A))\supset((A\supset(A\supset A))\supset(A\supset A))$. (1)

Подставляем в $A1$ вместо формулы B формулу $A\supset A$, получаем аксиому

$A\supset((A\supset A)\supset A)$. (2)

Из формул (1) и (2) по правилу MP получаем

$(A\supset(A\supset A))\supset(A\supset A)$. (3)

Подставляем в $A1$ вместо формулы B формулу A, получаем аксиому

$A\supset(A\supset A)$. (4)

Из формул (3) и (4) по правилу MP получаем $A\supset A$.

Теорема 10. Если $\Gamma|\!-A\supset B$ и $\Gamma|\!-A$, то $\Gamma|\!-B$.

Доказательство. Пусть $A_1, A_2,..., A_n$ – вывод формулы A из Γ, где A_n совпадает с A. Пусть $B_1, B_2,..., B_m$ – вывод формулы $A\supset B$ из Γ, где B_n совпадает с $A\supset B$. Тогда $A_1, A_2,..., A_n, B_1, B_2,..., B_m, B$ – вывод формулы B из Γ. Последняя формула в этом выводе получена применением правила MP к формулам A_n и B_m.

В исчислении высказываний импликация очень тесно связана с выводимостью.

Теорема 11 (о дедукции). Если $\Gamma, A|\!-B$, то $\Gamma|\!-A\supset B$ и наоборот.

Следствие.

1. $A\supset B, B\supset C\ |\!-A\supset C$.
2. $A\supset(B\supset C), B\ |\!-A\supset C$.

Полнота, разрешимость и непротиворечивость исчисления высказываний

В исчислении высказываний у нас есть два понятия, касающиеся формул: теорема и тавтология. Аксиомы и правило вывода придуманы так, что эти два понятия совпадают.

Наша цель показать, что формула исчисления высказываний является тавтологией тогда и только тогда, когда она есть теорема. В одну сторону это доказывается совсем просто.

Теорема 12.

1. Любая аксиома в исчислении высказываний является тавтологией.
2. Любая теорема в исчислении высказываний является тавтологией.

Доказательство. То, что каждая аксиома $A1$–$A3$ является тавтологией, легко проверить с помощью таблиц истинности. Для доказательства пункта 2 теоремы, предварительно докажем, что правило *MP*, примененное к тавтологиям, приводит к тавтологиям.

Действительно, пусть при произвольном распределении истинностных значений формулы A и $A \supset B$ являются тавтологиями. Тогда формула A истинна и, по свойствам импликации, B истинно. Следовательно, B – тавтология.

Теорема 13. (Пост, 1921) Если формула A логики высказываний является тавтологией, то A – теорема в исчислении высказываний.

Интерпретация формул исчисления высказываний проста – область интерпретации состоит из двух значений «истина» и «ложь»; поэтому пропозициональная переменная принимает только значения И и Л и интерпретация составной формулы вычисляется по известным законам с помощью логических операций над истинностными значениями. Поскольку любая формула содержит только конечное число пропозициональных переменных, то формула обладает только конечным числом различных интерпретаций. Следовательно, исчисление высказываний является, очевидно, разрешимой формальной теорией.

Легко убедиться, что исчисление высказываний является формально непротиворечивой теорией. Действительно, все теоремы исчисления высказываний суть тавтологии. Отрицание тавтологии не есть тавтология. Следовательно, никакая формула не может быть выведена вместе со своим отрицанием.

3.4 Теории первого порядка

Языки первого порядка

Чтобы описания математических объектов и математические рассуждения сделать точными в математической логике используются искусственные языки. Самый распространенный вид таких языков – так называемые логико-математические языки первого порядка. Каждый **язык первого порядка** задается своей сигнатурой – набором из трех множеств: множество **констант**; множество **функциональных символов** (**функторов**); множество **предикатных символов**. При этом с каждым предикатным или функциональным символом однозначно связано некоторое натуральное число – количество аргументов (местность, арность) этого символа. Функциональный или предикатный символ, местность которого равна k, называется k-местным.

Во всяком языке первого порядка имеется счетный набор **переменных.** Используя переменные и константы как аргументы функциональных символов, строят **термы**. Причем этот процесс рекуррентный, и из существующих термов, сделав их аргументами каких-то функциональных символов можно построить новые термы. Термы играют роль имен объектов.

Элементарные формулы определяются как выражения вида $P(t_1, t_2,\ldots, t_k)$, где P есть k-местный предикатный символ, а $t_1, t_2,\ldots, t_k$ – термы. Используя знаки логических операций $\neg$ (отрицание), $\vee$ (дизъюнкция), $\wedge$ (конъюнкция), $\supset$ (импликация), $\sim$ (эквиваленция), $\forall$ (квантор общности), $\exists$ (квантор существования) из элементарных формул строятся более сложные формулы.

Языки первого порядка используются для записи математических утверждений, причем для каждой конкретной области математики, или, как говорят математической

теории, выбирается язык с подходящей сигнатурой. Использование языка первого порядка для записи утверждений, относящейся к данной математической теории, становится возможным, если все основные понятия теории удается разбить на три категории: «объекты», «функции» и «предикаты». При этом функции и предикаты должны быть определены только на объектах, а значениями функций являются только объекты. В частности, не допускается рассматривать предикаты, заданные на функциях, или функции, заданные на предикатах.[2] Затем для некоторых конкретных, замечательных в том или ином отношении объектов, функций и предикатов, фиксируются их обозначения, которые и образуют сигнатуру языка.

В математической логике разработаны языки для многих математических теорий. Два из них играют наиболее важную роль в математике. Это язык формальной (элементарной) арифметики (о нем речь пойдет в разделе 5) и язык теории множеств.

Сигнатура языка теории множеств (точнее, **языка теории множеств Цермело–Френкеля**) содержит только два двуместных предиката: $\in$ (отношение принадлежности) и = (отношение равенства) и константу $\varnothing$ (пустое множество). Вместо $\in(t_1, t_2)$ и $=(t_1, t_2)$ пишут $t_1 \in t_2$ и $t_1 = t_2$, соответственно. Термами этого языка являются только переменные. Вот примеры формул языка теории множеств:

1. $\forall z\ (z \in x \supset z \in y)$. Имея в виду естественный подразумеваемый смысл символов языка, для этой формулы можно ввести сокращенное обозначение $x \subseteq y$.
2. $\neg\ (x = y)$. Эта формула имеет сокращенное обозначение $x \neq y$.
3. $\neg\exists y\ (y \in x)$. Эта формула обычно записывается как $x = \varnothing$.

Синтаксические свойства истинности теорий с языком первого порядка

Пусть нам даны некоторая формальная теория ***T*** с языком первого порядка и задана интерпретация этой теории. Через ***F*** множество всех формул теории ***T***, истинных в данной интерпретации. Перечислим те свойства ***F***, которые отражают заложенные в языки первого порядка логику, независящую от конкретных особенностей интерпретации.

- Для любой замкнутой формулы P, либо $P \in \boldsymbol{F}$, либо $\neg P \in \boldsymbol{F}$.
- Множество ***F*** не содержит противоречия, т. е. ни для какой формулы P не может быть, чтобы одновременно выполнялось $P \in \boldsymbol{F}$ и $\neg P \in \boldsymbol{F}$.
- Множество ***F*** содержит все тавтологии.
- Множество ***F*** содержит следующие общезначимые формулы (называемые «логические аксиомы с кванторами»):
  $\forall x\ A(x) \supset A(t)$,
  где $A(t)$ есть формула теории ***T*** и t есть терм теории ***T***, свободный для x в $A(x)$. Заметим, что t может совпадать с x, и тогда мы получаем аксиому $\forall x A(x) \supset A(x)$.
  $\forall x\ (A \supset B(x)) \supset (A \supset \forall x\ B(x))$,
  где A не содержит свободных вхождений переменной x.
- Множество ***F*** замкнуто относительно правил вывода *modus ponens* и обобщения. По определению это означает, что если $A \in \boldsymbol{F}$ и $A \supset B\ \in \boldsymbol{F}$, то также $B \in \boldsymbol{F}$; если $A \in \boldsymbol{F}$, то $\forall x A \in \boldsymbol{F}$ для любой переменной x.

Определение теории первого порядка

Нам понадобится следующее понятие. Пусть дана формула P, свободное вхождение переменной x в P и терм t. Мы говорим, что ***данное вхождение x не***

---

[2] Из-за этих ограничений язык называется *языком логики предикатов первого порядка* (или просто *языком первого порядка*).

связывает t в P, если оно не лежит в области действия ни одного квантора вида $\forall y$ и $\exists y$, где y - переменная, входящая в t.

Иными словами, после подстановки t вместо данного вхождения x все переменные, входящие в t, останутся свободными в P.

Чаще всего приходится подставлять терм вместо каждого из свободных вхождений данной переменной. Важно, что такая операция переводит термы в термы и формулы в формулы. Если каждое свободное вхождение x в P не связывает t, мы будем говорить просто, что ***терм t свободный для x в P.***

Пример

1. Терм y свободен для переменной x в формуле $P(x)$, но тот же терм y не свободен для переменной x в формуле $\forall y\, P(x)$.
2. Терм $f(x,z)$ свободен для переменной x в формуле $\forall y P(x, y) \supset Q(x)$, но тот же терм $f(x,z)$ не свободен для переменной x в формуле $\exists z \forall y\, P(x, y) \supset Q(x)$.

Аксиоматическая формальная теория с языком первого порядка называется **теорией первого порядка**, если она имеет следущие аксиомы и правила вывода.

Аксиомы теории первого порядка $\boldsymbol{T}$ разбиваются на два класса: логические аксиомы и собственные (или нелогические) аксиомы.

Логические аксиомы: каковы бы ни были формулы A, B и C теории $\boldsymbol{T}$, следующие формулы являются логическими аксиомами теории $\boldsymbol{T}$.

$\boldsymbol{A_1}$. $A \supset (B \supset A)$;

$\boldsymbol{A_2}$. $(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))$;

$\boldsymbol{A_3}$. $(\neg B \supset \neg A) \supset ((\neg B \supset A) \supset B)$;

$\boldsymbol{A_4}$. $\forall x\, A(x) \supset A(t)$, где $A(t)$ есть формула теории $\boldsymbol{T}$ и t есть терм теории $\boldsymbol{T}$, свободный для x в $A(x)$. Заметим, что t может совпадать с x, и тогда мы получаем аксиому $\forall x A(x) \supset A(x)$.

$\boldsymbol{A_5}$. $\forall x\, (A \supset B(x)) \supset (A \supset \forall x\, B(x))$, где A не содержит свободных вхождений переменной x.

Собственные аксиомы: таковые не могут быть сформулированы в общем случае, ибо меняются от теории к теории. Правилами вывода во всякой теории первого порядка являются

1. *Modus ponens*: $\dfrac{A,\ A \supset B}{B}$ MP
2. Правило обобщения: $\dfrac{A(x)}{\forall x A(x)}$ Gen

Теория первого порядка, которая не содержит собственных аксиом, называется **исчислением предикатов первого порядка**. **Чистым исчислением предикатов** называется исчисление предикатов первого порядка, не содержащее предметных констант и функторов.

Как мы увидим позже, логические аксиомы выбраны таким образом, что множество логических следствий аксиом теории в точности совпадает с множеством теорем теории. В частности, для исчисления предикатов первого порядка множество его теорем совпадает с множеством логически общезначимых формул.

Общезначимые формулы в теории первого порядка играют ту же роль, что тавтологии в логике высказываний. Между ними есть и формальная связь: если взять любую тавтологию и вместо входящих в нее пропозициональных переменных подставить произвольные формулы языка первого порядка, то получится общезначимая формула. В самом деле, пусть есть некоторая интерпретация теории, при которой как-то зафиксированы значения предметных переменных. Тогда каждая из подставленных формул станет истинной или ложной, а значение всей формулы определяется с помощью таблиц истинности для логических связок, то есть по тем же правилам, что и в логике высказываний.

Конечно, бывают и другие общезначимые формулы, не являющиеся частным случаем пропозициональных тавтологий. Например, формула

$$\forall xA(x)\supset\exists yA(y)$$

общезначима (здесь существенно, что универсум любой интерпретации не пуст).

Некоторые свойства теорий первого порядка

Для теорий первого порядка можно уточнить некоторые понятия, введенные в 4.1.

Формула P является **противоречием,** тогда и только тогда, когда формула $\neg P$ является общезначимой.

Формальная теория $\boldsymbol{T}$ называется (**формально**) **непротиворечивой**, если в ней не являются выводимыми одновременно формулы P и $\neg P$.

Имеет место **метатеорема**:

Теория $\boldsymbol{T}$ формально непротиворечива тогда и только тогда, когда она семантически непротиворечива.

Из определений логического следования и эквивалентности непосредственно вытекают следующие утверждения:

$A \Rightarrow B$ тогда и только тогда, когда $\models A \supset B$;

$A = B$ тогда и только тогда, когда $\models A \sim B$.

Имеют место также следующие логические следования и эквивалентности (формулы A и B – произвольны, формула C не содержит никаких вхождений переменной x):

$\neg\forall xA(x) = \exists x\neg A(x)$, $\neg\exists xA(x) = \forall x\neg A(x)$;

$\forall x(A(x)\&B(x)) = \forall x\, A(x)\&\forall xB(x)$,

$\exists x(A(x)\vee B(x)) = \exists xA(x)\vee\exists xB(x)$;

$\exists x(A(x)\&B(x)) \Rightarrow \exists xA(x)\&\exists xB(x)$,

$\forall xA(x)\vee\forall xB(x) \Rightarrow \forall x(A(x)\vee B(x))$;

$\forall x\forall yA(x,y) = \forall y\forall xA(x,y)$, $\exists x\exists yA(x,y) = \exists y\exists xA(x,y)$;

$\forall x(A(x)\&C) = \forall xA(x)\&C$, $\forall x(A(x)\vee C) = \forall xA(x)\vee C$;

$\exists x(A(x)\&C) = \exists xA(x)\&C$, $\exists x(A(x)\vee C) = \exists xA(x)\vee C$;

$C\supset\forall xA(x) = \forall x(C\supset A(x))$, $C\supset\exists xA(x) = \exists x(C\supset A(x))$;

$\forall xA(x)\supset C \Rightarrow \exists x(A(x)\supset C)$.

Имеет место следующее утверждение:

Теорема 14. Если теория первого порядка (формально) противоречива, то в ней выводима любая формула.

Доказательство. В самом деле, пусть формулы A и $\neg A$ выводимы в теории. Формула $\neg A\supset(A\supset B)$ является тавтологией в исчислении высказываний, следовательно, она выводима. Её вывод, поскольку он содержит только *MP*, остается выводом и в любой теории первого порядка. Поэтому формула $\neg A\supset(A\supset B)$ выводима в теории первого порядка. Дважды применяя *MP*, мы получаем вывод произвольной формулы B.

Таким образом, для доказательства непротиворечивости какой-либо теории первого порядка достаточно установить невыводимость в этой теории хотя бы одной формулы.

Теорема 15. (Гёдель, 1930). Всякая общезначимая формула теории $\boldsymbol{T}$ первого порядка является теоремой теории $\boldsymbol{T}$.

Значительная часть теорем логики состоит в доказательстве утверждений вида «$\Gamma\vdash P$» или «не $\Gamma\vdash P$» для разных теорий первого порядка, множеств Γ и разных (классов) формул P.

Результат «$\Gamma\vdash P$» может доказываться посредством предъявления описания вывода формулы P из Γ. Однако в мало-мальски сложных случаях оно оказывается настолько длинным, что заменяется инструкцией по составлению такого описания, более или менее полной. Наконец, доказательство «$\Gamma\vdash P$» может вообще не сопровождаться предъявлением

вывода *P* из Г, хотя бы и неполного. В этом случае мы «не доказываем *P*, а доказываем, что существует доказательство *P*» [16, с. 45].

