ФГБОУ ВПО «Московский государственный
университет имени М. В. Ломоносова»
Механико-математический факультет

На правах рукописи

Андреев Михаил Александрович

Сравнение игровыми методами понятий префиксной и обычной колмогоровской сложности

01.01.06 — математическая логика, алгебра и теория чисел

ДИССЕРТАЦИЯ
на соискание ученой степени
кандидата физико-математических наук

Научный руководитель:
доктор физико-математических наук,
профессор Николай Константинович Верещагин

Москва — 2016

Оглавление

1 Введение **4**

1.1 Актуальность темы 4

1.2 Цели и задачи работы 11

1.3 Основные результаты 11

1.4 Апробация работы и публикации автора 12

1.5 Используемые обозначения 13

2 Основные понятия и результаты **15**

2.1 Колмогоровская сложность 15

2.2 Перечислимые снизу дискретные и непрерывные полумеры 16

2.3 Очень большие числа 18

3 Области определения декомпрессоров **20**

3.1 Введение 20

3.2 Определения и обозначения 22

3.3 Основная лемма 23

3.4 Плотность 23

3.5 Разбиваем задачу на несколько 24

3.6 Построение полумер игровым методом 24

3.7 Выигрышная позиция 26

3.8 Выигрышная стратегия Алисы 27

3.9 Почему стратегия работает 27

4 Сравнение KA и KP **29**
4.1 Введение 29
4.2 Основной результат и набросок доказательства: игровая техника 33
4.3 Сведение к конечным играм 35
4.4 Выигрышная стратегия в конечной игре 37
4.5 Усиление: перечислимое множество, характеристическая последовательность которого имеет бесконечную сумму 42

5 Очень большие числа **46**
5.1 Введение 46
5.2 Верхние оценки 51
5.3 Нижние оценки 53
5.3.1 Доказательство утверждения (i) 53
5.3.2 Доказательство утверждения (ii) 56

Глава 1

Введение

1.1. Актуальность темы

Алгоритмическая теория информации изучает сложность и априорную вероятность конечных объектов. Обычно в качестве таких объектов рассматриваются конечные слова из нулей и единиц или натуральные числа. Кроме сложности конечных объектов, изучается случайность индивидуальных бесконечных объектов. В качестве бесконечных объектов обычно рассматриваются бесконечные последовательности битов. Начнем с обсуждения конечных объектов, а потом перейдем к бесконечным.

Колмогоровская сложность в одном из вариантов (другие варианты сложности будут обсуждаться дальше) была введена Колмогоровым в 1960-е годы [5]. Неформально говоря, колмогоровская сложность слова — это длина кратчайшего описания этого слова. Очевидно, что это определение зависит от того, что мы понимаем под «описанием». Колмогоров предложил рассматривать вычислимые функции (не обязательно всюду определенные) в качестве способов описания. Скажем, что слово p является *описанием* слова x относительно способа описания U, если $U(p) = x$. *Сложность слова x относительно способа описания U* равна длине кратчайшего U-прообраза слова x и обозначается $KS_U(x)$. Колмогоров доказал, что в этом классе способов описания существует *оптимальный*. Другими словами, существует такая вычислимая функция U, что для любой вычислимой функции U' неравенство

$$KS_U(x) < KS_{U'}(x) + c$$

выполняется для любого слова x и некоторой константы c, зависящей только от U' (но не от слова x). Вычислимые функции, удовлетворяющие этому свойству, называют *оптимальными способами описания* или *оптимальными*

декомпрессорами.

Зафиксируем какой-либо оптимальный способ описания. Сложность относительно этого способа описания называется *колмогоровской сложностью.* Мы будем обозначать её $KS(x)$. При замене оптимального способа описания функция сложности изменится не больше чем на константу, так что понятие колмогоровской сложности по существу определено с точностью до ограниченного слагаемого.

У колмогоровской сложности есть много разных вариантов, которые удобно использовать в разных ситуациях. Одним из самых часто используемых вариантов является префиксная колмогоровская сложность, введенная Левиным и позже Чейтиным [2, 6, 7, 12]. Для её определения будем рассматривать не все возможные вычислимые способы описания, а только беспрефиксные. Способ описания называется *беспрефиксным*, если его область определения не содержит двух слов, одно из которых является началом другого. Сложность относительно способа описания вводится так же, как и в случае обычной колмогоровской сложности. Как и раньше, среди беспрефиксных способов описания существует оптимальный (в том же смысле, что и в обычном случае, но для класса беспрефиксных способов описания). Сложность относительно него называется *префиксной сложностью* слова и обозначается KP. Как и обычная колмогоровская сложность $KS(x)$, префиксная сложность $KP(x)$ определена с точностью до ограниченного слагаемого.

Первый результат работы связан со сравнением префиксной и обычной колмогоровской сложности (точнее говоря, он сравнивает области определения оптимальных префиксных и обычных декомпрессоров).

Префиксная сложность тесно связана с так называемой максимальной дискретной полумерой (её называют также дискретной априорной вероятностью). Одно из определений этого понятия использует вероятностные алгоритмы. Такой алгоритм в качестве одной из базовых операций может получить случайный бит (бросить честную монету). Мы считаем, что алгоритм не имеет входа, а на выходе печатает какое-то двоичное слово и останавливается (или не печатает ничего). Если фиксирован такой вероятностный алгоритм, то возникает распределение на словах (для каждого слова есть число — вероятность получить данное слово на выходе). Формально говоря, это распределение не является вероятностным распределением на словах, поскольку алгоритм может не остановиться. Поэтому надо добавить специальный объект «неопределённость», или (как это обычно делают) рассматривать не

меры на словах, а *полумеры*, неотрицательные вещественные функции с суммой не больше 1.

Распределения на словах (полумеры), появляющиеся на выходе вероятностных алгоритмов, можно охарактеризовать по-другому. Напомним, что функция f называется *перечислимой снизу*, если она является поточечным пределом монотонно возрастающей (по второму аргументу) всюду определённой вычислимой функции двух аргументов F с рациональными значениями: $F(n,i) \to f(n)$ при $i \to \infty$. Оказывается, что вероятностные алгоритмы задают перечислимые снизу полумеры, и для любой перечислимой снизу полумеры есть алгоритм, который задает эту полумеру. Левин показал, что среди перечислимых снизу полумер существует максимальная с точностью до мультипликативной константы. Она называется *дискретной априорной вероятностью*, и мы будем обозначать её значение на слове x через $\mathbf{m}(x)$.

Левин показал, что она тесно связана с префиксной сложностью, а именно, верна следующая формула:

$$\mathbf{m}(x) = 2^{-KP(x)+O(1)},$$

где $O(1)$ обозначает некоторую ограниченную по модулю функцию от x.