Результат «не Г|–*P*» в редких случаях может устанавливаться чисто синтаксическим рассуждением, но обычно доказательство опирается на конструкцию модели, т.е. интерпретации, в которой Г истинно, а *P* ложно.

Непротиворечивость, полнота и неразрешимость исчисления предикатов

Теорема 16. Исчисление предикатов первого порядка непротиворечиво.

Теорема 17. (о полноте исчисления предикатов). В исчислении предикатов первого порядка теоремами являются те и только те формулы, которые общезначимы.

Доказательство. В силу теоремы 15 достаточно доказать только, что любая теорема исчисления предикатов первого порядка является общезначимой формулой. Но это следует из общезначимости аксиом исчисления предикатов и синтаксических свойств истинности в теориях первого порядка.

Теорема 18. (Теорема Чёрча о неразрешимости исчисления предикатов, 1936). Не существует алгоритма, который для любой формулы исчисления предикатов первого порядка устанавливает, общезначима она или нет.

Теорема Чёрча остается в силе и для любой теории первого порядка, содержащей все формулы исчисления предикатов первого порядка.

Теорема 17 показывает, что в логике предикатов синтаксический метод теорий первого порядка равносилен семантическому методу, использующему понятие интерпретации, модели, общезначимости и т. п. Для исчисления высказываний имеет место аналогичная эквивалентность семантических (тавтология и др.) и синтаксических (теорема и др.) понятий. Отметим также, что для исчисления высказываний теорема о полноте системы приводит к решению проблемы разрешения. Однако для теорий первого порядка мы не можем получить разрешающий алгоритм для общезначимости или, то же самое, для выводимости в любом исчислении предикатов первого порядка.

У логики один недостаток: она не
останавливается на полпути.
Д. Уиндем День триффидов.

4 Элементарная арифметика и неполнота

4.1 Элементарная арифметика

Теория элементарной арифметики

Попытаемся получить независимую формализацию средств математического рассуждения, использующих только понятие натурального числа (не упоминая ни действительных чисел, ни – тем более – произвольных множеств Кантора). Это самая надежная часть арсенала математики, не скомпрометировавшая себя парадоксами. Естественно назвать ее **элементарной арифметикой** (по аналогии с термином «элементарная теория чисел» – в отличие от аналитической теории чисел).

Язык элементарной арифметики – это язык первого порядка, который имеет:
предметную константу **0;**
двухместные функторы + и ×, одноместный функтор S;.
двухместный предикат =.

В качестве универсума при **стандартной интерпретации** элементарной арифметики рассматривается множество натуральных чисел, причем предметной константе **0** соответствует число ноль. Термы интерпретируются как натуральные числа. Значение функции $S(t)$ для любого терма t, обозначающего натуральное число, интерпретируется как (непосредственно) следующее число за t. Термы, имеющие вид $S(\mathbf{0})$, $S(S(\mathbf{0}))$, $S(S(S(\mathbf{0})))$ и т. д. интерпретируются как натуральные числа 1, 2, 3 и т.д. Знаки + и × интерпретируются как обозначения операций сложения и умножения.

Язык элементарной арифметики использует только предикатный символ равенства «=» (понимаемый как равенство натуральных чисел). Если t_1, t_2 – термы языка, то $t_1=t_2$ – элементарная формула. Из элементарных формул с помощью логических связок и кванторов строятся более сложные формулы языка, причем $\exists x$ мы понимаем как «существует натуральное число», а $\forall x$ – как «для всех натуральных чисел».

Средствами языка элементарной арифметики легко записываются простейшие утверждения о свойствах натуральных чисел, например:
«$x < y$» $\Leftrightarrow \exists z(\neg(z=\mathbf{0})\& \ y=x+z)$;
«x – четное число» $\Leftrightarrow \exists y(x=y+y)$;
«x – простое число» $\Leftrightarrow$ «$S(\mathbf{0})<x$» $\& \ \neg\exists y\exists z($«$y<x$»$\&$ «$z<x$» $\&\ x=y\times z)$;
«существует бесконечно много простых чисел» $\Leftrightarrow \forall x\exists y($«$x<y$» $\&$ «y – простое число»).

Все, что мы до сих пор говорили о нашем «понимании» языка элементарной арифметики, является нашим личным делом и к определению теории элементарной арифметики не относится. Аксиоматическое построение этой дисциплины известно как *теория элементарной арифметик*и (или *система аксиом Пеано*, названная в честь автора Джузеппе Пеано, 1891 г.).

Теория элементарной арифметики *EA* – это теория первого порядка с языком элементарной арифметики и аксиомами Пеано. В этой теории о свойствах введенных нами символов известно только то, что сформулировано в аксиомах. Часть этих свойств уже содержится в логических аксиомах и правилах вывода, которые безоговорочно

принимаются в *EA*. Для определения наиболее существенных свойств мы вводим собственные аксиомы *EA*:

Собственные аксиомы (схемы аксиом) элементарной арифметики:
P_1: $(P(\mathbf{0})\ \&\ \forall x(P(x)\supset P(S(x)))\supset\forall z P(z)$ (принцип математической индукции, P – произвольная формула),
P_2: $S(t_1) = S(t_2) \supset t_1 = t_2$,
P_3: $\neg(\ S(t)=\mathbf{0})$,
P_4: $t_1 = t_2 \supset (t_1 = t_3 \supset t_2 = t_3)$,
P_5: $t_1 = t_2 \supset S(t_1) = S(t_2)$,
P_6: $t+\mathbf{0} = t$,
P_7: $t_1 + S(t_2) = S(t_1+ t_2)$,
P_8: $\mathbf{0}\times t = \mathbf{0}$,
P_9: $S(t_1) \times t_2 = t_1\times t_2 + t_2$.

Аксиомы P_4 и P_5 обеспечивают некоторые необходимые свойства равенства. Аксиомы P_3 и P_2 обеспечивают существование нуля и операции «непосредственно следующий». Аксиомы P_6–P_9 представляют собой рекурсивные равенства, служащие определениями операций сложения и умножения.

С помощью *MP* из схемы аксиом P_1 мы можем получить следующее правило индукции: из $P(\mathbf{0})$ и $\forall x(P(x) \supset P(\mathrm{S}(x)))$ выводится $\forall x P(x)$.

Терм $S(\ldots S(\mathbf{0})\ldots)$, где символ S повторяется k раз, кратко будем обозначать $\boldsymbol{k}$. Таким образом, натуральное число k в *EA* интерпретируется как терм $\boldsymbol{k}$. Термы такого рода **0**, **1**, **2**, … принято называть нумералами – стандартными обозначениями конкретных натуральных чисел. Очевидно, термы *EA* – это обозначения полиномов (от нескольких, вообще говоря, переменных) с натуральными коэффициентами. Например, терм $((x\times x)+((\mathbf{2}\times x)\times y))+(y\times y)$ представляет полином $x^2+2xy+y^2$. Мы будем рассматривать теорию *EA* только в стандартной интерпретации.

Сейчас удобно на примере системы *EA* пояснить, почему необходима аксиома $\boldsymbol{A_4}$ для определения теории первого порядка – она предотвращает коллизию переменных.

Для теории первого порядка должна выполняться аксиома

$\boldsymbol{A_4}$: $\forall x\ A(x) \supset A(t)$, где $A(t)$ есть формула теории $\boldsymbol{T}$ и t есть терм теории $\boldsymbol{T}$, свободный для x в $A(x)$.

Рассмотрим теорию *EA*. Пусть $A(x) \equiv\exists b(b=x+1)$. Тогда $\forall x A(x) \equiv\forall x\exists b(b=x+1)$ – истинная формула при стандартной интерпретации *EA*. Возьмем терм $t \equiv b$, тогда для терма t имеем следующую формулу $\forall x\exists b(b=x+1) \supset \exists b(b=b+1)$. Формула $\exists b(b=b+1)$ ложна при стандартной интерпретации *EA*, следовательно, формула $\forall x\ A(x) \supset A(t)$ – не общезначима.

Вывод в теории *EA* организован таким образом, что все теоремы элементарной арифметики являются истинными утверждениями математики[17].

Основным средством вывода теорем в теории *EA* является, как и следовало ожидать, схема индукции. Рассмотрим в качестве примера вывод формулы $\mathbf{0}+x=x$ (она отличается от аксиомы $x+\mathbf{0}=x$!). Обозначим $\mathbf{0}+x=x$ через $A(x)$. Сначала мы должны доказать $A(\mathbf{0})$, т.е. $\mathbf{0}+\mathbf{0}=\mathbf{0}$, но это частный случай упомянутой только что аксиомы. Теперь можем доказать $A(x)\supset A(S(x))$. Предполагая воспользоваться теоремой дедукции, возьмем $A(x)$ в качестве гипотезы:
$A(x)$ или $\mathbf{0}+x=x$ (гипотеза),
$\mathbf{0}+S(x)=S(\mathbf{0}+x)$ (частный случай аксиомы),
$\mathbf{0}+x=x \supset S(\mathbf{0}+x)=S(x)$ (свойство равенства),
$S(\mathbf{0}+x)=S(x)$ (modus ponens),
$\mathbf{0}+S(x)=S(x)$ или $A(S(x))$ (транзитивность равенства).

По теореме дедукции отсюда следует $\vdash A(x)\supset A(S(x))$, а затем $\vdash \forall x(A(x)\supset A(S(x)))$. Так как $A(\mathbf{0})$ уже доказано, то по схеме индукции получаем $\vdash \forall x A(x)$ или $\vdash \mathbf{0}+x=x$.

Аналогично доказываются другие простые теоремы *EA*. Следует помнить, однако, что перед тем как доказывать какую-либо теорему (например, коммутативность умножения: $x\times y=y\times x$), полезно уже знать некоторые теоремы. Таким образом, даже доказательство простых теорем *EA* содержит в себе творческий момент – он состоит в наиболее рациональном выборе порядка, в котором эти теоремы следует доказывать.

Следующее утверждение является эмпирически установленным фактом: все рассуждения обычной (интуитивной) теории чисел, которые не апеллируют к произвольным действительным числам и функциям, могут быть формально воспроизведены в *EA*.

Язык теории *EA* (как и любой язык теории первого порядка) можно расширить, введя новые функциональные символы и константы, которые «доказуемо выразимы» в языке. Это просто формальный вариант «введения новых обозначений». Их добавление к алфавиту сокращает формульные выводы и записи формул, но не увеличивают множество выводимых формул (см. [16, стр. 75–76]).

Арифметические функции и отношения

Как можно трактовать в теории *EA* операцию возведения в степень, если соответствующего символа в языке теории нет? Другой вопрос: мы пишем формулы, заменяющие в *EA* интуитивно понимаемые предикаты вроде «x – простое число»:
«$\mathbf{1}<x$» & $\neg\exists y\exists z$(«$y<x$»& «$z<x$» & $x=y\times z$);

Какие требования мы должны предъявлять к такого рода формулам? Если кто-то предлагает формулу и утверждает, что она «выражает» то-то и то-то, как это проверить? Для таких ситуаций вводятся понятия выразимости предикатов и представимости функций.

Арифметическими функциями называются функции, у которых область определения и множество значений состоят из натуральных чисел, а арифметическим отношением является всякое отношение, заданное на множестве натуральных чисел. Так, например, умножение есть арифметическая функция с двумя аргументами, а выражение $x+y < z$ определяет некоторое арифметическое отношение с тремя аргументами. Арифметические функции и отношения являются понятиями интуитивными, и не связанными ни с какой формальной системой.

Выразимые отношения

Арифметическое отношение $R(x_1,\ldots, x_n)$ называется *выразимым* в *EA*, если существует формула $A(x_1,\ldots, x_n)$ теории *EA* с n свободными переменными такая, что для любых натуральных чисел $k_1,\ldots, k_n$
1) если $R(k_1,\ldots, k_n)$ истинно, то $\vdash A(\boldsymbol{k_1},\ldots, \boldsymbol{k_n})$,
2) если $R(k_1,\ldots, k_n)$ ложно, то $\vdash \neg A(\boldsymbol{k_1},\ldots, \boldsymbol{k_n})$.

Легко проверить, что формула $x=y$ выражает в *EA* обычное интуитивно понимаемое равенство натуральных чисел. В самом деле, для любых m, n: а) если $m=n$, то $\vdash \boldsymbol{m}=\boldsymbol{n}$ (термы $\boldsymbol{m}$, $\boldsymbol{n}$ в этом случае просто совпадают), б) если $m\neq n$, то $\vdash \neg(\boldsymbol{m}=\boldsymbol{n})$ (проверьте).

Интересно было бы получить какую-то характеристику класса тех предикатов, которые выразимы в *EA*. Если теория *EA* противоречива, то в ней доказуема любая формула и поэтому – в смысле принятого нами определения – в *EA* можно выразить любой предикат (например, формула $x=y$ выражает любой двухместный предикат). Многочисленные примеры выразимых отношений с доказательствами см. в [17, с.132–135]. Любое ли арифметическое отношение является выразимым в элементарной арифметике? Каждое одноместное отношение (свойство) натуральных чисел определяет некоторое множество натуральных чисел, а именно, состоящее из чисел, обладающих

данным свойством. Нетрудно проверить, что множество различных формул в теории первого порядка является счетным множеством. С другой стороны множество всех подмножеств, составленных из натуральных чисел, имеет мощность больше счетного (теорема Кантора), отсюда следует, что существуют свойства натуральных чисел, невыразимые в *EA*.

Представимые функции

Арифметическая функция $f(x_1,\ldots, x_n)$ называется *представимой* в *EA*, если существует формула $A(x_1,\ldots, x_n, y)$ формальной арифметики со свободными переменными $x_1,\ldots, x_n, y$ такая, что для любых натуральных чисел $k_1,\ldots, k_n, k_{n+1}$ таких, что $f(k_1,\ldots, k_n) = k_{n+1}$

1) $\vdash A(\boldsymbol{k_1},\ldots, \boldsymbol{k_n}, \boldsymbol{k_{n+1}})$,

2) $\vdash \forall y\ (\neg(\boldsymbol{k_{n+1}} = y) \supset \neg A(\boldsymbol{k_1},\ldots, \boldsymbol{k_n}, y))$.

Все функции, которые мы в интуитивном смысле считаем вычислимыми (т. е. значения которых вычисляются с помощью алгоритмов), оказываются представимыми в *EA*. Если теория *EA* противоречива, то в ней представима любая функция.

Теорема 19. [17, с. 158] Всякая функция, представимая в *EA*, является частично-рекурсивной и наоборот.

Оказывается, что класс выразимых отношений в *EA* также оказывается очень широк. Но сначала нам потребуется определение.

Характеристической функцией данного отношения $R(\boldsymbol{x})$ (в данном случае $\boldsymbol{x}$ сокращенная запись для $<x_1, \ldots, x_n>$) называется функция $C_R(\boldsymbol{x})$, задаваемая условиями:

$$C_R(\boldsymbol{x}) = \begin{cases} 1, \text{ если } R(\boldsymbol{x}) \text{ истинно,} \\ 0, \text{ если } R(\boldsymbol{x}) \text{ ложно.} \end{cases}$$

Теорема 20. [17, стр. 134] Отношение $R(\boldsymbol{x})$ выразимо в теории *EA* тогда и только тогда, когда характеристическая функция $C_R(\boldsymbol{x})$ представима в *EA*.

Отношение R называется **разрешимым**, если его характеристическая функция C_R является частично-рекурсивной.

Теорема 21. Всякое отношение, выразимое в *EA*, является разрешимым и наоборот. Справедливость этой теоремы следует из теорем 19 и 20.

4.2 Неполнота элементарной арифметики

Геделева нумерация

Пусть дана некоторая формальная система. Предложение, утверждающее, что некоторая последовательность формул образует (или не образует) вывод некоей формулы, уже не является доказательством в самой этой формальной системе. Это утверждение о системе; такие утверждения обычно называют **метаматематическими**. Необходимо тщательно различать математические и метаматематические рассуждения. Нарушение этого условия приводит к парадоксам. Пример см. в разделе 3.1.

У Курта Гёделя возникла великая идея **арифметизации метаматематики**, т.е. замены утверждений о формальной системе эквивалентными высказываниями о натуральных числах с последующим выражением этих высказываний в формальной системе. Идея арифметизации стала ключом к решению многих важных проблем математической логики. Арифметизацией данной теории первого порядка мы называем всякую функцию g, отображающую инъективно множество всех символов, выражений и конечных последовательностей выражений во множество целых положительных чисел.

При этом требуется: (*i*) чтобы значение функции *g* можно было вычислить с помощью некоторого алгоритма, (*ii*) чтобы существовал алгоритм, позволяющий для каждого *m* определить, является ли *m* значением функции *g*, и в случае, если является, то построить тот объект *x*, для которого $m = g(x)$. Значение функции *g* для символа (выражения, последовательности выражений) называется **гёделевым номером** соответствующего объекта.

Курт Гёдель

Существуют различные способы построения гёделевых номеров для системы *EA*[17, с. 151–152]. Изложим кратко одну из возможных гёделевых нумераций. Рассмотрим алфавит теории *EA*. Выберем 10 основных символов этого алфавита: **0**, *S*, (,), +, ×, ⊃, ¬, ∀ и =. Другие логические связки можно выразить через ⊃, ¬ и ∀. Например, $A \vee B \equiv \neg A \supset B$, а $\exists x A(x) \equiv \neg \forall x \neg A(x)$.

Таким образом, основной символ алфавита получает в качестве гёделева номера число, не превосходящее 10. Например, соответствие может быть задано таблицей:

0	*S*	(	)	+	×	⊃	¬	∀	=
1	2	3	4	5	6	7	8	9	10

Предметным переменным сопоставляются простые числа, большие 10 (скажем, *x* получает номер 11, *y* – номер 13 и т. д.).[3]

Формула получает номер по правилу, которое лучше всего пояснить на примере. Рассмотрим формулу $\forall x\ (x \times S(\mathbf{0}) = x)$, которая утверждает, что число не меняется при умножении на 1. Эта формула содержит 12 символов, причем некоторые из них (скажем, *x*) входят несколько раз. Возьмем первые 12 простых чисел, возведем каждое из них в степень, равную номеру соответствующего символа и перемножим полученные числа. Найденное так число

$$2^9 \times 3^{11} \times 5^3 \times 7^{11} \times 11^6 \times 13^2 \times 17^3 \times 19^1 \times 23^4 \times 29^{10} \times 31^{11} \times 37^4 =$$
$$351246696379113496496255178346205341140836151634379449650656150131286227200 0$$

и будет гёделевым номером этой формулы.

Последовательность формул (которая может составлять доказательство) $F_1, F_2, \ldots, F_m$ получает гёделевый номер

$$2^{G_1} \times 3^{G_2} \times \ldots \times p_m^{\,G_m},$$

где p_m есть *m*-е простое число, а $G_1, G_2, \ldots$ – гёделевы номера соответственно формул $F_1, F_2, \ldots$.

Таким образом каждой формуле или последовательности формул ставится в соответствие единственное натуральное число. Не всякое натуральное является геделевым номером, но всякий гёделевый номер однозначно определяет соответствующее выражение. Это следует из теоремы о единственности разложения на простые множители: для целого числа, большего 1, существует единственный (с точностью до порядка) способ записать его в виде произведения степеней простых чисел.

---

[3] Если мы расширяем язык, добавляя счетное число новых функциональных символов, то нумерацию символов алфавита меняем следующим образом: предметным переменным сопоставляются простые числа, большие 10 и имеющие вид 3n+1, а новым функторам сопоставляются простые числа, большие 10 и имеющие вид 3n+2.

Оказывается функцию, осуществляющую гёделеву нумерацию формул элементарной арифметики, можно сделать даже примитивно-рекурсивной.

Гёделева нумерация устанавливает также счетность формул теории первого порядка.

Заметим, что метаматематическое утверждение «$\forall x$ есть начальная часть формулы $\forall x\ (x \times S(\mathbf{0}) = x)$» отражается внутрь теории, переходя в чисто арифметическое предложение «гёделевый номер выражения $\forall x$, равный $2^9 \times 3^{11}$, является делителем гёделевого номера полной формулы».

Лемма о рефлексии

– Он целовал вас, кажется?
– Боюсь, что это так.
– Но как же вы позволили?
– Ах, он такой чудак.
Он думал, что уснула я
И всё во сне стерплю,
Иль думал, что я думала,
Что думал он: я сплю!
Ковентри Патмор «О поцелуе»

Формулы *EA* «умеют говорить» только о натуральных числах. Чтобы формула Q могла говорить «о формулах и о себе», все формулы должны быть закодированы натуральными числами с помощью гёделевой нумерации. Пусть функция g осуществляет гёделеву нумерацию. Если теперь для некоторой формулы $A(x)$ и другой формулы B удается установить, что $\vdash A(g(B))$, то можно сказать: в теории *EA* доказано, что формула B «обладает свойством A».

Теорема 22 (лемма о рефлексии). Пусть $A(x)$ произвольная формула формальной арифметики, имеющая единственную свободную переменную x. Тогда можно построить замкнутую формулу B, такую, что $\vdash B \sim A(g(B))$.

Неформально говоря, для любого выразимого свойства формул, можно подобрать замкнутую формулу, «утверждающую», что она этим свойством обладает.

В своей знаменитой теореме о неполноте (см. ниже) К. Гёдель использовал конструкцию, которая составляет доказательство леммы о рефлексии, однако в общем виде он эту лемму не сформулировал. Доказательство этой леммы в различных формах см. в [16, с. 79; 9, с. 153–154; 21, с. 128; 25, с.15–16].

Мы приведем еще один вариант доказательства. Предполагаем, что в ***EA*** существует счетное множество функций (операций), причем каждый функциональный символ имеет свой код Гёделя. Если s – синтаксический объект в ***EA***, то, ради краткости записи, пусть $\#s$ обозначает его геделевский номер («код»), а $\langle s \rangle \equiv \langle \#s \rangle$ – соответствующий терм в ***EA***. В этих обозначениях формулировка теоремы следующая:

Пусть $A(x)$ – формула системы ***EA*** с единственной свободной переменной x. Тогда существует формула B системы ***EA***, такая, что

$\vdash_{EA} B \sim A(\langle B \rangle)$.

(B говорит: «*Я* обладаю свойством A»)

Доказательство.

Введем на множестве натуральных чисел $\boldsymbol{N}$ бинарное отношение $\sim$:

$n_1 \sim n_2 \Leftrightarrow$ существуют формулы A_1 и A_2 в ***EA***, такие, что $n_1 = \#A_1$, $n_2 = \#A_2$ и $\vdash A_1 \sim A_2$. Для теоремы достаточно показать, что для данной формулы $A(x)$ существует такое $n \in \boldsymbol{N}$, что

$$n \sim \#A(\langle n \rangle). \qquad (1)$$

Действительно, после этого взяв формулу B такую, что $n = \#B$ будем иметь $\vdash B \sim A(\langle n \rangle)$. Так как $\langle n \rangle = \langle \#B \rangle \equiv \langle B \rangle$, то получим, что $\vdash B \sim A(\langle B \rangle)$.

Для доказательства (1) введем следующее определение.

Для n, $m \in \boldsymbol{N}$ положим $h(n, m)$ = код результата подстановки терма $<m>$ вместо единственной свободной переменной в формулу с кодом n.

Функция h является вычислимой. Действительно, зная n, можно проверить, является ли n формулой. Если это так, то алгоритмическим способом можно отыскать свободную переменную в формуле и убедиться, что она единственная. По номеру m определяем соответствующий терм. Подстановка терма $<m>$ вместо свободной переменной в формулу описывается алгоритмом, как и последующее вычисление геделевского номера. По тезису Черча, функция h – частично рекурсивна. Она не определена в тех случаях, когда n не является номером формулы, или формула не обладает единственной свободной переменной, или m не является номером терма.

Когда значение функции определено, то имеем равенство

$$h(\#F(x), m) = \#F(<m>). \qquad (2)$$

Поскольку h – частично рекурсивная функция, то она представима в ***EA***, т. е. существует формула $H(x_1, x_2, x_3)$ в ***EA*** со свободными переменными x_1, x_2, x_3, такая, что

$\vdash H(<n>, <m>, <h(n, m)>)$,

$\vdash \forall y(\neg(y = <h(n, m)>) \supset \neg H(<n>, <m>, y))$.

Для любых термов x_1, x_2 если и существует, то единственный терм y, обладающим свойством $H(x_1, x_2, y)$. Поэтому мы можем ввести в ***EA*** (не усиливая первоначальную теорию) новый двухместный функциональный символ sub, такой, что

$$\vdash sub(<n>, <m>) = <h(n, m)>. \qquad (3)$$

Положим теперь натуральные $v = \#A(sub(x, x)))$ и $n = h(v, v)$. Тогда

$n = h(v, v) = h(\#A(sub(x, x)), v) = \#A(sub(<v>, <v>)$ (в силу (2)) $\sim \#A(<h(v,v)>) = \#A(<n>)$.

Отношение $\sim$ между числами имеет место потому, что в силу (3) из свойства равенства следует, что

$$\vdash A(sub(<n>, <m>)) \sim A(<h(n, m)>).$$

Теорема Гёделя о неполноте

Когда вопрос простой – ответ таким же будет.
Когда ж вопрос скрипит в зубах
Песком, попавшим в кашу,
Ответ – как прут, торчащий из земли.
Дверь без двери (Мумонкан)

Итак, лемма об рефлексии позволяет воспроизвести различные утверждения о свойствах формул. Можно ли воспроизвести парадокс лжеца средствами *EA*? Какие это будет иметь последствия?

Формула–аналог утверждения «Я лгу» должна утверждать: «Меня можно опровергнуть в *EA*» (вместо истинности и ложности в *EA* фигурируют доказуемость и опровержимость). Если

F: «$\neg F$ доказуема в *EA*»,

то

$\neg F$: «$\neg F$ недоказуема в *EA*»,

поэтому с тем же успехом мы можем рассматривать формулу

F: «F недоказуема в *EA*»,

она также «равносильна» парадоксу лжеца. Именно такой формулой занимался в свое время К. Гёдель, и мы не будем нарушать традицию.

Формулу F: «F недоказуема в EA» мы могли бы получить из леммы о рефлексии, если бы сумели изобразить в виде формулы EA свойство «формулу с номером x можно доказать в EA».

Определим отношение на множестве замкнутых термов в элементарной арифметике $Prove(d, t)$:

« формула с гёделевым номером t имеет доказательство с номером d».

Разумеется, по номеру можно определить, является ли он номером доказательства, принадлежащего теории EA. В самом деле, по номеру последовательности можно восстановить все входящие в нее формулы и порядок их расположения. Затем мы можем проверить, является ли каждая из этих формул аксиомой EA (логической или собственной) или же она получена из предыдущих формул последовательности с помощью правил вывода. Если для всех формул это так – анализируемый номер представляет доказательство. Можно построить даже программу, реализующую эту процедуру проверки без всякого вмешательства человека.

Геделева нумерация формул и доказательств элементарной арифметики проводится таким образом, чтобы было выразимым отношение $Prove$, т.е. чтобы выполнялось утверждение

«теорема y имеет доказательство x тогда и только тогда, когда $\vdash Prove(g(x),g(y))$».

Введем формулу $Pr(t)$ как сокращение для формулы $\exists d\ Prove(d,t)$. Предикат $Pr(t)$ утверждает доказуемость теоремы с гёделевым номером t.

Описанную выше гёделеву нумерацию можно провести таким образом, чтобы для любой формулы A элементарной арифметики из $\vdash A$ следует $\vdash Pr(g(A))$.

Нам понадобится следующее определение.

ω-противоречие

Формула в теории первого порядка $C(y)$ с одной свободной переменной y такая, что
а) $\vdash \exists y C(y)$,
б) для каждого n: $\vdash \neg C(\boldsymbol{n})$
называется ω-противоречием.

Формула $C(y)$ не дает «настоящего» противоречия (доказательства формулы вместе с ее отрицанием). Однако, если подобная формула $C(y)$ имеется, это означает все же, что в теории «не все в порядке».

Можно доказать, что «настоящее» противоречие (т.е. формула D такая, что $\vdash D$ и одновременно $\vdash \neg D$) означает также ω-противоречие.

Теорема 23 (теорема Гёделя о неполноте, 1931). Можно построить замкнутую формулу G из языка EA, такую, что

1) если G доказуема в EA, то теория EA противоречива,
2) если $\neg G$ доказуема в EA, то теория EA ω-противоречива.

Доказательство. Взяв в лемме о рефлексии формулу $\neg Pr(t)$ («формула с номером t не имеет доказательства в EA»), получим замкнутую формулу G, такую, что

$$\vdash G \sim \neg Pr(g(G)).$$

Формула G утверждает, что она недоказуема в EA. Теперь можно попытаться выяснить какое из двух утверждений $\vdash G$ или $\vdash \neg G$ выполнено. Разбор случаев приводит соответственно к утверждениям, что теория EA противоречива или ω-противоречива. Но мы не можем узнать, как «на самом деле» обстоит с доказуемостью G.

Если теория EA ω-непротиворечива, то теорема Гёделя утверждает, что обе формулы G и $\neg G$ нельзя доказать средствами EA. Но тогда недоказуемая формула G при стандартной интерпретации на универсуме натуральных чисел является истинной, поскольку она утверждает свою недоказуемость.

Таким образом, у нас выбор: 1) теория *EA* ω-противоречива или 2) *EA* ω-непротиворечива (и, следовательно, просто непротиворечива), но существует истинная формула, которую доказать нельзя. Второй вариант говорит, что теория *EA* неполна.[4]

Почему этой теореме придается такое большое значение? Сначала введем общепринятый термин. Замкнутую формулу *F* из языка теории ***T*** называют **неразрешимой** *в* ***T***, если ни *F*, ни ¬*F* нельзя доказать средствами ***T*** (*F* предсказывает вполне определенное свойство «объектов» теории ***T***, однако это предсказание нельзя средствами ***T*** ни доказать, ни опровергнуть).

Не следует, однако, думать, что нами *доказана* неполнота теории *EA*. Неразрешимость средствами *EA* формулы *G* будет доказана только..., если удастся доказать, что теория *EA* ω-непротиворечива (т.е. что в ней не могут возникать ω-противоречия). До тех пор мы вправе утверждать только, что доказали **несовершенство** аксиом *EA* – эти аксиомы либо ω-противоречивы, либо с их помощью нельзя решить некоторые проблемы, касающиеся натуральных чисел (одна такая проблема выражена в формуле *G* – несмотря на все наши разговоры о том, что *G* «занимается» собственной доказуемостью, *G* – замкнутая формула в языке *EA* и как таковая выражает вполне определенное свойство натуральных чисел).

Несовершенную систему аксиом следует совершенствовать. Может быть, мы «забыли» какие-то важные аксиомы? Следует найти их, присоединить к аксиомам *EA*, и в результате мы получим... совершенную систему?

К сожалению, рассуждения К. Гёделя проходят и для любого расширения *EA*. Никакие новые аксиомы не могут привести к «совершенной» системе аксиом арифметики. Метод Гёделя позволяет доказать *принципиальное* несовершенство всякой системы аксиом арифметики: каждая такая система неизбежно является либо ω-противоречивой, либо недостаточной для решения некоторых проблем, касающихся свойств натуральных чисел.

Из теоремы Гёделя вытекает следующий интересный факт. Согласно закону исключенного третьего, принятого в классической логике,

$\vdash A \vee \neg A$

для любой формулы *A*. Однако следует ли отсюда, что если *A* – замкнутая формула, то либо $\vdash A$, либо $\vdash \neg A$? Если взять формулу *G*, то из теоремы Геделя о неполноте следует, что ни *EA*, ни какая-либо другая серьезная математическая теория этим идеальным свойством обладать не могут – несмотря на постулирование закона исключенного третьего в их аксиомах!