Помимо дискретной априорной вероятности можно рассматривать так называемую *непрерывную априорную вероятность*. Чтобы определить её, вместо вероятностных алгоритмов, которые печатают слово на выходе и останавливаются, будем рассматривать вероятностные алгоритмы, которые печатают на выходе бит за битом и никогда не останавливаются (хотя, возможно, печатают только конечное число битов при некоторых исходах бросаний). Для каждого такого алгоритма рассмотрим функцию на словах. Её значение на слове x есть вероятность события «на выходе в какой-то момент появляется слово x» (а после этого могут появиться и другие биты — или не появиться). Каков бы ни был вероятностный алгоритм описанного вида, возникающая функция A обладает двумя свойствами:

1. $A(\Lambda) = 1$, где Λ — пустое слово;
2. $A(x) \geq A(x0) + A(x1)$ для любого слова x.

Все неотрицательные вещественнозначные функции, обладающие этими двумя свойствами, называют *непрерывными полумерами*. Как и в случае дискретных полумер, вероятностные алгоритмы порождают все перечислимые снизу непрерывные полумеры и только их. Звонкин и Левин в работе 1970

года [4] доказали, что существует максимальная перечислимая снизу непрерывная полумера. Мы будем обозначать ее через **a** и называть *непрерывной априорной вероятностью*. Минус двоичный логарифм этой полумеры называют *априорной сложностью* и обозначают KA.

Оба варианта априорной вероятности (непрерывный и дискретный) могут быть использованы для того, чтобы дать критерий случайности бесконечных последовательностей в смысле Мартин-Лёфа. Прежде чем формулировать соответствующие результаты, скажем несколько слов об этом понятии.

Определяя случайность, мы пытаемся выделить класс объектов, про которые мы готовы поверить, что они получены в результате случайного процесса. Например, странно считать случайным длинное слово из всех нулей или слово 0101 ... 01 (повторяющуюся группу цифр 01): если кто-то утверждает, что он многократно бросал монету и получил такую последовательность орлов и решек, мы вряд ли ему поверим. В случае конечных слов естественно объявить «случайными» те слова, у которых сложность близка к длине (несжимаемые слова), и «неслучайными» все остальные (те слова, которые хорошо сжимаются). Понятно, что в случае конечных объектов невозможно провести чёткую границу между случайными и неслучайными словами — как её ни проводи, обязательно найдутся два слова, отличающиеся в одном символе, одно из которых считается случайным, а второе нет, что противоречит интуиции. В случае бесконечных последовательностей этой проблемы нет, и соответствующее определение случайности (наиболее распространённое) было дано Мартин-Лёфом в 1960-е годы.

Определяя случайный бесконечный объект, надо прежде всего фиксировать вероятностное пространство. Мы рассматриваем канторовское пространство (пространство всех бесконечных последовательностей из нулей и единиц) с естественной σ-алгеброй: базовыми множествами являются интервалы — множества $x\Omega$ бесконечных последовательностей с данным началом x. Мера на канторовском пространстве называется *вычислимой*, если есть алгоритм, вычисляющий рациональное приближение меры любого базового интервала с любой наперед заданной точностью. Мартин-Лёф [16] ввел понятие эффективно нулевого множества относительно данной вычислимой меры. А именно, *эффективно нулевым* называется множество, для которого существует алгоритм, который по любому рациональному $\varepsilon > 0$ перечисляет покрытие этого множества базовыми интервалами с суммарной мерой не

больше ε. Мартин-Лёф показал, что для любой вычислимой меры существует максимальное (по включению) эффективно нулевое множество (естественно, зависящее от меры). Последовательности, не входящие в него, называются *случайными* относительно этой меры.

Шнорр, Левин и Гач [8, 20] в 1970-е годы обнаружили, что у случайных последовательностей есть естественная характеризация в терминах сложности. А именно, последовательность ω случайна относительно вычислимой меры μ тогда и только тогда, когда отношение максимальной полумеры и μ-меры соответствующего интервала ограничено на всех её началах:

$$\frac{\mathbf{m}(x)}{\mu(x\Omega)} < c$$

для всех начал x последовательности ω и некоторой константы c.

Было показано также [14], что в этом определении $\limsup$ можно заменить на сумму: последовательность ω случайна относительно меры μ тогда и только тогда, когда

$$\sum_{x \sqsubset \omega} \frac{\mathbf{m}(x)}{\mu(x\Omega)} < \infty, \qquad (*)$$

где $x \sqsubset \omega$ обозначает, что слово x является (конечным) началом последовательности ω. Оба выражения (с максимумом и суммой), точнее, их логарифмы, обычно называют «дефектом случайности» последовательности ω. Можно показать, что для вычислимых мер μ эти выражения отличаются на ограниченный множитель (см. раздел 2.4 обзора [1]) Таким образом, для вычислимой меры μ последовательность случайна тогда и только тогда, когда ее дефект случайности ограничен сверху.

Понятие случайности многократно пытались расширить на более широкий класс распределений, чем класс вычислимых мер. Например, Левин [9] ввел определение *равномерной случайности* относительно произвольной (не обязательно вычислимой) меры P. Левин доказал, что существует «нейтральная» *мера* N, для которой любая бесконечная последовательность равномерно случайна относительно N.

Было бы интересно обобщить понятие случайности на случай непрерывных перечислимых снизу полумер. Неформально говоря, мы пытаемся ответить на вопрос «про какие бесконечные последовательности мы готовы поверить, что они появились на выходе данного вероятностного алгоритма (который печатает ответ символ за символом)?». Их можно было бы назвать слу-

чайными относительно соответствующей этому алгоритму непрерывной полумеры. Кажется естественным объявить, что последовательность случайна относительно полумеры, если она является образом случайной по равномерной мере последовательности относительно алгоритма, порождающего эту полумеру. К сожалению (см. раздел 5.9.5 в английской версии книги Успенского и др. [22]), это определение некорректно в том смысле, что существуют два алгоритма, порождающих одно и то же распределение, но для которых образы множества случайных последовательностей отличаются.

Возникает естественная идея использовать формулу (∗) или аналогичную формулу с максимумом вместо суммы для измерения неслучайности последовательности ω относительно произвольной перечислимой снизу полумеры μ. Однако не ясно, какую из этих двух формул выбрать. С одной стороны, формула (∗) представляется более правильной по следующей причине. Для случая вычислимых мер эта формула дает наиболее естественное выражение для максимальной с точностью до ограниченного множителя перечислимой снизу функции $t(\omega)$, интеграл которой по мере μ конечен. Такие функции называются тестами случайности и служат для измерения неслучайности; бесконечность значения такой функции на данной последовательности принято считать свидетельством «неслучайности» последней. С другой стороны, посмотрим, что дают обе формулы в случае максимальной непрерывной полумеры **a**. Формула с максимумом дает ограниченное значение на любой последовательности ω, поскольку максимальная дискретная полумера на любом слове не превосходит максимальной непрерывной полумеры (с точностью до ограниченного множителя). Таким образом, если использовать формулу с макисумом, то любая последовательность оказывается случайной (относительно максимальной непрерывной полумеры). Пожалуй, это согласуется с нашей интуицией: максимальная непрерывная полумера устроена столь сложно, что у нас нет никаких средств установить неслучайность какой-либо последовательности относительно нее.