Мы сформулировали теорему Гёделя для теории *EA*. Формальная арифметика *EA* представляет простейший уровень математических рассуждений – в которых участвуют только целые числа (и не участвуют произвольные действительные числа, не говоря уже о произвольных множествах Кантора). Более сложные рассуждения формализуются и более сложными (по сравнению с *EA*) формальными теориями. «Силу» этих более сложных теорий составляет прежде всего их способность обсуждать более сложные объекты (действительные числа, функции действительных и комплексных переменных и т.д.), которые недоступны в *EA*.

Каким образом выделить в некоторой формальной теории ***T*** ту ее часть, которая относится к компетенции *EA*? Этот вопрос решается очень естественно с помощью так называемых **относительных интерпретаций**. Чтобы воспроизвести в теории ***T*** арифметику, прежде всего какие-то объекты из области значений переменных ***T*** должны быть объявлены натуральными числами. Это связано с выделением в языке ***T*** некоторой

---

[4] Это очень близко к парадоеку лжеца, в котором мы встречаем предложение, выражающее свою собственную ложность. Но теперь из-за сделанной Гёделем замены «ложности» на недоказуемость появляется выход. Все доказуемые в *EA* формулы истинны (мы верим, что это так). Тогда, если бы все истинные формулы были доказуемы, получилось бы, что «ложно = недоказуемо», и возник бы парадокс лжеца. Выход состоит теперь в том, что не всн истинные формулы доказуемы.

формулы $N(x)$ (с единственной свободной переменной x), которая «утверждает», что x является натуральным числом. Далее, необходимо отобразить в теории $\boldsymbol{T}$ элементарные формулы EA, т.е. формулы вида $t_1 = t_2$, трактующие о значениях полиномов t_1, t_2 с натуральными коэффициентами.

Имея относительную интерпретацию EA в $\boldsymbol{T}$, в теории $\boldsymbol{T}$ можно доказать любое свойство натуральных чисел, которое доказуемо в EA. Учитывая роль системы натуральных чисел в математике, формальную теорию, в которой относительно интерпретируема теория EA (и которая содержит в этом смысле полноценное понятие натурального числа), будем называть **фундаментальной теорией**. Простейшей из фундаментальных теорий является, конечно, сама теория EA.

Теорема Гёделя, оказывается, справедлива для любой фундаментальной теории.

Понятие об ω-противоречивости несколько «портит» теорему К. Гёделя о неполноте. Однако сам Гёдель не пытался освободиться от него, впервые это удалось только Б. Россеру в 1936 г. – в теореме Россера вместо ω-противоречивости фигурирует «обычная» противоречивость.

Теорема 24 (теорема Гёделя в форме Россера). В языке всякой фундаментальной теории $\boldsymbol{T}$ найдется замкнутая формула $R_{\boldsymbol{T}}$ («выражающая» некоторое свойство натуральных чисел), такая, что если $\vdash_{\boldsymbol{T}} R_{\boldsymbol{T}}$ или $\vdash_{\boldsymbol{T}} \neg R_{\boldsymbol{T}}$, то теория $\boldsymbol{T}$ противоречива.

В этой формулировке сделано еще одно усиление теоремы о неполноте, не имеющее отношения к методу Россера. Сейчас речь идет о теории $\boldsymbol{T}$ с *произвольным* языком первого порядка, которая содержит в себе элементарную арифметику. Это освобождает нас от подозрений, что принципиальное несовершенство всякой системы аксиом арифметики кроется в неудачном выборе языка EA.

«Принцип несовершенства» К. Гёделя

Всякая фундаментальная теория несовершенна – она либо противоречива, либо недостаточна для решения всех возникающих в ней проблем.

В частности, если удалось аксиоматизировать всю математику, то она была бы несовершенной.

В качестве примера нефундаментальной теории можно назвать **арифметику Пресбургера**, которая получается из EA удалением символа умножения. В 1929 г. М. Пресбургер доказал полноту и непротиворечивость этой теории. В силу теоремы Геделя–Россера тем самым была доказана и ее нефундаментальность.

Нестандартное расширение *EA*

Поскольку EA является неполной системой (формула G – неразрешима), то ее можно пополнить (как можно дополнить абсолютную геометрию).[5]

Стандартное пополнение: добавить в качестве аксиомы формулу G. (Это соответствует расширению абсолютной геометрии в эвклидовом смысле.) Такое добавление кажется довольно безвредным и даже желательным, поскольку G всего навсего утверждает некоторую истину о системе натуральных чисел.

Нестандартное пополнение: (если следовать аналогии с ситуацией аксиомы параллельности) добавить в качестве аксиомы формулу $\neg G$. Но как мы можем даже подумать о такой ужасной, отвратительной вещи? В конце концов, если перефразировать Саккери[6], не является ли то, что утверждает $\neg G$, «противным самой природе натуральных чисел».

---

[5] Этот пункт излагается по книге Д. Хофштадтера [27].

[6] Джованни Саккери (1967–1733) – итальянский математик. Сделал попытку доказать пятый постулат Евклида о параллельных от противного. Саккери не допускал возможности отказа от пятого постулата о параллельных.

Проблема с подходом Саккери к геометрии (при доказательстве постулата о параллельных) заключалась в том, что, что он основывался на жестком понятии о том, что истинно и что ложно; он хотел доказать только то, что он считал истинным с самого начала. Рассмотрим беспристрастно, что означает добавление к системе *EA* новой аксиомы $\neg G$.

Подумаем только, на что была бы похожа современная математика, если бы люди не решили в свое время добавить к ней аксиом типа:

$\exists a(a+a=\mathbf{1})$ (рациональные числа);

$\exists a(\boldsymbol{S}a=\mathbf{0})$ (отрицательные числа);

$\exists a(a\times a=\mathbf{2})$ (иррациональные числа);

$\exists a(\boldsymbol{S}(a\times a)=\mathbf{0})$ (мнимые числа).

Хотя каждое из этих утверждений «противно природе ранее известных числовых систем», каждое из них в то же время означает значительное и замечательное расширение понятия целых чисел. $\neg G$ пытается открыть нам глаза на такую возможность. В прошлом каждое новое расширение системы натуральных чисел встречалось в штыки. Это заметно по названиям: «иррациональные», «мнимые». Оставаясь верными традиции, давайте назовем числа, которые порождает $\neg G$, супернатуральными, поскольку они противоречат всем понятиям разума и здравого смысла.

Если мы собираемся добавить $\neg G$ в качестве новой аксиомы *EA*, мы должны постараться понять, каким образом эта строчка может существовать с ω-противоречием. Ведь $\neg G$ утверждает, «что выполнено $\vdash \exists y Prove(y, g(G))$». При этом члены пирамидальной семьи с успехом утверждают, что

$\vdash \neg Prove(\mathbf{0}, g(G))$

$\vdash \neg Prove(\mathbf{1}, g(G))$

$\vdash \neg Prove(\mathbf{2}, g(G))$

$\vdash \neg Prove(\mathbf{3}, g(G))$

…

Это сбивает с толку, поскольку кажется совершеннейшим противоречием. Наша проблема заключается в том, что, так же как и в случае с расширенной геометрией, мы упрямо отказываемся модифицировать интерпретацию символов, несмотря на то, что прекрасно понимаем, что имеем дело с модифицированной системой. Мы хотим обойтись без добавления *хотя бы одного* символа – что, разумеется, оказывается невозможным.

Проблема разрешается, если мы интерпретируем ∃ как «существует некое *обобщенное* натуральное число» вместо «существует некое натуральное число». Одновременно с этим нам потребуется соответствующим образом изменить интерпретацию ∀. Это означает, что, кроме натуральных, мы открываем дверь для неких новых чисел. Это супернатуральные числа. Натуральные и супернатуральные числа вместе составляют *обобщенные натуральные числа*.

Кажущее противоречие теперь испаряется, поскольку пирамидальная семья все еще утверждает, «что никакое натуральное число не является номером доказательства *G*». Строчки этой семьи ничего не упоминают о супернатуральных числах, поскольку для них не существует *символов*. С другой стороны, $\neg G$ утверждает, что существует такое обобщенное натуральное число, которое является номером доказательства *G*. Противоречия больше нет. $EA+\neg G$ превращается в непротиворечивую систему, если её интерпретация включает супернатуральные числа.

Об аксиоматизации

Теорема К. Гёделя о неполноте породила множество рассуждений о том, что аксиоматический метод недостаточен для реконструкции «живого, содержательного» математического мышления. Аксиоматику сравнивали с прокрустовым ложем, которое не в состоянии вместить все богатство содержательной математики. Но разве могут в

математике какие-либо доказательные рассуждения происходить иначе, как по схеме «посылки – заключение»? Если так и всякое математическое рассуждение сводится к цепи заключений, то можно спросить: эти заключения происходят по *определенным* правилам (т.е. таким, которые не меняются от одного случая к другому и от одного математика к другому)? И если правила являются определенными, то, будучи функцией человеческого мозга, могут ли они быть такими, что их нельзя никак явно сформулировать? Если какие-либо «правила» нельзя явно сформулировать, то, следовательно, нельзя *доказать* их определенность! Ну, а полагать, что в математике кроме рассуждений (по определенным правилам) имеются «объекты», существующие независимо от этих рассуждений, означает впасть в обыкновенный платонизм работающего математика.

Таким образом, преждевременно говорить об ограниченности аксиоматизации – границы ее применимости, по-видимому, совпадают с границами применимости самой математики.

В процессе развития математических теорий аксиоматизация и интуиция взаимодействуют. Аксиоматизация «проясняет» интуицию, когда та «запуталась в себе». Но аксиоматизация влечет за собой и неприятные последствия: многие рассуждения, которые в интуитивной теории опытный специалист проводит очень быстро и представляет компактно, в аксиоматической теории оказываются очень громоздкими. Поэтому после замены интуитивной теории аксиоматической (особенно если эта замена неэквивалентна по причине недостатков интуитивной теории) специалисты развивают новую интуицию, которая восстанавливает способность теории к творческому развитию. Пример тому – история аксиоматизации теории множеств. Когда в интуитивной теории множеств Кантора в 1890-х гг. были обнаружены противоречия, от них удалось избавиться путем аксиоматизации. Естественно, что созданная аксиоматическая теория множеств Цермело–Френкеля отличалась от интуитивной теории Кантора не только формой, но и отдельными аспектами содержания. Для работы в новой теории специалисты развили модифицированную интуицию (в том числе особую интуицию множеств и классов). Ныне вполне нормальной считается работа в теории Цермело–Френкеля на интуитивном уровне. Именно на таком уровне доказываются серьезные новые теоремы этой теории.

Какую пользу дает аксиоматизация?

Во-первых, аксиоматизация позволяет «подправить» интуицию: устранить неточности, двусмысленности и парадоксы, которые иногда возникают из-за неполной контролируемости бессознательных процессов. Самый впечатляющий пример – историю аксиоматизации теории множеств, мы только что отметили.

Во-вторых, аксиоматизация позволяет подвергнуть подробному исследованию отношения между принципами теории (прежде всего, установить их зависимость или независимость), а также между этими принципами и теоремами теории. Для доказательства конкретной теоремы иногда требуются не все аксиомы теории, а только их часть. Исследования такого рода могут привести к созданию более общих теорий, которые применимы в различных конкретных теориях. Характерными примерами являются теория групп и многочисленные ее алгебраические ответвления.

В-третьих, нередко после аксиоматизации удается установить недостаточность данной теории для решения отдельных проблем, естественно возникающих в ней. Именно так произошло с континуум–гипотезой в теории множеств. В таких случаях можно ставить вопрос о необходимости совершенствования системы аксиом теории, о развитии альтернативных вариантов теории и т.д.

Построение формальной системы определяется, как правило, нуждами практики, но аксиоматические системы можно изучать независимо от каких-либо прикладных задачах. Поэтому существует, например, несколько вариантов формальных теорий множеств, причем не все они приводят к одним результатам (например, есть теории, в которых существуют множества с промежуточной мощностью между счетными и

континуумом, и есть теории, в которых таких множеств нет). Можно спросить, «а что на самом деле?». Каков «подлинный мир множеств»? Никакого «подлинного мира множеств», не зависящего от аксиом, с помощью которых он исследуется, разумеется, не существует. Математика даёт ответы на вопросы о реальном мире только после того, как мы выбрали, какой тип математики мы используем в данный момент.

> … почему математики не спотыкаются да и не спотыкались на протяжении нескольких веков ни об один из этих недоказуемых постулатов, почему и после Гёделя, то есть в настоящее время, математика способна идти своей дорогой – в любом направлении, куда ей заблагорассудится?
> … если математическая теорема может быть сформулирована в границах той же «шкалы», что и аксиомы, она принадлежит обычному миру математиков и будет иметь либо доказательство, либо опровержение. Но если ее изложение требует иного масштаба, тогда возникает опасность, что она – часть глубинного мира бесконеч<ных> … величин, но в любом случае латентного, того, что невозможно ни доказать, ни опровергнуть.
>
> Гильермо Мартинес
> Незаметные убийства

4.3 Теорема Гудстейна

Хотя неразрешимое самоссылочное утверждение Гёделя, несомненно, также говорит о каком-то свойстве натуральных чисел, математикам хотелось бы также обнаружить более «естественное» верное, но недоказуемое в арифметике Пеано утверждение. Одно из таких утверждений есть теорема Гудстейна.


Лъюис Гудстейн

Наследственное представление

Наследственным представлением натурального числа называется его представление в виде суммы степеней с основанием b, причем показатели степени также представляются в виде суммы степеней числа b и т. д., пока процесс не остановится.

Например, наследственное представление 266 по основанию 2 есть

$$\begin{aligned} 266 &= 2^8 + 2^3 + 2 \\ &= 2^{(2^{2+1})} + 2^{2+1} + 2 \end{aligned}$$

Последовательность Гудстейна

Для данного наследственного представления числа n по основанию b пусть $F_b(n)$ – неотрицательное целое число равное результату синтаксической замены в представлении n каждого b на b+1 (т. е. F_b есть оператор замены b на b+1).

Так как $266 = 2^{(2^{2+1})} + 2^{2+1} + 2$, то замена основания 2 на 3 дает

$$F_2(266) = 3^{(3^{3+1})} + 3^{3+1} + 3.$$

В построении следующей последовательности повторяется применение оператора F_b с последующим вычитанием 1. Первые девять членов суть

$G_0(266) = 266 = 2^{(2^{2+1})} + 2^{2+1} + 2$

$G_1(266) = F_2(266)-1 = 3^{(3^{3+1})} + 3^{3+1} + 2$
$= 443426488243037769948249630619149892886$
(39 цифр)

$G_2(266) = F_3(G_1(266))-1 = 4^{(4^{4+1})} + 4^{4+1} + 1$
$= 3231700607\ldots859611059596231681$ (617 цифр)

$G_3(266) = F_4(G_2(266))-1 = 5^{(5^{5+1})} + 5^{5+1}$ (10922 цифры)

$G_4(266) = F_5(G_3(266))-1 = 6^{(6^{6+1})} + 6^{6+1} - 1$
$= 6^{(6^{6+1})} + 5\cdot 6^6 + 5\cdot 6^5 + \ldots + 5\cdot 6 + 5 \approx 4\cdot 10^{217832}$

$G_5(266) = F_6(G_4(266))-1$
$= 7^{(7^{7+1})} + 5\cdot 7^7 + 5\cdot 7^5 + \ldots + 5\cdot 7 + 4 \approx 10^{4871822}$

$G_6(266) = F_7(G_5(266))-1$
$= 8^{8^{8+1}} + 5\cdot 8^8 + 5\cdot 8^5 + \ldots + 5\cdot 8 + 3 \approx 2\cdot 10^{121210686}$

$G_7(266) = F_8(G_6(266))-1$
$= 9^{9^{9+1}} + 5\cdot 9^9 + 5\cdot 9^5 + \ldots + 5\cdot 9 + 2 \approx 5\cdot 10^{3327237896}$

$G_8(266) = F_9(G_7(266))-1$
$= 10^{10^{10+1}} + 5\cdot 10^{10} + 5\cdot 10^5 + \ldots + 5\cdot 10 + 1 \approx 10^{10^{11}}$

Последовательность $\{G_k(n)\}$ называется последовательностью Гудстейна.

Теорема 25. (Гудстейн).

Для любого n существует такое k, что $G_k(n) = 0$.

Кажется невероятным, но это так. А чтобы в это поверить, мы рекомендовали бы читателю проделать вышеописанную процедуру, для начала – с числом «3».

$G_0(3) = 2 + 1$
$G_1(3) = F_2(3) - 1 = 3$
$G_2(3) = F_3(G_1) - 1 = 4 - 1 = 3$
$G_3(3) = F_4(G_2) - 1 = 2$
$G_4(3) = F_5(G_3) - 1 = 1$
$G_5(3) = F_6(G_6) - 1 = 0$

Попробуем проверить утверждение теоремы для n=4:

2^2
$3^3 - 1 = 2\times 3^2 + 2\times 3 + 2$
$2\times 4^2 + 2\times 4 + 1$
$2\times 5^2 + 2\times 5$
$2\times 6^2 + 6 + 5$
$2\times 7^2 + 7 + 4$
$2\times 8^2 + 8 + 3$
$2\times 9^2 + 9 + 2$
$2\times 10^2 + 10 + 1$
$2\times 11^2 + 11$
$2\times 12^2 + 12 - 1 = 2\times 12^2 + 11$
$2\times 13^2 + 10$

…
2×23^2
$2\times24^2 - 1 = 24^2 + 23\times24 + 23$
$25^2 + 23\times25 + 22$
…
$47^2 + 23\times47$
$48^2 + 23\times48 - 1 = 48^2 + 22\times48 + 47$
$49^2 + 22\times49 + 46$
…
$95^2 + 22\times95$
$96^2 + 22\times96 - 1 = 96^2 + 21\times96 + 95$
$97^2 + 21\times97 + 94$
…
$191^2 + 21\times191$
$192^2 + 21\times192 - 1 = 192^2 + 20\times192 + 191$
$193^2 + 20\times193 + 190$
…
$383^2 + 20\times383$
$384^2 + 20\times384 - 1 = 384^2 + 19\times384 + 383$
$385^2 + 19\times385 + 382$
…
$767^2 + 19\times767$
$768^2 + 19\times768 - 1 = 768^2 + 18\times768 + 767$
$769^2 + 18\times769 + 766$
…
$1535^2 + 18\times1535$
$1536^2 + 18\times1536 - 1 = 1536^2 + 17\times1536 + 1535$
$1537^2 + 17\times1537 + 1534$
…
…
…

Эта последовательность доходит до числа из 121 210 695 цифр (для сравнения заметим, что 1000000! содержит около 5 с половиной миллионов цифр), но потом числа только уменьшаются (так как уже не содержат в наследственном представлении основания) вплоть до 0. $G_k(4)$ впервые достигает 0 для $k=3(2^{402653211}-1) \approx 10^{121210695}$.

Набросок доказательства теоремы Гудстейна

Будем использовать трансфинитные ординалы.[7] Пусть, как обычно, ω обозначает ординал, равный порядковому типу множества натуральных чисел; а ординал ε_0 обозначает $\sup(\omega, \omega^{\omega}, \omega^{\omega^{\omega}}, ...)$.

Для любого натурального числа n и любого основания b пусть $B_b(n)$ обозначает выражение, полученное из наследственного представления числа n по основанию b, после синтаксической замены b на ординал ω. Например, так как наследственное представление 266 есть $2^{2^{2+1}} + 2^2 + 2$, то $B_2(266) = \omega^{\omega^{\omega+1}} + \omega^{\omega} + \omega$.

Очевидно, $B_b(n)$ есть ординал, меньший ε_0 и представленный в виде «экспоненциального полинома» относительно ω (канторова нормальная форма).

Имеем следующие свойства (для любых натуральных n и m и любого основания $b>1$):

1. $B_b(n)=0 \Leftrightarrow n=0$;

[7] См., например, [18].

2. $m < n \Leftrightarrow B_b(m) < B_b(n)$;
3. $B_b(n) = B_{b+1}(F_b(n))$ (поскольку в правой части основание b в наследственном представление n сначала заменяется на b+1, а потом b+1 заменяется на ω).

Зафиксируем n>0 и пусть $g_0, g_1, g_2, \ldots$ – последовательность Гудстейна для данного числа n, т.е. $g_k = G_k(n)$ при k=0,1,2,… Определим теперь соответствующую последовательность ординалов $a_0, a_1, a_2, \ldots$ по правилу $a_k = B_{k+2}(g_k)$ при k=0,1,2,… Пусть для всех k=0,1,2,… выполнено $g_k > 0$. Придем к противоречию, и тем самым теорема будет доказана.

Сравним a_k и a_{k+1}. Имеем $g_{k+1} = F_{k+2}(g_k)$-1 (здесь как раз используется предположение, что последовательность Гудстейна не содержит 0), поэтому

$$a_{k+1} = B_{k+3}(g_{k+1}) = B_{k+3}(F_{k+2}(g_k)\text{-}1) < B_{k+3}(F_{k+2}(g_k)) = (\text{по свойству (3)})\ B_{k+2}(g_k) = a_k.$$

Проиллюстрируем предыдущее неравенство на примере:

Пусть n=266. Тогда $g_0 = 266 = 2^{2^{2+1}} + 2^2 + 2$, и $a_0 = B_2(g_0) = \omega^{\omega^{\omega+1}} + \omega^\omega + \omega$.Также имеем $g_1 = 3^{3^{3+1}} + 3^3 + 2$, и $a_1 = B_3(g_1) = \omega^{\omega^{\omega+1}} + \omega^\omega + 2$. Очевидно, $a_1 < a_0$.

Таким образом, получили бесконечную убывающую последовательность ординалов $a_0 > a_1 > \ > a_2 > a_3 > \ldots$, что невозможно.

Гудстейн доказал теорему в 1944 году, используя трансфинитную индукцию [3]. В 1982 году Л. Кирби и Дж. Парис получили результат о том, что теорема Гудстейна формально недоказуема в рамках аксиоматической системы элементарной арифметики [4].

5 Сложность вычислений

Применение математики во многих приложениях требует, как правило, использования различных алгоритмов. Для решения многих задач не трудно придумать комбинаторные алгоритмы, сводящиеся к полному перебору вариантов. Но здесь вступает в силу различие между математикой и информатикой: в информатике недостаточно высказать утверждение о существовании некоторого объекта в теории и даже недостаточно найти конструктивное доказательство этого факта, т.е. алгоритм. Мы должны учитывать ограничения, навязываемые нам миром, в котором мы живем: необходимо, чтобы решение можно было вычислить, используя объем памяти и время, приемлемые для человека и компьютера.

Если дана задача, как найти для её решения эффективный алгоритм? А если алгоритм найден, как сравнить его с другими алгоритмами, решающими ту же задачу? Как оценить его качество? Вопросы такого рода интересуют и программистов, и тех, кто занимается теоретическим исследованием вычислений.

Введем в первую очередь некоторые понятия, связанные с асимптотической оценкой функций.

5.1 Асимптотические обозначения

Введем в первую очередь обозначение, связанное с асимптотической оценкой функций. Хотя во многих случаях эти обозначения используются неформально, полезно начать с точных определений [13]. На протяжении этого раздела встречающиеся в тексте функции отображают целые числа в действительные.

Θ–обозначение

Если $f(n)$ и $g(n)$ – некоторые функции, то запись $\boldsymbol{f(n) = \Theta(g(n))}$ означает, что найдутся такие $c_1, c_2 > 0$ и такое n_0, что $0 \le c_1 g(n) \le f(n) \le c_2 g(n)$ для всех $n \ge n_0$.
В этом случае говорят, что $g(n)$ является **асимптотически точной оценкой** для $f(n)$.

Разумеется, это обозначение следует употреблять с осторожностью: Установив, что $f_1(n) = \Theta(g(n))$ и $f_2(n) = \Theta(g(n))$, не следует заключать, что $f_1(n) = f_2(n)$!

Определение $\Theta(g(n))$ предполагает, что функции $f(n)$ и $g(n)$ асимптотически неотрицательны, т.е. неотрицательны для достаточно больших значений n. Заметим, что если f и g строго положительны, то можно исключить n_0 из определения (изменив константы c_1 и c_2 так, чтобы для малых n неравенство также выполнялось).

Это отношение симметрично: если $f(n) = \Theta(g(n))$, то $g(n) = \Theta(f(n))$.

Пример. Проверим, что $(1/2)n^2 - 3n = \Theta(n^2)$. Согласно определению надо указать положительные константы c_1, c_2 и число n_0 так, чтобы неравенства

$$c_1 n^2 \le n^2/2 - 3n \le c_2 n^2$$

выполнялось для всех $n \ge n_0$. Разделим на n^2:

$$c_1 \le \frac{1}{2} - \frac{3}{n} \le c_2.$$

Видно, что для выполнения второго неравенства достаточно положить $c_2 = ½$. Первое будет выполнено, если, например, $n_0 = 7$ и $c_1 = 1/14$.

Другой пример использования формального определения: покажем, что $6n^3 \ne \Theta(n^2)$. В самом деле, пусть найдутся такие c_2 и n_0, что $6n^3 \le c_2 n^2$ для всех $n \ge n_0$. Но тогда $n \le c_2/6$ для всех $n \ge n_0$ – что явно не так.

Отыскивая асимптотически точную оценку для суммы, мы можем отбрасывать члены меньшего порядка, которые при больших n становятся малыми по сравнению с основным слагаемым. Заметим также, что коэффициент при старшем члене роли не играет

(он может повлиять только на выбор констант c_1 и c_2). Например, рассмотрим квадратичную функцию $f(n) = an^2 + bn + c$, где a, b и c – некоторые константы и $a>0$. Отбрасывая члены младших порядков и коэффициент при старшем члене, находим, что $f(n) = \Theta(n^2)$. Чтобы убедиться в этом формально, можно положить $c_1 = a/4$, $c_2 = 7a/4$ и $n_0 = 2\times\max(|b|/a, \sqrt{|c|/a})$. Вообще, для любого полинома $p(n)$ степени d с положительным старшим коэффициентом имеем $p(n) = \Theta(n^d)$.

Упомянем важный частный случай использования Θ-обозначений: $\Theta(1)$ обозначает ограниченную функцию, отделенную от нуля некоторой положительной константой при достаточно больших значениях аргумента.

O– и Ω-обозначения

Запись $f(n) = \Theta(g(n))$ включает в себя две оценки: верхнюю и нижнюю. Их можно разделить.

Если $f(n)$ и $g(n)$ – некоторые функции, то

- запись $\boldsymbol{f(n) = O(g(n))}$ означает, что найдётся такая константа $c > 0$ и такое n_0, что $f(n) \le cg(n)$ для всех $n \ge n_0$;
- запись $\boldsymbol{f(n) = \Omega(g(n))}$ означает, что найдётся такая константа $c > 0$ и такое n_0, что $0 \le cg(n) \le f(n)$для всех $n \ge n_0$.

Теорема 26.

1. Для любых двух функций $f(n)$ и $g(n)$ свойство $f(n) = \Theta(g(n))$ выполнено тогда и только тогда, когда $f(n) = O(g(n))$ и $f(n) = \Omega(g(n))$.
2. Для любых двух функций $f(n)$ и $g(n)$ свойство $f(n) = O(g(n))$ выполнено тогда и только тогда, когда $g(n) = \Omega(f(n))$.

Как мы видели, $an^2 + bn + c = \Theta(n^2)$ (при $a>0$). Поэтому $an^2 + bn + c = O(n^2)$. Другой пример: при $a>0$ можно записать $an + b = O(n^2)$ (положим $c = a + |b|$ и $n_0 = 1$). Заметим, что в этом случае $an + b \ne \Omega(n^2)$ и $an + b \ne \Theta(n^2)$.

Работая с символами O, Θ и Ω мы имеем дело с *односторонними* равенствами – эти символы могут стоять только справа от знака =.

Введенные нами определения обладают некоторыми свойствами транзитивности, рефлексивности и симметричности.

Транзитивность

$f(n) = \Theta(g(n))$ и $g(n) = \Theta(h(n))$ влечет $f(n) = \Theta(h(n))$.
$f(n) = O(g(n))$ и $g(n) = O(h(n))$ влечет $f(n) = O(h(n))$.
$f(n) = \Omega(g(n))$ и $g(n) = \Omega(h(n))$ влечет $f(n) = \Omega(h(n))$.

Рефлексивность

$f(n) = \Theta(f(n)), f(n) = O(f(n)), f(n) = \Omega(f(n))$

Симметричность

$f(n) = \Theta(g(n)) \Leftrightarrow g(n) = \Theta(f(n))$

Сравнение роста функций

Определение. Говорят, что функция g **мажорирует** функцию f, если $f(n) = O(g(n))$.

Символ $O(g(n))$ читается как **«*O* большое от *g*(*n*)»**; при этом говорят, что $f(n)$ имеет порядок O большое от $g(n)$.

Отношение «О большое» наиболее полезно для сравнения асимптотического роста функции. Рассмотрим это отношение для конкретных функций.

Для положительных целых чисел r и s следующую теорему можно доказать методом индукции. Справедливость утверждения теоремы для положительных рациональных чисел r и s можно показать, не используя логарифмы.

Теорема 27. Если r и s – действительные числа, $r \le s$ и $n > 1$, тогда $n^r \le n^s$. Следовательно, $n^r = O(n^s)$.

Доказательство. Функция $\ln(x)$ – возрастающая, поэтому $a \le b$ тогда и только тогда, когда $\ln(a) \le \ln(b)$. Отсюда $n^r \le n^s$ тогда и только тогда, когда $\ln(n^r) \le \ln(n^s)$, что, в свою очередь, выполняется тогда и только тогда, когда $r \ln(n) \le s \ln(n)$, т.е. тогда и только тогда, когда $r \le s$, поскольку $\ln(n)$ для $n > 1$ – величина положительная.

Следующие теоремы показывают, что свойство функции иметь порядок $O(g(n))$ замкнуто относительно операций сложения и умножения на число.

Теорема 28. Если $f(n) = O(g(n))$, то $cf(n) = O(g(n))$.

Доказательство. По определению, $|f(n)| \le k\,|g(n)|$ для некоторого действительного числа k и всех n, больших или равных некоторому целому числу m. Поэтому

$$|cf(n)| \le k|c|\,|g(n)|$$

и $cf(n) = O(g(n))$.

Теорема 29. Если $f(n) = O(g(n))$ и $h(n) = O(g(n))$, то $(f+h)(n) = O(g(n))$.

Доказательство. По определению, для некоторого постоянного k и некоторого целого числа m_1 имеем $|f(n)| \le k\,|g(n)|$ для всех $n > m_1$. Опять же по определению, для некоторого постоянного l и некоторого целого числа m_2 имеем $|h(n)| \le l\,|g(n)|$ для всех $n > m_2$. Пусть $m = \max(m_1, m_2)$. Следовательно, для всех $n > m$

$$\begin{gathered}|f(n)+h(n)| \le |f(n)|+|h(n)| \le \\ \le k\,|g(n)| + l\,|g(n)| = \\ = (k + l)|g(n)|\end{gathered}$$

и $(f+g)(n) = O(g(n))$.

Теорема 30. Если $f(n) = O(g(n))$ и $h(n) = O(e(n))$, то $(f\times h)(n) = O(g\times e)(n))$.