В этой связи возникают следующие вопросы:

- отличаются ли значения, даваемые этими формулами, на ограниченный множитель;
- если это не так, то задают ли они один и тот же класс случайных последовательностей;
- наконец, если и это не так, то верно ли, что любая последовательность

имеет ограниченный дефект случайности в смысле (∗) относительно максимальной непрерывной полумеры **a**?

Вторым результатом диссертации являются отрицательные ответы на все эти вопросы. А именно, в главе 4 доказано, что существует последовательность ω, для которой формула (∗) даёт бесконечное значение в случае максимальной непрерывной полумеры **a**. Таким образом, формулы с суммой и максимумом могут давать существенно разные значения и задают разные классы случайных последовательностей в случае максимальной непрерывной полумеры.

Еще одним приложением теории Колмогоровской сложности является классификация очень больших чисел. К этой области относится третий основной результат диссертации. Прежде чем перейти к деталям, отметим, что между двоичными словами и натуральными числами есть вычислимая биекция, поэтому можно говорить о сложности и априорной (дискретной) вероятности натуральных чисел.

В 1960-е годы Радо [19] предложил рассматривать наибольшее число единиц, которое может напечатать перед остановкой машина Тьюринга с алфавитом $\{0,1\}$ и не более чем n внутренними состояниями, начиная с пустой ленты. Он показал, что получившаяся функция растет быстрее любой вычислимой. У этого определения есть много вариаций. Например, можно рассматривать одностороннюю ленту или двустороннюю, одну или несколько лент и т.п. Теория сложности позволяет дать более инвариантное определение, рассматривая наибольшее число сложности не больше n. Другими словами, для каждого n можно рассмотреть самое большое число, которое может напечатать программа длины не больше n. Этот вариант построения быстро растущих функций рассматривал Гач [15]. Конечно, это определение зависит от того, какой именно оптимальный способ описания мы будем использовать. Поэтому, желая получить инвариантные результаты, мы должны записывать все соотношения (неравенства) с точностью до изменения аргумента быстро растущих функций на константу.

Остается вопрос, какой именно вариант сложности использовать. П. Гач исследовал функцию, которая получается, если использовать префиксную сложность. Мы будем обозначать эту функцию $BP(n)$. Второй вариант быстрорастущей функции, который тоже рассматривал Гач — это регулятор сходимости ряда $\sum_x \mathbf{m}(x)$. Имеется в виду, что для каждого n мы рассматриваем

первое такое N, что

$$\sum_{k>N} \mathbf{m}(k) < 2^{-n}.$$

Будем обозначать это число $BP'(n)$. Ясна связь его с величиной $BP(n)$, которую можно определить как первое такое N, что $\mathbf{m}(k) < 2^{-n}$ для всех $k > N$. Из этого описания сразу видно, что $BP(n) \leq BP'(n)$. Гач показал, что функции BP и BP' не слишком далеки друг от друга (отличаются не больше чем на $KP(n)$ в аргументе: $BP'(n - KP(n) - c) \leq BP(n)$ для некоторой константы c и всех n). Но некоторая (и не слишком маленькая) разница между ними есть: как показал Гач, если вычислимый ряд 2^{-a_n} расходится, то существует такое n, что $BP'(n - a_n) > BP(n)$. Мы добавляем к этой классификации функцию $B(n)$, определённую с использованием обычной (не префиксной) сложности.

1.2. Цели и задачи работы

Цель работы — получить новые результаты в области алгоритмической теории информации с помощью игровой техники. Игровая техника восходит к Ан.А. Мучнику. Работает она следующим образом: для доказательства утверждения про сложность строится комбинаторная игра (не упоминающая сложность явно). Доказывается, что сложностное утверждение является следствием существования выигрышной стратегии для одного из игроков. После этого для этого игрока предъявляется выигрышная стратегия.

1.3. Основные результаты

Работа содержит три результата, два совместных и один полностью принадлежащий автору.

Первый результат (получен совместно с И. Разенштейном и А. Шенем) дает отрицательный ответ на вопрос, поставленный Калюде с соавторами [13]. Доказано существование такого оптимального декомпрессора (для обычной сложности, без требования беспрефиксности), что его область определения не содержит в себе область определения никакого оптимального беспрефиксного декомпрессора.

Второй результат (получен совместно с А. Кумком), показывает, что если использовать формулу (*) на с. 8 в качестве определения случайности относительно непрерывной полумеры, то существует последовательность, ко-

торая не случайна в этом смысле относительно максимальной перечислимой снизу непрерывной полумеры. Другими словами, существует последовательность ω, для которой сумма $\sum_{x \sqsubset \omega} \mathbf{m}(x)/\mathbf{a}(x)$ (по всем конечным началам x последовательности ω) бесконечна. Более того, показано, что последовательность с таким свойством можно выбрать среди характеристических последовательностей перечислимых множеств натуральных чисел.

Третий результат (получен без соавторов) является обобщением результата Гача. Кроме уже описанных функций $BP(n)$ (максимального числа, префиксная колмогоровская сложность которого не больше n) и $BP'(n)$ (регулятора сходимости априорной вероятности), рассматривается функция $B(n)$, равная максимальному числу (обычной) колмогоровской сложности не больше n. Мы доказываем, что все три функции B, BP и BP' достаточно близки: выполнено неравенство

$$BP(n) < BP'(n + O(1)) < B(n + O(1)) < BP(n + KP(n) + O(1)).$$

Однако при этом *оба* разрыва (между BP и BP', как уже было показано Гачем, а также между BP' и B) могут приближаться к даваемой этим неравенством верхней оценке: для любой перечислимой последовательности a_n со свойством $\sum 2^{-a_n} = \infty$ существует такое n, что $BP'(n - a_n) > B(n)$ и существует такое n, что $BP(n - a_n) > BP'(n)$. Заметим, что эти две разницы вместе уже превышают верхнюю оценку, поэтому максимум в том и другом случае необходимо должен достигаться для разных значений n.

В диссертации также предложена более симметричная формулировка этой теоремы (в исходной формулировке не вполне ясно, существенно ли то, что a_n вычитается из аргумента слева, а не прибавляется к аргументу справа): для любой последовательности различных пар натуральных чисел (x_n, y_n), в которой $x_n \leq y_n$, последовательность x_n перечислима снизу, последовательность y_n перечислима сверху и сумма $\sum 2^{x_n - y_n} = +\infty$, найдется такое n, что $B(x_n) > BP'(y_n)$ и найдется такое n (отличное от первого), что $BP'(x_n) > BP(y_n)$.

1.4. Апробация работы и публикации автора

Основные результаты, полученные в работе, были изложены на международных конференциях «Computability in Europe» в 2010 году (Понта Дельгадо,

Португалия) и в 2016 году (Париж, Франция). На последней конференции работа автора [23] получила «Best Student Paper Award».