Доказательство. По определению, для некоторого постоянного k и некоторого целого числа m_1 имеем $|f(n)| \le k\,|g(n)|$ для всех $n > m_1$. Опять же по определению, для некоторого постоянного l и некоторого целого числа m_2 имеем $|h(n)| \le l\,|e(n)|$ для всех $n > m_2$. Пусть $m = \max(m_1, m_2)$. Следовательно, для всех $n > m$

$$\begin{gathered}|f(n)\times h(n)| = |f(n)|\times|h(n)| \le \\ \le k\,|g(n)|\;\; l\,|e(n)| = \\ = (k\,l)|g(n)\times e(n)|\end{gathered}$$

и $(f\times g)(n) = O(g\times e)(n))$.

Следующая теорема устанавливает мажоранту для полинома.

Теорема 31. Если $p(n) = a_k n^k + a_{k-1}n^{k-1} + \ldots + a_1 n + a_0$, то $p(n)=O(n^k)$.

Доказательство.

$$|p(n)| \le |a_k n^k + a_{k-1}n^{k-1} + \ldots + a_1 n + a_0| \le$$

(в силу неравенства треугольника: $|A+B| \le |A|+|B|$)

$$\le |a_k n^k| + |a_{k-1} n^{k-1}| + \ldots + |a_1 n| + |a_0| =$$
$$= |a_k|\, n^k + |a_{k-1}|\, n^{k-1} + \ldots + |a_1| n + |a_0| \le$$
(по теореме 1)
$$\le |a_k|\, n^k + |a_{k-1}|\, n^k + \ldots + |a_1|\, n^k + |a_0|\, n^k =$$
$$= (|a_k| + |a_{k-1}| + \ldots + |a_1| + |a_0|)\, n^k$$

и $p(n)=O(n^k)$.

Теорема 32. Для целых чисел a и b, больших единицы, $\log_a(n) = O(\log_b(n))$.

Доказательство. Следует непосредственно из равенства $\log_a(n) = \dfrac{\log_b(n)}{\log_a(b)}$.

Теорема 33. Пусть n – неотрицательное целое число. Тогда $n < 2^n$ и, следовательно, $n = O(2^n)$.

Доказательство. Воспользуемся индукцией, имея для $n = 0$, $0 < 2^0 = 1$. Допустим, что $k < 2^k$, тогда

$$k+1 \le k+k \le 2^k + 2^k = 2^{k+1}$$

и, по индукции, $n < 2^n$.

Пусть R – отношение на множестве функций, определенное как fRg, если $f(n) = O(g(n))$. Очевидно, отношение R рефлексивно. Приведенная ниже теорема утверждает, что это отношение также и транзитивно. Доказательство теоремы предоставляется читателю.

Теорема 34. Если $f(n) = O(g(n))$ и $g(n) = O(h(n))$, то $f(n) = O(h(n))$.

Следующая теорема дает ответ на вопрос о том, какие функции могут выступать в роли мажорант для других функций.

Теорема 35. Для целых чисел a, больших единицы, $\log_a(n) = O(n)$.

Доказательство. Согласно теореме 33, имеет место неравенство $n < 2^n$. Поэтому $\log_2(n) < \log_2(2^n) = n$ и $\log_2(n) = O(n)$. Поскольку по теореме 32 имеем $\log_a(n) = O(\log_2(n))$, то по теореме 34 получаем $\log_a(n) = O(n)$.

Доказательство сформулированных ниже теорем предоставляем читателям.

Теорема 36. Пусть n – неотрицательное целое число, тогда $n! < n^n$ и, следовательно, $n! = O(n^n)$.

Теорема 37. Пусть $a > 1$ и n – неотрицательное целое число, тогда $\log_a(n!) \le n \log_a(n)$ и, следовательно, $\log_a(n!) = O(n \log_a(n))$.

Пример 1. Определим число арифметических операций, необходимых для сложения двух матриц.

Пусть матрицы имеют размеры $m \times k$. Тогда алгоритм сложения матриц $A + B = C$ можно описать на Паскале следующим образом:

```
for i:=1 to m do
     for j:= 1 to k do
           C[i,j]:=A[i,j] + B[i,j];
```

Как видим, сложение выполняется для каждого i и каждого j. Поскольку i принимает m значений, а j принимает k значений, то выполняется mk операций сложения. Пусть $n = \max(m, k)$. Тогда число выполняемых арифметических операций имеет порядок $O(n^2)$.

Пример 2. Определим число арифметических операций, необходимых для умножения двух матриц.

Пусть матрицы A и B имеют размеры $m \times p$ и $p \times k$, соответственно. Тогда алгоритм умножения матриц $A \times B = C$ можно описать на Паскале следующим образом:

```
for i:=1 to m do
     for j:= 1 to k do
        begin
```

```
C[i,j]:=0;
for s:= 1 to p do
   C[i,j]:= C[i,j]+A[i,s]* B[s,j];
end;
```

Поскольку s принимает значения от 1 до p, то выполняется p операций сложения и p операций умножения. Величина s изменяется от 1 до p для каждого i и каждого j, поэтому s пробегает значения от 1 до p mk раз. Таким образом, выполняется mkp операций сложения и столько же операций умножения. Следовательно, всего выполняется $2mkp$ операций. Пусть $n = \max(m,k,p)$. Тогда число выполняемых арифметических операций имеет порядок $O(n^3)$.

Пример 3. Сравним количество операций, которое требуется для непосредственного вычисления значения многочлена традиционным способом и по схеме Горнера.

Пусть требуется вычислить $p(c)$, где $p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$. Если $p(c)$ вычисляется непосредственно, то для подсчета c^k требуется выполнить k–1 операций умножения. Еще одна операция нужна для умножения на a_k, так что вычисление $a_k x^k$ требует k операций умножения. Таким образом, нужно выполнить $1 + 2 + \ldots + n = n(n\text{-}1)/2$ умножений. Для того чтобы найти сумму n+1 слагаемых, требуется выполнить n сложений, так что общее число арифметических операций равно $n(n\text{-}1)/2 + n$ и имеет порядок $O(n^2)$.

При вычислении полинома $a_4x^4 + a_3x^3 + a_2x^2 + a_1x + a_0$ по схеме Горнера, мы переписываем полином в виде $x(x(x(a_4x + a_3) + a_2) + a_1) + a_0$ и замечаем, что выражение включает четыре операции умножения и четыре операции сложения. Очевидно, в общем случае $p(x) = x(x(\ldots (x(a_nx+a_{n-1}) + a_{n-2}) + \ldots a_2) + a_1) + a_0$ включает n операций сложения и n операций умножения. Таким образом, общее число арифметических операций равно $2n$ и имеет порядок $O(n)$.

Задача. Расположите следующие функции в порядке увеличения скорости роста (каждая функция есть O(следующая)), не исключено, что некоторые функции имеют одинаковую скорость:

2^n, $(1+n)!$, $\ln(n!)$, e^n, $n\,2^n$, $n^{\ln(\ln n)}$, 1, $n(\ln n)$, $2^{(2^n)}$, $(\mathrm{trunc}(\ln n))!$, $\ln n$, n^2, $\ln(\ln n)$, $\sqrt{2}^{\ln n}$, $n!$, e, $4^{\ln n}$, $2^{(2^{1+n})}$, $(\ln n)^2$, n, $\sqrt{\ln n}$, $2^{\ln n}$, $(3/2)^n$, n^3, $(\ln n)^{\ln n}$, $2^{\sqrt{\ln n}}$.

5.2 Алгоритмы и их сложность

Класс однородных вычислительных задач мы будем называть **проблемой** (также используется понятие **массовая задача** или **абстрактная задача**). Индивидуальные случаи проблемы Q мы будем называть **частными случаями** проблемы Q. Мы можем, например, говорить о проблеме умножения матриц. Частные случаи этой проблемы суть пары матриц, которые нужно перемножить.

Более формально, мы принимаем следующую абстрактную модель вычислительной задачи. Абстрактная задача есть произвольное бинарное отношение Q между элементами двух множеств – множества **условий** (или входных данных) I и множества **решений** S. Например, в задаче умножения матриц входными данными являются две конкретные матрицы – сомножители, а матрица – произведение является решением задачи. В задаче SHORTEST-PATH (поиска кратчайшего пути между двумя заданными вершинами некоторого неориентированного графа $G = (V,E)$) условием (элементом I) является тройка, состоящая из графа и двух вершин, а решением (элементом S) – последовательность вершин, составляющих требуемый путь в графе. При этом один элемент множества I

может находиться в отношении Q с несколькими элементами множества S (если кратчайших путей между данными вершинами несколько).

Нам бы хотелось связать с каждым частным случаем проблемы некоторое число, называемое его **размером**, которое выражало бы меру количества входных данных. Например, размером задачи умножения матриц может быть наибольший размер матриц-сомножителей. Размером задачи о графах может быть число ребер данного графа.

Решение задачи на компьютере можно осуществлять с помощью различных алгоритмов. Прежде чем подавать на вход алгоритма исходные данные (то есть элемент множества I), надо договориться о том, как они представляются «в понятном для компьютера виде»; мы будем считать, что исходные данные закодированы последовательностью битов. Формально говоря, **представлением** элементов некоторого множества M называется отображение e из M во множество битовых строк. Например, натуральные числа 0, 1, 2, 3, … обычно представляют битовыми строками 0, 1, 10, 11, 100, … (при этом, например, $e(17) = 10001$).

Фиксировав представление данных, мы превращаем абстрактную задачу в **строковую**, для которой входным данным является битовая строка, представляющая исходное данное абстрактной задачи. Естественно считать размером строковой задачи длину строки.

Асимптотическая временная сложность

Будем говорить, что алгоритм **решает** строковую задачу за **время $O(T(n))$**, если на входном данном битовой строки длины n алгоритм работает время $O(T(n))$.

В качестве временной оценки работы алгоритма вместо общего числа шагов мы можем подсчитывать число шагов некоторого вида, таких как арифметические операции при алгебраических вычислениях, число сравнений при сортировке или число обращений к памяти.

Можно было подумать, что колоссальный рост скорости вычислений, вызванный появлением нынешнего поколения компьютеров, уменьшит значение эффективных алгоритмом. Однако происходит в точности противоположное. Так как компьютеры работают все быстрее и мы можем решать все большие задачи, именно сложность алгоритма определяет то увеличение размера задачи, которое можно достичь с увеличением скорости машины.

Следуя [8], рассмотрим это более подробно. Допустим, у нас есть пять алгоритмов A_1–A_5 со следующими временными сложностями:

Алгоритм	Асимптотическая временная сложность
A_1	$O(n)$
A_2	$O(n \log n)$
A_3	$O(n^2)$
A_4	$O(n^3)$
A_5	$O(2^n)$

Пусть единицей времени будет одна миллисекунда и мультипликативные константы в точных оценках временной сложности во всех алгоритмах равны 1. Тогда алгоритм A_1 может обработать за одну секунду вход размера 1000, в то время как A_5 – вход размера не более 9. В таблице 1 приведены размеры задач, которые можно решить за одну секунду, одну минуту и один час каждым из этих пяти алгоритмов.

Таблица 1. Границы размеров задач, определяемые скоростью роста сложности.

Алгоритм	Асимптотическая временная сложность	Максимальный размер задачи		
		1 с	1 мин	1 ч
A_1	O(n)	1000	6×10^4	$3{,}6\times10^6$
A_2	O($n \log n$)	140	4893	$2{,}0\times10^5$
A_3	O(n^2)	31	244	1897
A_4	O(n^3)	10	39	153
A_5	O(2^n)	9	15	21

Предположим, что следующее поколение компьютеров будет в 10 раз быстрее нынешнего. В таблице 2 показано, как возрастут размеры задач, которые мы сможем решить благодаря этому увеличению скорости.

Таблица 2. Эффект десятикратного ускорения

Алгоритм	Асимптотическая временная сложность	Максимальный размер задачи	
		До ускорения	После ускорения
A_1	O(n)	S_1	10 S_1
A_2	O($n \log n$)	S_2	Примерно 10 S_2 для больших S_2
A_3	O(n^2)	S_3	3,16 S_3
A_4	O(n^3)	S_4	2,15 S_4
A_5	O(2^n)	S_5	S_5 + 3,3

Заметим, что для алгоритма A_5 десятикратное увеличение скорости увеличивает размер задачи, которую можно решить, только на три, тогда как для алгоритма A_3 размер задачи более чем утраивается.

Вместо эффекта увеличения скорости рассмотрим теперь эффект применения более действенного алгоритма. Вернемся к таблице 1. Если в качестве основы для сравнения взять 1 мин, то, заменяя алгоритм A_4 алгоритмом A_3, можно решить задачу в 6 раз большую, а заменяя A_4 на A_2, можно решить задачу, большую в 125 раз. Эти результаты производят гораздо большее впечатление, чем двукратное улучшение, достигаемое за счет десятикратного увеличения скорости. Если в качестве основы для сравнения взять 1 ч, то различие оказывается еще значительнее. Отсюда мы заключаем, что асимптотическая скорость алгоритма служит важной мерой качественности алгоритма, причем такой мерой, которая обещает стать еще важнее при последующем увеличении скорости вычислений.

Несмотря на то, что основное внимание здесь уделяется порядку роста величин, надо понимать, что больший порядок сложности алгоритма может иметь меньшую мультипликативную постоянную, чем малый порядок роста сложности другого алгоритма. В таком случае алгоритм с быстро растущей сложностью может оказать предпочтительнее для задач с малым размером – возможно, даже для всех задач, которые нас интересуют.

5.3 Сложность задач

Сложность задачи – это асимптотическая временная сложность наилучшего алгоритма, известного для ее решения.

Основной вопрос теории сложности: насколько успешно или с какой стоимостью может быть решена заданная проблема *Q*? Мы не имеем в виду никакого конкретного алгоритма решения *Q*. Наша цель – рассмотреть все возможные алгоритмы решения *Q* и попытаться сформулировать утверждение о вычислительной сложности, внутренне присущей *Q*. В то время как всякий алгоритм *A* для *Q* дает верхнюю оценку величины сложности *Q*, нас интересует нижняя оценка. Знание нижней оценки представляет интерес математически и, кроме того, руководит нами в поиске хороших алгоритмов, указывая, какие попытки заведомо будут безуспешны.

Быстрыми являются линейные алгоритмы, которые обладают сложностью порядка O(*n*), где *n* – размерность входных данных. К линейным алгоритмам относится школьный алгоритм нахождения суммы десятичных чисел, состоящих из n_1 и n_2 цифр. Сложность этого алгоритма – $O(n_1 + n_2)$. Есть алгоритмы, которые быстрее линейных, например, алгоритм двоичного поиска в линейном упорядоченном массиве имеет сложность O(log *n*), *n* – длина массива.

Другие хорошо известные алгоритмы – деление, извлечение квадратного корня, решение систем линейных уравнений и др. – попадают в более общий класс полиномиальных алгоритмов.

Полиномиальным алгоритмом (или алгоритмом полиномиальной временной сложности, или алгоритмом принадлежащим классу **P**) называется алгоритм, у которого временная сложность равна $O(n^k)$, где *k* – положительное целое число. Алгоритмы, для временной сложности которых не существует такой оценки, называются *экспоненциальными* и такие задачи считаются **труднорешаемыми**. Понятие полиномиально разрешимой задачи принято считать уточнением идеи «практически разрешимой» задачи. Чем объясняется такое соглашение?

Во-первых, используемые на практике полиномиальные алгоритмы обычно действительно работают довольно быстро. Конечно, трудно назвать практически разрешимой задачу, которая требует времени $\Theta(n^{100})$. Однако полиномы такой степени в реальных задачах почти не встречаются.