Работа, посвященная очень большим числам, опубликована в сборнике трудов конференции «Computability in Europe» 2016 года [23] (серия Lecture Notes in Computer Science, входит в систему Scopus). Работа, посвященная случайности относительно полумеры (совместная с А. Кумком), опубликована в журнале Theoretical Computer Science [24] (входит в систему Scopus). Работа, посвященная декомпрессорам (совместная с И. Разенштейном и А. Шенем), опубликована в Theory of Computing Systems [25] (входит в систему Web of Science).

1.5. Используемые обозначения

К сожалению, в литературе по колмогоровской сложности нет устойчивых обозначений. Мы будем в основном следовать обозначениям из [3]. Приведем некоторые принятые нами соглашения:

- Все логарифмы берутся по основанию 2, если явно не указано другое.
- Двоичные слова мы будем обозначать маленькими латинскими буквами. Множество всех двоичных слов длины l мы будем обозначать $\{0,1\}^l$, всех конечных слов — $\{0,1\}^*$. Длину слова x мы будем обозначать $l(x)$.
- Множества слов мы будем обозначать заглавными латинскими буквами.
- Множество всех бесконечных двоичных последовательностей мы будем обозначать Ω. Множество всех бесконечных двоичных последовательностей, начинающихся с фиксированного слова x, мы обозначаем через $x\Omega$. Мы будем отождествлять слова с вершинами бесконечного двоичного дерева, растущего вверх, а бесконечные двоичные последовательности — с бесконечными путями в этом дереве.
- $KS(x)$ — (обычная) колмогоровская сложность слова x.
- $KP(x)$ — префиксная колмогоровская сложность слова x.
- $KA(x)$ — априорная сложность слова x.
- $\mathbf{m}(\cdot)$ — максимальная перечислимая снизу дискретная полумера.

- $\mathbf{a}(\cdot)$ — максимальная перечислимая снизу непрерывная полумера.
- Наконец, мы будем использовать стандартную O-нотацию и писать $f(n) = O(g(n))$, если существует такая константа c, что $f(n) \leq c \cdot g(n)$ для всех достаточно больших n.

Благодарности

Автор выражает глубокую благодарность своим научным руководителям Верещагину Николаю Константиновичу и Шеню Александру Ханьевичу за помощь с постановкой задач, постоянное внимание, поддержку в работе, обсуждения и ценные советы.

Автор благодарен своим соавторам Илье Разенштейну и Акиму Кумку за плодотворное сотрудничество, приведшее к получению и публикации результатов.

Автор благодарен одногруппникам и коллегам, с которыми он обсуждал математические результаты, включая С. Артамонова, Л. Биенвеню (L. Bienvenu), М. Дектярева, К. Салихова.

Наконец, автор очень благодарен семье и друзьям (в частности, П. Вахрушевой, М. Вроде, Н. Горященко, М. Мовшицу) за постоянную поддержку.

Глава 2

Основные понятия и результаты

В этой главе мы дадим все необходимые определения, о которых мы неформально говорили в прошлой главе. Также мы дадим точные формулировки результатов диссертации.

Мы начнем с определения необходимых нам вариантов колмогоровской сложности, после этого перейдем к перечислимым снизу полумерам, и закончим эту главу определениями, связанными с очень большими числами.

2.1. Колмогоровская сложность

В этом разделе мы дадим определения обычной и префиксной колмогоровской сложности и введем понятия оптимальных обычных и беспрефиксных декомпрессоров.

***Определение* 2.1.** *Пусть* $D\colon \{0,1\}^* \to \{0,1\}^*$ *— вычислимая не обязательно тотальная функция. Любую такую функцию мы будем называть* декомпрессором. *Колмогоровская сложность слова* $x \in \{0,1\}^*$ *относительно способа описания* D *определяется как длина кратчайшего* D*-описания:*

$$KS_D(x) := \min_{y:D(y)=x} l(y).$$

***Определение* 2.2.** *Декомпрессор называется* беспрефиксным*, если в его области определения нет двух слов, одно из которых является началом другого (т.е. его область определения является беспрефиксным множеством).*

Оказывается, существуют оптимальные (с точностью до аддитивной константы, зависящей от способа описания, но не от слова) декомпрессоры.

***Теорема* 2.3** ([3]). *Существует* оптимальный декомпрессор U*: для любого декомпрессора* V *существует такая константа* c*, что для всех слов* x *выполнено*

$$KS_U(x) < KS_V(x) + c.$$

Несложно заметить, что множество оптимальных декомпрессоров бесконечно, и сложность любого слова относительно любых двух оптимальных декомпрессоров отличается не больше чем на константу, зависящую от выбранных декомпрессоров, но не от слова. Зафиксируем какой-нибудь оптимальный декомпрессор. Будем называть сложность относительно него *колмогоровской сложностью* и обозначать $KS(\cdot)$.

Верен аналог теоремы 2.3 для множества беспрефиксных декомпрессоров.

***Теорема* 2.4** ([3]). *Существует* оптимальный беспрефиксный декомпрессор U*: для любого беспрефиксного декомпрессора* V *существует такая константа* c*, что для всех слов* x *выполнено*

$$KS_U(x) < KS_V(x) + c.$$

Оптимальных беспрефиксных декомпрессоров бесконечно много, зафиксируем один из них. Будем называть сложность относительно него префиксной колмогоровской сложностью и обозначать $KP(\cdot)$.

Основным результатом главы 3, является следующая теорема:

Теорема (3.1). *Существует такой оптимальный декомпрессор, что его область определения не содержит область определения никакого оптимального беспрефиксного декомпрессора.*

2.2. Перечислимые снизу дискретные и непрерывные полумеры

Префиксная сложность тесно связана с понятием *дискретной априорной вероятности* [7, 6, 12]. Рассмотрим неотрицательную вещественнозначную функцию m, определенную на словах.

***Определение* 2.5**. *Если* $\sum_x m(x) \leq 1$*, то* m *называется* дискретной полумерой.

***Определение* 2.6**. *Вещественнозначная функция* f *называется* перечислимой снизу*, если она представляется в виде предела монотонно возрастающей последовательности* $F(x,0), F(x,1), \ldots$ *где* F *— вычислимая функция двух аргументов с рациональными значениями.*

***Теорема* 2.7** (Левин; см. например [3]). *Среди всех перечислимых снизу дискретных полумер существует максимальная с точностью до мультипликативной константы. Более того, эта мера равна* $2^{-KP(x)+O(1)}$.

Мы зафиксируем максимальную перечислимую снизу полумеру, согласованную с нашим выбором префиксной сложности (т.е. положим её равной $2^{-KP(x)}$ без дополнительного множителя), и будем называть её *дискретной априорной вероятностью* (определение непрерывной априорной вероятности будет дано дальше). Будем обозначать её $\mathbf{m}(x)$.

Её максимальность можно переформулировать так:

***Следствие* 2.8**. *Пусть q — перечислимая снизу полумера. Тогда найдется такая константа c, что для всех слов x выполнено $KP(x) \le -\log_2 q(x) + c$.*

Дискретные перечислимые снизу полумеры могут рассматриваться как распределения на выходах, создаваемые вероятностными машинами, которые печатают свой выход целиком (например, печатают двоичное слово и после этого останавливаются). Также можно рассматривать вероятностные машины, которые печатают выход символ за символом (и никогда не останавливаются, хотя, возможно, печатают лишь конечное слово). Распределения на входах, создаваемое такими машинами, называются перечислимыми снизу *непрерывными полумерами* (т.е. полумерами на двоичных деревьях, подробнее см. [4].) Более формально:

***Определение* 2.9**. *Непрерывной полумерой мы называем неотрицательную вещественнозначную всюду определенную функцию a, определенную на всех двоичных словах, со следующими свойствами:*

- $a(\Lambda) = 1$, *где Λ — пустое слово;*
- $a(x) \ge a(x0) + a(x1)$ *для любого слова x.*

Как и в случае дискретных полумер, существуют максимальные (с точностью до мультипликативного множителя) перечислимые снизу непрерывные полумеры.

***Теорема* 2.10** (Звонкин, Левин, см. [4]). *Среди всех перечислимых снизу непрерывных полумер есть максимальная с точностью до мультипликативной константы.*

Зафиксируем одну из максимальных непрерывных полумер. Будем называть её *непрерывной априорной вероятностью* и обозначать $\mathbf{a}(x)$. Величину $-\log_2 \mathbf{a}(x)$ часто называют *априорной сложностью* слова x и обозначают $KA(x)$.

Основным результатом диссертации в этой области является следующая теорема:

Теорема (4.2, глава 4). *Существует бесконечная последовательность ω из нулей и единиц, такая что*

$$\sum_{x \sqsubset \omega} \frac{\mathbf{m}(x)}{\mathbf{a}(x)} = \infty.$$

Более того, такая последовательность может быть выбрана среди характеристических последовательностей перечислимых множеств.

2.3. Очень большие числа

Введем формально определения, которые мы будем использовать в главе 5. Гач [15] формулирует эти определения в терминах обратных (медленно растущих) функций.

Максимальное число, которое может быть получено на выходе оптимального декомпрессора для входа длины не больше n, мы будем обозначать $B(n)$. Более формально:

Определение **2.11**. $B(n) = \max\{N \mid KS(N) \le n\}$.

Максимальное число, которое может быть получено на выходе оптимального беспрефиксного декомпрессора для входа длины не больше n, мы будем обозначать $BP(n)$. Более формально:

Определение **2.12**. $BP(n) = \max\{N \mid KP(N) \le n\}$.

По теореме 2.7 это определение может быть переписано в терминах априорной дискретной вероятности: $BP(n) = \max\{N \mid \mathbf{m}(N) \ge 2^{-n}\}$ или $BP(n) = \min\{N \mid \max_{k>N} \mathbf{m}(k) < 2^{-n}\}$. Вслед за Гачем рассмотрим определение BP', получающееся заменой максимума в этом определении на сумму:

Определение **2.13**. $BP'(n) = \min\{N \mid \sum_{k>N} \mathbf{m}(k) < 2^{-n}\}$

Гач не рассматривал функцию B, в статье [15] изучалась связь между BP и BP':

***Теорема* 2.14** (Гач, [15]).

(i) *Существует такое c, что $BP(n) \leq BP'(n + c)$ при всех n.*

(ii) *Существует такое c, что $BP(n) \leq BP(n + KP(n) + c)$ при всех n.*

(iii) *Пусть a_n — перечислимая сверху последовательность натуральных чисел и $\sum 2^{-a_n} = +\infty$. Тогда найдется n, при котором $BP'(n - a_n) > BP(n)$.*

Доказательство этой теоремы будет дано в главе 5 диссертации. Там же будет рассмотрена функция B, и доказаны две следующие теоремы:

Теорема (5.1).

(i) *Существует такое c, что $BP(n) \leq BP'(n + c)$ и $BP'(n) \leq B(n + c)$ при всех n.*

(ii) *Существует такое c, что $B(n) \leq BP(n + KP(n) + c)$ при всех n.*

(iii) *Пусть (x_n, y_n) — последовательность пар натуральных чисел, для которых $x_n \leq y_n$, последовательность x_n перечислима снизу, а последовательность y_n перечислима сверху. Пусть при этом $\sum_n 2^{x_n - y_n} < +\infty$. Тогда существует такое c, что $B(x_n) \leq BP(y_n + c)$ при всех n.*

Теорема (5.2). *Пусть (x_n, y_n) — последовательность различных пар натуральных чисел, причем $x_n \leq y_n$, последовательность x_n перечислима снизу, а последовательность y_n перечислима сверху. Пусть при этом $\sum 2^{x_n - y_n} = +\infty$. Тогда*:

(i) *найдется n, при котором $B(x_n) > BP'(y_n)$;*

(ii) *найдется n, при котором $BP'(x_n) > BP(y_n)$.*

Глава 3

Области определения декомпрессоров

3.1. Введение

Напомним необходимые определения.

Пусть $D\colon \{0,1\}^* \to \{0,1\}^*$ — вычислимая (не обязательно тотальная) функция. Колмогоровская сложность слова $x \in \{0,1\}^*$ относительно способа описания D определяется как длина кратчайшего прообраза:

$$KS_D(x) := \min_{y:D(y)=x} l(y).$$

Существует *оптимальный декомпрессор* U такой, что KS_U минимально (с точностью до аддитивной константы). $KS_U(x)$ называется *(обычной) сложностью* слова x и обозначается $KS(x)$.

Декомпрессор называется *беспрефиксным*, если его область определения является беспрефиксным множеством (т.е. если слово u является началом слова v, то декомпрессор не может быть определенным одновременно и на u, и на v). И в этом случае может быть доказано существование оптимального (с точностью до аддитивной константы) беспрефиксного декомпрессора V. Величина $KS_V(x)$ называется *префиксной сложностью* слова x и обозначается $KP(x)$, см. например, [21].

В работе [13] Calude и др. описали области определения оптимальных и оптимальных беспрефиксных декомпрессоров. В ней поставлен вопрос: «Верно ли, что для любого оптимального декомпрессора существует оптимальный беспрефиксный декомпрессор, чья область определения является подмножеством области определения исходного декомпрессора?». Мы даем отрицательный ответ на этот вопрос:

***Теорема* 3.1**. *Существует такой оптимальный декомпрессор D с областью определения S, что никакое множество $T \subset S$ не может являться областью определения оптимального беспрефиксного декомпрессора.*

Заметим, что для любого разрешимого множества A, которое содержит фиксированную долю (скажем, треть) слов длины n для всех n, существует такой оптимальный (обычный) декомпрессор, что его область определение является подмножеством множества A. В самом деле, поскольку множество A содержит треть слов, то существует такое вычислимое инъективное отображение $p \mapsto a(p)$, что для любого слова p слово $a(p)$ двумя символами длиннее чем p и лежит в A. Возьмем любой оптимальный декомпрессор U и заменим k-символьные описания на $(k+2)$-символьные описания, лежащие в A: положим $V(a(p))$ равным $U(p)$ (и не определенным на остальных словах). Тогда V — оптимальный декомпрессор, чья область определения является подмножеством A. (Ответ на более общий вопрос: «Какие множества могут быть областью определения оптимального декомпрессора?» дан в [13].)