Второй аргумент в пользу рассмотрения класса полиномиальных алгоритмов – тот факт, что объем этого класса не зависит от выбора конкретной модели вычислений (для достаточно широкого класса моделей) [13]. Например, класс задач, которые могут быть решены за полиномиальное время на последовательной машине с произвольным доступом (RAM), совпадает с классом задач, полиномиально разрешимых на машинах Тьюринга. Класс будет тем же и для моделей параллельных вычислений, если, конечно, число процессоров ограничено полиномом от длины входа.

В-третьих, класс полиномиально разрешимых задач обладает естественными свойствами замкнутости. Например, композиция двух полиномиальных алгоритмов (выход первого алгоритма подается на вход второго) также работает полиномиальное время. Объясняется это тем, что сумма, произведение и композиция многочленов снова есть многочлен.

Приведем примеры классификации задач по их сложности.

Класс P

- Рассортировать множество из *n* чисел. Сложность поведения в среднем порядка O(*n* log *n*) для быстрого алгоритма Хоара [22, стр.316–321].
- Найти эйлеровый цикл на графе из *m* ребер. В силу теоремы Эйлера мы имеем необходимое и достаточное условие для существования эйлерова цикла и проверка этого условия есть алгоритм порядка O(*m*).
- Задача Прима-Краскала. *Дана плоская страна и в ней n городов. Нужно соединить все города телефонной связью так, чтобы общая длина телефонных линий была минимальной.* В терминах теории графов задача Прима-Краскала выглядит следующим образом: *Дан граф с n вершинами; длины ребер заданы матрицей* (*d[i,j]*)*, i,j = 1,...,n.*

Найти остовное дерево минимальной длины. Эта задача решается с помощью жадного алгоритма сложности O(n log n)[22, стр.357–358].

- Кратчайший путь на графе, состоящем из n вершин и m ребер. Сложность алгоритма O(m n) [22, стр.377–382].
- Связные компоненты графа. Определяются подмножества вершин в графе (связные компоненты), такие, что две вершины, принадлежащие одной и той же компоненте, всегда связаны цепочкой дуг. Если n – количество вершин, а m – количество ребер, то сложность алгоритма O(n+m) [22, стр.364–365].
- Быстрое преобразование Фурье [8, стр. 284–302], требующее O(n log n) арифметических операций, – один из наиболее часто используемых алгоритмов в научных вычислениях.
- Умножение целых чисел. Алгоритм Шёнхаге-Штрассена [8, стр. 304–308]. Сложность алгоритма порядка O(n log n log log n). Отметим, что школьный метод для умножения двух n-разрядных чисел имеет сложность порядка O(n^2).
- Умножение матриц. Алгоритм Штрассена [8, стр. 259–261] имеет сложность порядка O($n^{\log 7}$), для умножения двух матриц размера n×n. Очевидный алгоритм имеет порядок сложности O(n^3).

Класс E: задачи, экспоненциальные по природе

К экспоненциальным задачам относятся задачи, в которых требуется построить множество всех подмножеств данного множества, все полные подграфы некоторого графа или же все поддеревья некоторого графа.

Существует масса примеров задач с экспоненциальной сложностью. Например, чтобы вычислить $2^{(2^k)}$ для заданного натурального k, нам только для записи конечного ответа потребуется около 2^n шагов (где n – число цифр в двоичной записи k), не говоря даже о самом вычислении.

Задачи не попадающие ни в класс P, ни в класс E

На практике существуют задачи, которые заранее не могут быть отнесены ни к одному из рассмотренных выше классов. Хотя в их условиях не содержатся экспоненциальные вычисления, однако для многих из них до сих пор не разработан эффективный (т.е. полиномиальный) алгоритм.

К этому классу относятся следующие задачи [14, с. 207]:

- задача о выполнимости: существует ли для данной булевской формулы, находящейся в КНФ, такое распределение истинностных значений, что она имеет значение истина?
- задача коммивояжера (Коммивояжер хочет объехать все города, побывав в каждом ровно по одному разу, и вернуться в город, из которого начато путешествие. Известно, что переезд из города i в город j стоит $c(i, j)$ рублей. Требуется найти путь минимальной стоимости.);
- решение систем уравнений с целыми переменными;
- составление расписаний, учитывающих определенные условия;
- размещение обслуживающих центров (телефон, телевидение, срочные службы) для максимального числа клиентов при минимальном числе центров;
- оптимальная загрузка емкости (рюкзак, поезд, корабль, самолёт) при наименьшей стоимости;
- оптимальный раскрой (бумага, картон, стальной прокат, отливки), оптимизация маршрутов в воздушном пространстве, инвестиций, станочного парка;
- задача распознавания простого числа; самый лучший в настоящее время тест на простоту имеет сложность порядка O($L(n)^{L(L(L(n)))}$), где $L(n)$ – количество цифр в числе n

(выражение $L(L(L(n)))$ стремиться к бесконечности очень медленно; первое число, для которого $L(L(L(n))) = 2$, равно $10^{999999999}$) [7, стр. 102].

6 *NP*-полнота

В этом разделе мы рассмотрим класс задач, называемых «***NP***-полными». Для этих задач не найдены полиномиальные алгоритмы, однако не доказано, что таких алгоритмов не существует. Примеры таких задач приведены в конце предыдущего раздела. Изучение ***NP***-полных задач связано с так называемым вопросом ***P*≠*NP***. Этот вопрос был поставлен в 1971 году и является сейчас одной из наиболее интересных и сложных проблем теории вычислений.

Большинство специалистов полагают, что ***NP***-полные задачи нельзя решить за полиномиальное время. Дело в том, что если хотя бы для одной ***NP***-полной задачи существует решающий ее полиномиальный алгоритм, то и для всех ***NP***-полных задач такие алгоритмы существуют. В настоящее время известно очень много ***NP***-полных задач – многие из них практически важные. Все попытки найти для них полиномиальные алгоритмы оказались безуспешными. По-видимому, таких алгоритмов нет вовсе.

Зачем программисту знать о ***NP***-полных задачах? Если для некоторой задачи удается доказать ее ***NP***-полноту, есть основания считать ее практически неразрешимой. В этом случае лучше потратить время на построение приближенного алгоритма, чем продолжать искать быстрый алгоритм, решающий ее точно.

Содержание этого раздела взято из [13]. Для понимания материала могут потребоваться элементарные понятия теории графов.

6.1 Задачи разрешения и задачи оптимизации

В теории ***NP***-полноты рассматриваются только **задачи разрешения** – задачи, в которых требуется дать ответ «да» или «нет» на некоторый вопрос. Формально задачу разрешения можно рассматривать как функцию, отображающую множество условий I во множество $\{0,1\}$ (1 = «да», 0 = «нет»). Многие задачи можно тем или иным способом преобразовать к такому виду. Например, с задачей SHORTEST-PATH поиска кратчайшего пути в графе связана задача разрешения PATH: «по заданному графу $G = (V, E)$, паре вершин $u, v \in V$ и натуральному числу k определить, существует ли в G путь между вершинами u и v, длина которого не превосходит k». Пусть четверка $i = \langle G, u, v, k \rangle$ является условием задачи PATH. Тогда $\mathrm{PATH}(i) = 1$, если длина кратчайшего пути между вершинами u и v не превосходит k, и $\mathrm{PATH}(i) = 0$ в противном случае.

Часто встречаются **задачи оптимизации**, в которых требуется минимизировать или максимизировать значение некоторой величины. Прежде чем ставить вопрос о ***NP***-полноте таких задач, их следует преобразовать в задачу разрешения. Обычно в качестве такой задачи разрешения рассматривают задачу проверки, является ли некоторое число верхней (или нижней) границей для оптимизируемой величины. Так, например, мы перешли от задачи оптимизации SHORTEST-PATH к задаче разрешения PATH, добавив в условие задачи границу длины пути k.

Если после этого получается задача, не имеющая полиномиального алгоритма разрешения, то тем самым установлена трудность исходной задачи. В самом деле, если для оптимизационной задачи имеется быстрый алгоритм, то и полученную из неё задачу разрешения можно решить быстро (надо просто сравнить ответ этого алгоритма с заданной границей). Таким образом, можно исследовать временную сложность задач оптимизации.

Мы будем использовать строковое представление задач. Для каждого представления e множества I входов абстрактной задачи Q мы получаем свою строковую задачу, которую мы в дальнейшем обозначаем $e(Q)$. Мы не будем подробно описывать используемое представление в конкретных задачах, считая, что оно выбрано достаточно разумно и экономно (целые числа задаются двоичной записью, конечные множества – списком элементов и т.п.). Нам понадобятся также следующие определения.

Будем говорить, что функция f: $\{0,1\}^* \to \{0,1\}^*$ ($\{0,1\}^*$ обозначает множество битовых строк) **вычислима за полиномиальное время**, если существует полиномиальный алгоритм A, который для любого $x \in \{0, 1\}^*$ выдает результат $f(x)$.

Рассмотрим теперь множество I условий произвольной абстрактной задачи разрешения. Два представления e_1 и e_2 этого множества называются **полиномиально связанными**, если существуют две вычислимые за полиномиальное время функции f_{12} и f_{21}, для которых $f_{12}(e_1(i)) = e_2(i)$ и $f_{21}(e_2(i)) = e_1(i)$ для всякого $i \in I$. Это значит, что e_1–представление входа может быть за полиномиальное время получено из e_2–представления и наоборот. В этом случае не имеет значения, какое из двух полиномиально связанных представлений выбрать, как показывает следующая теорема.

Теорема 38. Пусть Q – абстрактная задача разрешения с множеством условий I, а e_1 и e_2 – полиномиально связанные представления для элементов множества I. Предположим, что множество всех строк, которые являются e_1–представлениями элементов Q, разрешимо за полиномиальное время, и что аналогичное свойство выполнено для представления e_2. Тогда свойства $e_1(Q) \in P$ и $e_2(Q) \in P$ равносильны.

Доказательство. Утверждение симметрично, так что достаточно доказать его в одну сторону. Предположим, что задача $e_1(Q)$ разрешима за время $O(n^k)$ для некоторого фиксированного числа k. По предположению для всякого условия $i \in I$ представление $e_1(i)$ может быть получено из представления $e_2(i)$ за время $O(n^c)$ (где c – некоторая константа, $n = |e_2(i)|$, $|s|$ – обозначает длину строки s). Для решения задачи $e_2(Q)$, получив на входе $e_2(i)$, мы сперва вычислим $e_1(i)$, а затем применим алгоритм, разрешающий $e_1(Q)$ к строке $e_1(i)$. Сколько времени займет наше вычисление? Преобразование $e_2(i)$ в $e_1(i)$ требует полиномиального времени. Следовательно, $|e_1(i)| = O(n^c)$, поскольку длина выхода алгоритма не превосходит времени его работы. Решение задачи с условием $e_1(i)$ занимает $O(|e_1(i)|^k) = O(n^{ck})$ времени. Итак, время вычисления оказалось полиномиальным. (Мы пропустили важный момент: получив на входе некоторую строку, мы должны сначала проверить, что она является e_2–представлением некоторого входа; по предположению это можно сделать за полиномиальное время.)

Представление объекта будем обозначать угловыми скобками: $\langle G \rangle$ – это стандартное представление объекта G. При этом множество всех строк, являющихся представлениями, является полиномиально разрешимым (существует полиномиальный алгоритм, проверяющий по строке, представляет ли она какой-либо объект), а различные «разумные» способы представления данных оказываются полиномиально связанными, так что можно воспользоваться теоремой 38 и не описывать представление детально, если нас интересует лишь вопрос о полиномиальности задачи. Таким образом, в дальнейшем мы не будем делать различия между абстрактной задачей и её строковым представлением, как это обычно и делают.

6.2 Формальные языки

Для задач разрешения удобно использовать терминологию теории формальных языков. **Алфавитом** Σ называется любой конечный набор символов. **Языком** L **над алфавитом** Σ называется произвольное множество строк символов из алфавита Σ (такие строки называются **словами** в алфавите Σ). Например, можно рассмотреть $\Sigma = \{0,1\}$ и язык $L = \{10, 11, 101, 111, 1011, 1101, 10001,\ldots\}$, состоящий из двоичных записей простых чисел.

Задача разрешения (точнее, соответствующая ей строковая задача разрешения) является языком над алфавитом $\Sigma = \{0,1\}$.

Например, задаче PATH соответствует язык PATH = {<G, u, v, k>: $G = (V,E)$ – неориентированный граф, $u, v \in V$, $k \geq 0$ – целое число, и в графе G существует путь из u в v, длина которого не превосходит k}.

Мы будем использовать одно и то же название – в данном случае PATH – для обозначения задачи и соответствующего языка.

Алгоритм допускает слово

Говорят, что алгоритм A *допускает слово* $x \in \{0, 1\}^*$, если на входе x алгоритм выдает результат 1 ($A(x) = 1$).

Алгоритм отвергает слово

Говорят, что алгоритм A *отвергает слово* $x \in \{0, 1\}^*$, если на входе x алгоритм выдает результат 0 ($A(x) = 0$).

Заметим, что алгоритм может не остановиться на входе x или дать ответ, отличный от 0 и 1. В этом случае он и не допускает и не отвергает слово x.

Алгоритм допускает язык

Говорят, что алгоритм A *допускает язык* L, если алгоритм допускает те и только те слова, которые принадлежат языку L.

Алгоритм A, допускающий некоторый язык L, не обязан отвергать всякое слово $x \notin L$.

Алгоритм распознает язык

Говорят, что алгоритм A *распознает язык* L, если A допускает все слова из L, а все остальные слова отвергает.

Язык допускается за полиномиальное время

Язык L *допускается за полиномиальное время*, если имеется алгоритм A, который допускает данный язык, причем всякое слово $x \in L$ допускается алгоритмом за время $O(n^k)$, где n – длина слова x, а k – некоторое не зависящее от x число.

Язык распознается за полиномиальное время

Язык L *распознается за полиномиальное время*, если имеется алгоритм A, который распознает данный язык, причем время работы алгоритма на каждом слове длины n не больше $O(n^k)$.

Теперь можно переформулировать определение сложностного класса ***P***.

$P = \{L \subset \{0, 1\}^*$: существует алгоритм A, распознающий язык L за полиномиальное время$\}$.

На самом деле в данной ситуации нет разницы между языками, допускаемыми и распознаваемыми за полиномиальное время.

Теорема 39. $P = \{L$: L допускается за полиномиальное время$\}$.

Доказательство. Если язык распознается некоторым алгоритмом, то он и допускается тем же алгоритмом. Остается доказать, что если язык L допускается полиномиальным алгоритмом A, то он распознается некоторым (возможно, другим) полиномиальным алгоритмом B. Пусть алгоритм A допускает язык L за время $O(n^k)$. Это значит, что существует константа c, для которой A допускает любое слово длины n из L, сделав не более $T = cn^k$ шагов. (Формально говоря, это верно для достаточно длинных слов

x; мы опускаем очевидные детали.) С другой стороны, слова не из L алгоритм не допускает (ни за какое время).

Новый алгоритм B моделирует работу алгоритма A и считает число шагов этого алгоритма, сравнивая его с известной границей T. Если за время T алгоритм A допускает слово x, алгоритм B также допускает это слово и выдает 1. Если же A не допускает x за указанное время, алгоритм B прекращает моделирование и отвергает слово (выдает 0). Замедление работы за счёт моделирования и подсчета шагов не так уж велико и оставляет время работы полиномиальным.

6.3 Проверка принадлежности языку и класс NP

Для определения сложностного класса ***NP*** нам потребуется ряд новых понятий, в том числе рассмотрение задачи о гамильтоновом цикле.

Гамильтоновым циклом в неориентированном графе $G = (V, E)$ называется простой цикл, который проходит через все вершины графа. Графы, в которых есть гамильтонов цикл, называются **гамильтоновыми**.