Таким образом, достаточно доказать, что существует такое достаточно большое (содержащее хотя бы треть слов для любой длины n) разрешимое множество A, что никакой оптимальный беспрефиксный декомпрессор не имеет областью определения подмножество A. С этого момента мы забываем про обычные декомпрессоры: мы будем строить множество A. Это множество будет построено в следующем разделе. В этом разделе мы обсудим результат и то, на каких идеях построено доказательство.

Полезным инструментом в теории префиксной сложности является наблюдение, называемое *леммой Крафта-Чейтина*. Рассмотрим следующую игру «выделения памяти»: на каждом ходу Алиса выбирает натуральное число n, Боб обязан ответить словом длины n. Числа, выбранные Алисой, не должны быть слишком маленькими: сумма 2^{-n} по всем выбранным ею n не должна превосходить 1. Бобу запрещено выбирать два слова таких, что одно является началом другого. Лемма Крафта-Чейтина утверждает, что у Боба есть вычислимая выигрышная стратегия в этой игре (см. например, [21, с. 28]).

Неформально говоря, вопрос из [13] можно сформулировать так: остается ли это утверждение (про существование выигрышной стратегии) истинным, если некоторые слова (фиксированная часть слов каждой длины) запрещены для Боба (и сумма, которую может использовать Алиса, тоже уменьшена). Ответ на этот вопрос оказывается отрицательным: можно так выбрать запре-

щенные слова, чтобы Боб не мог выиграть при сколь угодно малом ограничении на сумму у Алисы. Строго говоря, нам нужно будет рассмотреть более сложную игру, так как сложность определена с точность до аддитивной константы. Мы не будем описывать эту игру подробно (хотя стоит заметить, что игра, основанная на идеях Андрея Мучника [18] была полезным инструментом, который мы изначально использовали). Вместо рассмотрения этой игры мы дадим доказательство, которое комбинирует рассуждения из теории вычислимости и игровые методы.

Поставленный в [13] вопрос может показаться не очень естественным. Более естественным выглядит вопрос «Верно ли, что каждый оптимальный декомпрессор можно сузить (на некоторое меньшее перечислимое множество) так, чтобы результатом стал беспрефиксный оптимальный декомпрессор?». Однако отрицательный ответ на этот вопрос очевиден: рассмотрим оптимальный декомпрессор U у которого два разных слова s и t имеют ровно по одному описанию p_s и p_t, и, скажем, p_s является началом p_t.

3.2. Определения и обозначения

Напомним, что мы отождествляем двоичные слова с вершинами полного бесконечного двоичного дерева, растущего вверх: пустое слово является корнем, дети слова x отождествлены со словами $x0$ и $x1$. Множество $\Omega = \{0,1\}^\infty$ всех бесконечных двоичных слов отождествляется с $[0,1]$. Для каждого слова x мы определяем интервал $x\Omega \subset [0,1]$; пустому слову соответствует интервал $[0,1]$ целиком; интервалам $x0\Omega$ и $x1\Omega$ соответствуют в точности левая и правая половина $x\Omega$. В Ω интервалу $x\Omega$ соответствует поддерево, состоящее из двоичных слов с началом x. Мы будем использовать обозначение $x\Omega$ и для интервалов в $[0,1]$, и для интервалов в Ω.

В рамках этой главы мы будем называть интервалы $x\Omega$ *базовыми. Базовым подмножеством* Ω мы будем называть конечное объединение базовых интервалов; без ограничения общности можно считать, что эти базовые интервалы не пересекаются и имеют одинаковую длину (т.е. представляют собой поддеревья с корнями на одной высоте). Если базовое множество V равно объединению $\cup_{x\in X} x\Omega$ где $X \subset \{0,1\}^n$, то мы говорим, что X *представляет* V *на высоте* n. Любое базовое множество может быть представлено на любой достаточно большой высоте.

3.3. Основная лемма

Как было показано в прошлом разделе, для доказательства теоремы 3.1 достаточно доказать следующую лемму:

***Лемма* 3.2.** *Существует разрешимое множество A конечных слов, обладающее следующими двумя свойствами:*

(1) *Для всех n в множестве A есть хотя бы треть всех слов длины n;*

(2) *Нет оптимального беспрефиксного декомпрессора, чья область определения является подмножеством A.*

Мы создаем множество A уровень за уровнем так, чтобы каждое базовое множество меры не меньше 1/3 было представлено на некоторой высоте: для каждого базового множества V меры не меньше 1/3 должно существовать такое n, что $A \cap \{0,1\}^n$ представляет V на высоте n. (Это делает A «универсальным»: любое возможное ограничение возникает хотя бы раз.) Мы потребуем дополнительно, чтобы каждое базовое множество V (размера хотя бы 1/3) было представлено на бесконечном количестве уровней, причем каждый раз на большом отрезке последовательных уровней: для любого базового множества V должно существовать бесконечно много таких n, что A представляет V на уровнях $n, n+1, \ldots, 2n$. Легко найти разрешимое множество A с этим свойством (семейство всех базовых множеств перечислимо и может быть эффективно перечислено, таким образом, мы можем выделять бесконечно много таких групп для каждого базового множества.)

Остается показать (предполагая, что A обладает указанным выше свойством), что нет оптимального беспрефиксного декомпрессора с областью определения, являющейся подмножеством A.

3.4. Плотность

Предположим, что D — оптимальный беспрефиксный декомпрессор, чья область определения является подмножеством A. Множество слов x, на которых определен D, является беспрефиксным множеством. Соответствующие интервалы $x\Omega$ не пересекаются; положим $\mathbf{D} \subset \Omega$ равным объединению этих интервалов.

***Лемма* 3.3** (О плотности). *Определенное выше множество $\mathbf{D}$ пересекается с любым базовым множеством меры не меньше* 1/3.

Доказательство. Пусть V — базовое множество с мерой не меньше 1/3. В соответствии с предположением, существует бесконечно много таких n, что V представлено A на высоте $n, n+1, \ldots, 2n$. Если $\mathbf{D}$ не пересекается с V, это означает, что D не определено на всех словах, имеющих длину от n до $2n$, что невозможно для оптимального D (для достаточно больших n большинство слов длины $1.5n$ должны иметь сложность между n и $2n$, а значит описания такой длины должны существовать).

□

3.5. Разбиваем задачу на несколько

Напомним, что следствие 2.8 утверждает, что для любой перечислимой снизу полумеры q существует такая константа c, что $KP(x) < -\log q(x) + c$ для всех слов x. Мы хотим прийти к противоречию с этим следствием.

Пусть D — оптимальный беспрефиксный декомпрессор, чья область определения лежит внутри A. Разность $KS_D(x) - KP(x)$ ограничена константой, не зависящей от x. Мы хотим построить перечислимую снизу полумеру q со следующим свойством: для любого $c > 0$ должно существовать такое x (зависящее от c), что $KS_D(x) \geq -\log_2 q(x) + c$.

Для этого мы построим семейство (для всех целых положительных c) равномерно перечислимых снизу полумер q_c, обладающих следующими двумя свойствами:

- $\sum_x q_c(x) \leq 2^{-c}$,
- $KS_D(x) \geq -\log_2 q(x) + c$ для некоторого x.

Легко видеть что этого достаточно. Функция $q = q_1 + q_2 + \ldots$ является полумерой, перечислима снизу (в силу равномерной перечислимости q_c) и обладает необходимым нам свойством.

3.6. Построение полумер игровым методом

Осталось показать, как построить перечислимую снизу функцию q_c, обладающую требуемыми свойствами, наблюдая за перечислением D. Очевидно, что вычислимые функции являются перечислимыми снизу. Мы предъявим вычислимую функцию (и даже функцию с конечным носителем), обладающую этими двумя свойствами (и сделаем это равномерно по c). Для этого

мы применим игровую технику. Предположим, что у Алисы есть некоторый «капитал» 2^{-c}, и она может распределять его по словам x (заметим, что она распределяет капитал по словам, которые лежат в образе D, а не в области определения); её цель выделить не меньше чем $2^c \cdot 2^{-KS_D(x)}$ для некоторого слова x. Конечно, Алиса не знает окончательного значения $KS_D(x)$; оно может уменьшиться впоследствии (уже после того, как она выделила свой капитал). Поэтому Алисе необходимо гарантировать, что это свойство сохранится для какого-то слова x вне зависимости от того, что случится после того, как она потратила свой капитал. Более того, её стратегия должна быть вычислимой.

Как Алисе выиграть в этой игре? Чтобы объяснить её стратегию, введем несколько вспомогательных определений. Вершины (слова) в A называются *разрешенными*, а слова вне A — *запрещенными*. (На любой высоте не меньше 1/3 всех слов разрешены.)

Эти два определения не зависят от времени (т.е. от того, какая часть области определения D перечислена). Другие понятия являются динамическими. Обозначим через $\bar{D}$ ту часть области определения D, которая уже возникла в процессе перечисления области определения D. Слово u называется *свободным* (в некоторый момент времени), если $\bar{D} \cup \{u\}$ является беспрефиксным множеством. (Слова, которые не являются свободными, не могут возникнуть позже в области определения D, поскольку она должна оставаться беспрефиксной.) В терминах Ω это определение может быть сформулировано следующим образом: u — свободно, если $u\Omega$ и множество $\bar{\mathbf{D}}$ всех последовательностей, чьи начала лежат в $\bar{D}$, не пересекаются.

Если на некотором уровне нет свободных разрешенных слов, это гарантирует, что слов этой длины не возникнет в области определения D.

Свободное слово может стать несвободным, но не наоборот. Заметим, что продолжения свободных слов являются свободными, поэтому доля свободных слов на фиксированном уровне не убывает как функция уровня n (в любой фиксированный момент времени).

Только разрешенные свободные слова могут быть использованы (когда-либо в будущем) как описания, поэтому если на некоторой высоте (точнее, на некотором отрезке высот) такие слова составляют очень маленькую долю от всех слов, Алиса может использовать это, чтобы достичь своей цели. В следующем разделе мы формализуем это утверждение.

3.7. Выигрышная позиция

Предположим, что на всех высотах из некоторого интервала (скажем, между l и L) разрешенные слова представляют одно и то же базовое множество. Дополнительно предположим, что на максимальном уровне L эта доля меньше, чем некоторое небольшое $\varepsilon > 0$.

Что Алиса делает в этом случае? Она может выделить вес $2^c \cdot 2^{-L}$ нескольким (скажем, N; значение N будет выбрано позже) разным словам, у которых еще нет описаний (они не лежат в области значений текущего приближения D). Если этого недостаточно для выигрыша, то каждое из N слов впоследствии получит описание длины не больше L (иначе у Алисы по-прежнему есть слово с требуемым свойством). Эти описания различны (и более того, ни одно не может являться началом другого). Только 2^l описаний могут иметь длину меньше l, поэтому не меньше $N - 2^l$ из них должны лежать в нашем интервале (иметь длину между l и L). Все эти описания должны были быть свободны и разрешены, когда Алиса сделала свой ход, поэтому доля свободных разрешенных слов длины L не меньше

$$(N - 2^l)/2^L$$

(Если были использованы свободные разрешенные слова на промежуточной высоте, это только увеличивает это отношение, так как каждое из них может быть заменено свободным словом на высоте L.)

Мы приходим к противоречию, если

$$(N - 2^l)/2^L \geq \varepsilon,$$

другими словами, если

$$N \geq \varepsilon \cdot 2^L + 2^l$$

Вспомним, что общий капитал Алисы ограничен 2^{-c}, а капитал, который она потратила, равен

$$(\varepsilon \cdot 2^L + 2^l) \cdot 2^c \cdot 2^{-L} = \varepsilon \cdot 2^c + 2^c/2^{L-l}.$$

Поэтому Алиса выигрывает, если обе величины $\varepsilon \cdot 2^c$ и $2^c/2^{L-l}$ не больше $2^{-c}/2$. Оба эти ограничения выполнены, например, если

$$\varepsilon = 2^{-3c} \text{ и } L - l \geq 3c.$$

3.8. Выигрышная стратегия Алисы

В результате мы имеем следующую стратегию для Алисы.

Для фиксированного c Алиса ждет, пока появится такой интервал $[l, L]$, что выполнены следующие три свойства:

- $L - l \geq 3c$;
- разрешенные слова представляют одно и то же базовое множество на всех уровнях между l и L;
- (текущая) доля свободных разрешенных слов на высоте L мала (меньше $\varepsilon = 2^{-3c}$).

Как только такой интервал появляется, Алиса выделяет меру $2^c \cdot 2^{-L}$ на каждое из $N = \varepsilon \cdot 2^L + 2^l$ новых слов (слов, у которых еще нет описания). Эта стратегия очевидно вычислима по c.

Как мы уже показали, это гарантирует победу Алисы, требуемое неравенство $q_c(u) \geq 2^c \cdot 2^{-KS_D(u)}$ будет выполнено для одного из этих слов и для D, возникающего в пределе.