Задача о гамильтоновом цикле состоит в выяснении, имеет ли данный граф G гамильтонов цикл. Формально говоря,

$$\text{HAM-CYCLE} = \{<G>: G \text{ – гамильтонов граф}\}.$$

Как решать такую задачу? Можно перебрать все перестановки вершин данного графа и проверить, является ли хотя бы одна из них гамильтоновым циклом. Оценим время работы такого алгоритма. Если мы используем представление графа с помощью матрицы инцидентности, то число вершин m в графе будет $\Omega(\sqrt{n})$, где $n = |<G>|$ – длина представления графа G. Имеется $m!$ различных перестановок вершин графа, и время работы алгоритма равно $\Omega(m!) = \Omega(\sqrt{n}\,!) = \Omega(2^{\sqrt{n}})$, т.е. не является полиномиальным. Таким образом, наивный алгоритм не дает эффективного решения задачи. (На самом деле, есть основания предполагать, что полиномиального алгоритма для неё вообще не существует.)

Пусть вы заключили пари с приятелем, который утверждает, что (нарисованный перед вами на доске) граф является гамильтоновым. При этом вы не можете быстро проверить так это или нет. Тем не менее приятель может выиграть пари, если каким-то образом отгадает гамильтонов цикл и предъявит его вам: проверка того, что данный цикл является гамильтоновым, проста. Нужно лишь проверить, что предъявленный цикл проходит через все вершины графа и что он действительно идет по рёбрам.

Проверяющий алгоритм

- Назовем проверяющим алгоритмом алгоритм A с двумя аргументами; первый аргумент мы будем называть (как и раньше) входной строкой, а второй – **сертификатом**. Мы говорим, что алгоритм A с двумя аргументами **допускает** вход x, если существует сертификат y, для которого $A(x, y) = 1$.
- **Языком, проверяемым алгоритмом** A, мы назовем язык

$L = \{x \in \{0, 1\}^*: \exists y \in \{0, 1\}^*$, для которого $A(x, y) = 1\}$.

Другими словами, алгоритм A проверяет язык L, если для любой строки $x \in L$ найдется сертификат y, с помощью которого A может проверить принадлежность x к языку L, а для $x \notin L$ такого сертификата нет. Например, в задаче HAM-CYCLE сертификатом была последовательность вершин, образующая гамильтонов цикл.

Класс *NP*

Сложностный класс ***NP*** – это класс языков, для которых существуют проверяющие алгоритмы, работающие полиномиальное время, причем длина сертификата также ограничена некоторым полиномом.
Более точно, язык L принадлежит классу ***NP***, если существует такой полиномиальный алгоритм A с двумя аргументами и такой многочлен $p(x)$ с целыми коэффициентами, что

$L = \{x \in \{0, 1\}^*$: существует сертификат y, для которого $|y| \le p(|x|)$ и $A(x, y) = 1\}$.

В этом случае мы говорим, что алгоритм A **проверяет язык L за полиномиальное время**.

Несколько слов о названии: сокращение ***NP*** происходит от английских слов Nondeterministic Polynomial (time), что переводится как недетерминированное полиномиальное время. Первоначально класс ***NP*** определялся в терминах так называемых недетерминированных вычислений.

Мы уже знаем одну задачу из класса ***NP*** – это задача HAM-CYCLE. Кроме того, всякая задача из ***P*** принадлежит также и ***NP***. Действительно, если есть полиномиальный алгоритм, распознающий язык, то легко построить проверяющий алгоритм для того же языка – проверяющий алгоритм может просто игнорировать свой второй аргумент (сертификат). Таким образом, $\boldsymbol{P} \subseteq \boldsymbol{NP}$.

В данное время неизвестно, совпадают ли классы ***P*** и ***NP***, но большинство специалистов полагают, что нет.

Интуитивно класс ***P*** можно представлять себе как класс задач, которые можно быстро решить, а класс ***NP*** – как класс задач, решение которых может быть быстро проверено.

На практике решить самому задачу часто намного труднее, чем проверить уже готовое решение, особенно если время работы ограничено. По аналогии можно думать, что в классе ***NP*** имеются задачи, которые нельзя решить за полиномиальное время.

6.4 NP-полнота и сводимость

По-видимому, наиболее убедительным аргументом в пользу того, что классы ***P*** и ***NP*** различны, является существование так называемых ***NP***-полных задач. Этот класс обладает удивительным свойством: если какая-нибудь ***NP***-полная задача разрешима за полиномиальное время, то и все задачи из класса ***NP*** разрешимы за полиномиальное время, т.е. $\boldsymbol{P} = \boldsymbol{NP}$. Несмотря на многолетние исследования, ни для одной ***NP***-полной задачи не найден полиномиальный разрешающий алгоритм.

В частности, задача HAM-CYCLE (о гамильтоновом цикле) является ***NP***-полной. Таким образом, научившись решать её за полиномиальное время, мы получим полиномиальные алгоритмы для всех задач класса ***NP***. Переформулировка: если множество $\boldsymbol{NP} \backslash \boldsymbol{P}$ не пусто, то HAM-CYCLE $\in \boldsymbol{NP} \backslash \boldsymbol{P}$.

Неформально говоря, ***NP***-полные языки являются самыми «трудными» в классе ***NP***. При этом трудность языков можно сравнивать, сводя один язык к другому. В этом разделе мы даём определение сводимости, затем определяем класс ***NP***-полных задач и устанавливаем полноту многих задач.

Говоря неформально, задача Q сводится к задаче R, если задачу Q можно решить для любого входа, считая известным решение задачи R для какого-то другого входа. Например, задача решения линейного уравнения сводится к задаче решения квадратного уравнения (линейное уравнение можно превратить в квадратное, добавив фиктивный старший член). Если задача Q сводится к задаче R, то любой алгоритм, решающий R, можно использовать для решения задачи Q, т.е. Q «не труднее» R.

Сводимость за полиномиальное время

Говорят, что язык L_1 сводится за полиномиальное время к языку L_2 (запись: $L_1 \leq_P L_2$), если существует такая функция f: $\{0, 1\}^* \to \{0, 1\}^*$, вычислимая за полиномиальное время, что для любого $x \in \{0, 1\}^*$,

$x \in L_1 \Leftrightarrow f(x) \in L_2$.

Функцию f называют **сводящей функцие**й, а полиномиальный алгоритм F, вычисляющий функцию f, – **сводящим алгоритмом**.

Теорема 40. Если язык $L_1 \subseteq \{0, 1\}^*$ сводится за полиномиальное время к языку $L_2 \subseteq \{0, 1\}^*$ и $L_2 \in \boldsymbol{P}$, то $L_1 \in \boldsymbol{P}$.

Доказательство. Пусть A_2 – полиномиальный алгоритм, распознающий язык L_2, а F – полиномиальный алгоритм, сводящий язык L_1 к языку L_2. Построим алгоритм A_1, который будет за полиномиальное время разрешать язык L_1.

Рис. 2 иллюстрирует построение. Получив вход $x \in \{0, 1\}^*$, алгоритм A_1 (с помощью алгоритма F) получает $f(x)$ и с помощью алгоритма A_2 проверяет, принадлежит ли слово $f(x)$ языку L_2. Результат работы алгоритма A_2 на слове $f(x)$ и выдается алгоритмом A_1 в качестве ответа.

Определение полиномиальной сводимости гарантирует, что алгоритм A_1 дает правильный ответ; он полиномиален, поскольку полиномиальны алгоритмы F и A_1.

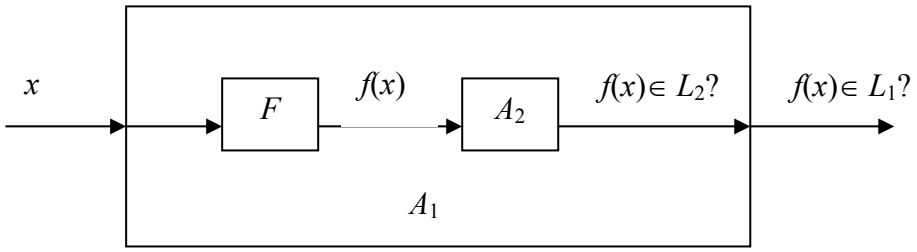


Рис. 2. Принадлежность слова x к языку L_1 можно проверить, использовав сводящий алгоритм F и алгоритм A_2, разрешающий язык L_2.

Понятие сводимости позволяет придать точный смысл утверждению о том, что один язык не менее труден, чем другой (с точностью до полинома). Запись $L_1 \leq_P L_2$ можно интерпретировать так: сложность языка L_1 не более чем полиномиально превосходит сложность языка L_2. Наиболее трудны в этом смысле ***NP*-полные задачи**.

***NP*-полнота**

Язык $L \subseteq \{0, 1\}^*$ называется ***NP*-полным**, если

1) $L \in \boldsymbol{NP}$;
2) $L' \leq_P L$ для любого $L' \in \boldsymbol{NP}$.

Класс ***NP*-полных языков будем обозначать *NPC***.

Основное свойство ***NP*-полных языков состоит в следующем.

Теорема 41. Если некоторая ***NP*-полная задача разрешима за полиномиальное время, то $\boldsymbol{P} = \boldsymbol{NP}$. Другими словами, если в классе $\boldsymbol{NP}$ существует задача, не разрешимая за полиномиальное время, то все ***NP*-полные задачи таковы.

Доказательство. Пусть L – ***NP*-полный язык, который одновременно оказался разрешимым за полиномиальное время ($L \in \boldsymbol{P}$ и $L \in \boldsymbol{NPC}$). Тогда для любого языка L' по свойству 2 определения ***NP*-полного языка имеем $L' \leq_P L$. Следовательно, $L' \in \boldsymbol{P}$ (теорема 40), и первое утверждение теоремы доказано.

Второе утверждение теоремы является переформулировкой первого.

Таким образом, гипотеза ***P*** ≠ ***NP*** означает, что ***NP***-полные задачи не могут быть решены за полиномиальное время. Большинство экспертов полагают, что это действительно так; предполагаемое соотношение между классами ***P***, ***NP*** и ***NPC*** показано на рис. 3.

Конечно, мы не можем быть уверены, что однажды кто-нибудь не предъявит полиномиальный алгоритм для решения ***NP***-полной задачи и не докажет тем самым, что ***P***=***NP***. Но пока этого никому не удалось, и поэтому доказательство ***NP***-полноты некоторой задачи является убедительным аргументом в пользу того, что она является практически неразрешимой.

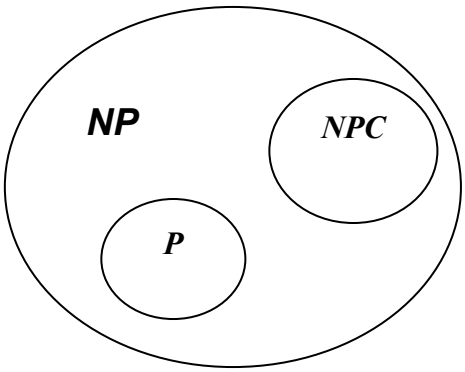


Рис. 3. Предполагаемое соотношение между классами ***P***, ***NP*** и ***NPC***.
Классы ***P*** и ***NPC*** содержатся в ***NP*** (что очевидно), и, можно полагать, не пересекаются и не покрывают всего ***NP***.

Первой задачей, для которой была доказана её ***NP***-полнота была задача о выполнимости пропозициональных формул:

SAT = {<φ>: φ – пропозициональная и выполнимая формула}.

В неформальной постановке:

Условие. Дана формула исчисления высказываний φ.

Вопрос. Существует ли такое распределение истинностных значений высказывательных переменных, при которых формула φ выполнима?

Теорема 42(теорема Кука). Задача SAT является ***NP***-полной [10, с.56–63].

Можно сказать так: после доказательства этой теоремы была пробита брешь в стене и установлена ***NP***-полнота одной задачи, после этого было доказано с помощью сводимости ***NP***-полнота большого числа задач [10, 13]. В разделе 6.3 мы перечислили некоторые задачи, которые не попадают ни в класс ***P***, ни в класс ***E***. Все они являются ***NP***-полными.

Забудь, что прочел.
Учись читать заново.
Так сказал сэнсэй.
Б. Акунин
Алмазная колесница

Автор пишет только половину книги:
другую половину пишет читатель.
Джозеф Конрад

Литература

1. Curry H. B., Feys R. Combinatory Logic, vol. I, Amsterdam: North Holland Co., 1958.
2. Church A. The Calculi of Lambda Conversion. Princeton University Press, Princeton, 1941.
3. Goodstein, R. L. On the Restricted Ordinal Theorem. // J. Symb. Logic. 1944. V. 9.– P. 33–41.
4. Kirbi L., Paris J. Accessible independence result for Peano arithmetic. // Bulletin of the London Mathematical Society. 1982. – V. 14. – P. 285–293.
5. Schönfincel M. Über die Bausteine der mathematischen Logik. Math. Annalen, 92, 1924, s. 305–316.
6. Turner D. A. A new implementstion technique for applicative languages. Software–Practice and Experience, 9, 1979, pp. 31-49.
7. Акритас А. Основы компьютерной алгебры с приложениями: пер. с англ. – М.: Мир, 1994. – 544 с.
8. Ахо А., Хопкрофт Дж., Ульман Дж. Построение и анализ вычислительных алгоритмов. М.: Мир, 1979. – 536с.
9. Барендрегт Х. Ламбда–исчисление. Его синтаксис и семантика. – М.: Мир, 1985.–606с.
10. Гэри М., Джонсон Д. Вычислительные машины и труднорешаемые задачи. – М.: Мир, 1982.
11. Катленд Н. Вычислимость. Введение в теорию рекурсивных функций: Пер. с англ.–М.: Мир, 1983.–256с.
12. Кац М., Улам С. Математика и логика. Ретроспектива и перспективы: Пер. с англ. – М.: Мир, 1971. –254с.
13. Кормен Т., Лейзерсон Ч., Ривест Р. Алгоритмы: построение и анализ. М.: МЦНМО, 2001.– 960с.
14. Лорьер Ж.-Л. Системы искусственного интеллекта – М.: Мир, 1991. – 568с.
15. Манин Ю. И. Вычислимое и невычислимое. – М.: «Советское радио», 1980. – 128с.
16. Манин Ю. И. Доказуемое и недоказуемое. – М.: «Советское радио», 1979. – 168с.
17. Мендельсон Э. Введение в математическую логику - М.: Наука, 1976.– 320с.
18. Н. К. Верещагин, А. Шень. Лекции по математической логике и теории алгоритмов. Часть 1. Начала теории множеств. М.: МЦМНО, 1999, 128 стр.
19. Оуэн Н. Магические метафоры. 77 историй для учителей, терапевтов и думающих людей / пер. с англ. Е. Рачковой. – М.: Изд-во Эксмо, 2002. – 320 с.
20. Пенроуз Р. Тени разума: в поисках науки о сознании. Часть 1: Понимание разума и новая физика. – Москва-Ижевск: Институт компьютерных исследований, 2003, 368с.
21. Подниекс К. М. Вокруг теоремы Гёделя. Рига: «Зинатне», 1992. – 191с.
22. Рейнгольд Э., Нивергельт Ю., Део Н. Комбинаторные алгоритмы. Теория и практика. Пер. с англ. – М.: Мир, 1980. – 478 с.
23. Смаллиан Р. Принцесса или тигр? – М.: Мир, 1985. – 221с.
24. Справочная книга по математической логике: В 4-х частях /Под ред. Дж. Барвайса. – Ч. III. Теория рекурсии: Пер. с англ.– М.: Наука, 1982.– 360с.
25. Справочная книга по математической логике: В 4-х частях /Под ред. Дж. Барвайса. – Ч. IV. Теория доказательств и конструктивная математика: Пер. с англ.– М.: Наука, 1983. – 392с.
26. Филд А., Харрисон П. Функциональное программирование. – М.: Мир, 1993. – 637 с.
27. Хофштадтер Д. Гёдель, Эшер, Бах: эта бесконечная гирлянда. – Самара: Издательский Дом «Бахрах-М», 2001. – 752 с.
28. Хьювенен Э., Сеппянен Й. Мир Лиспа. В 2-х т.–М.:Мир,1990.