3.9. Почему стратегия работает

Осталось показать, что событие, которого ожидает Алиса, действительно случится. Предположим, что это не так. Вспомним, что (по построению) каждое базовое множество меры больше чем треть представлено бесконечно много раз на высотах, создающие непрерывные блоки, и почти все такие блоки (кроме конечного числа) достаточно большие (имеют размер больше чем $3c$). Поэтому доля свободных разрешенных вершин на самых высоких уровнях каждого блока никогда не опускается ниже $\varepsilon = 2^{-3c}$.

Покажем, что это предположение противоречит предположению об оптимальности D. Зафиксируем некоторый блок («первый блок»), который является достаточно большим (не важно, какое именно множество представляют все его уровни), и дождемся, пока доля свободных разрешенных вершин на его верхнем уровне не стабилизируется (мы не можем сделать это вычислимо, но это не важно). Пусть B_0 — базовое множество, которое представляется множеством свободных разрешенных вершин на этой высоте. По предположению его размер не меньше ε.

Если размер B_0 хотя бы 1/3, то мы приходим к противоречию с леммой о плотности. Значит, его размер меньше 1/3 (а, значит, и 2/3), и есть второй блок над ним, где запрещенные (=неразрешенные) элементы представляют собой B_0. В верхнем уровне этого блока доля свободных разрешенных слов не опускается ниже ε. Дождемся, пока она стабилизируется и положим B_1 равным базовому множеству, которое представляется этими свободными разрешенными словами. По построению B_0 и B_1 не пересекаются (так как во время построения B_1, все слова из B_0 были запрещены).

Если суммарная мера $B_0 \cup B_1$ не меньше 1/3, мы опять получили противоречие с леммой о плотности (поскольку $B_0 \cup B_1$ и $\mathbf{D}$ не пересекаются, мы дожидались, пока результат стабилизируется). Таким образом, мы можем найти третий блок, где $B_0 \cup B_1$ запрещены, дождаться стабилизации его верхнего уровня, построить B_2 и т.д.

В результате мы приходим к противоречию: каждый блок добавляет хотя бы ε к множеству и в какой-то момент это множество станет больше, чем 1/3.

Технические замечания: (1) Граница 1/3 может быть изменена на любое другое значение, меньшее 1/2: мы хотим прийти к противоречию до того, как запрещенная часть станет слишком большой. В нашем рассуждении мы можем запретить до двух теретей от всех слов, но уже одной трети достаточно чтобы прийти к противоречию.

(2) Множества $B_0, B_1, \ldots$ не являются вычислимыми, но это не необходимо, они не используются в стратегии. Мы всего лишь доказываем (неэффективно), что существует момент, когда доля свободных разрешенных слов упадет ниже ε для какого-то блока.

Глава 4

Сравнение KA и KP

4.1. Введение

В этом разделе мы напомним определения, необходимые для понимания результата, дадим более подробную мотивацию, и обсудим сам результат.

Напомним, что дискретная априорная вероятность $\mathbf{m}$ — это наибольшая (с точностью до мультипликативной константы) перечислимая снизу полумера. При этом её логарифм равен префиксной колмогоровской сложности (верна формула $\mathbf{m}(x) = 2^{-KP(x)}$).

Непрерывной полумерой мы называем вещественнозначную неотрицательную функцию A, определенную на всех двоичных словах, со следующими свойствами:

- $A(\Lambda) = 1$, где Λ — пустое слово;
- $A(x) \geq A(x0) + A(x1)$ для любого слова x.

Среди непрерывных перечислимых снизу полумер есть максимальные, одна из которых фиксирована. Она называется непрерывной априорной вероятностью, обозначается $\mathbf{a}$, её логарифм называется априорной сложностью KA.

В главе 1 мы обсуждали понятие случайности последовательности по Мартин-Лёфу. Как мы уже обсуждали, Гач и Левин нашли эквивалентную формулировку в терминах сложности. Они же изучали промежуточную формулировку в терминах функции *дефекта случайности*. Опишем ее более подробно. *Тестом Левина–Гача* относительно меры μ называется перечислимая снизу функция t на канторовском пространстве, принимающая вещественные значения (возможно, бесконечные), с конечным интегралом

$\int t(\omega)\, d\mu(\omega)$. Для фиксированной вычислимой меры μ существует максимальный (с точностью до мультипликативной константы) тест Левина–Гача. Будем обозначать его $\mathbf{t}(\omega)$. Левин и Гач показали, что последовательности, на которых конечен максимальный тест — это в точности случайные по Мартин-Лёфу последовательности.

Характеристика случайных последовательностей в терминах сложности (которую мы уже обсуждали в главе 1), является следствием того, что максимальный тест может быть представлен в терминах априорной вероятности[1]:

$$\mathbf{t}(\omega) = \sum_{x \sqsubset \omega} \frac{\mathbf{m}(x)}{\mu(x\Omega)};$$

где $x \sqsubset \omega$ означает, что конечное слово x является началом бесконечного слова ω. Заметим, что $\mathbf{t}$ с левой стороны равенства зависит от μ (хотя это и не отражено в обозначениях). Как мы уже обсуждали, сумма в этой формуле может быть заменена на супремум.

Для равномерной меры на Канторовском пространстве этот результат может быть переписан в следующем виде (все равенства верны с точностью до мультипликативной константы):

$$\mathbf{t}(\omega) = \sum_n 2^{n - KP(\omega_1 \ldots \omega_n)} = 2^{\sup_n (n - KP(\omega_1 \ldots \omega_n))}.$$

Из этого равенства следуют оба критерия случайности Шнорра–Левина (подробнее см. [8, 20]; критерий с префиксной сложностью утверждает, что ω является случайной по Мартин-Лёфу относительно равномерной меры тогда и только тогда, когда $n - KP(\omega_1 \ldots \omega_n)$ ограничена, см. [12]) и “ample excess lemma” Миллера–Ю [17, Section 2] утверждающая, что сумма с правой стороны конечна для случайных ω.

Равномерная случайность, которую мы уже упоминали, получается, если использовать равномерные тесты (являющиеся функциями двух аргументов: меры и последовательности) вместо обычных. Больше деталей про равномерную случайность и нейтральную меру есть в статьях Левина [8, 9].

Как мы обещали, в этой главе мы будем изучать обобщение понятия случайности на случай непрерывных полумер. Мы попробуем использовать формулу, введенную Гачем, в качестве определения случайной последовательности. Можно было бы сказать, что последовательность ω случайна относительно непрерывной полумеры A, если конечна сумма $\sum_{x \sqsubset \omega} \mathbf{m}(x)/A(x)$,

[1] Эта формула впервые встречается в статье [14], но Гач в ней использовал другие более громоздкие обозначения, см. [1].

или конечен соответствующий супремум $\sup_{x\sqsubset\omega} \mathbf{m}(x)/A(x)$. Если использовать эти определения для максимальной полумеры $\mathbf{a}$, то для определения с супремумом все последовательности оказываются случайными (более того, дефект равномерно ограничен: легко видеть, что $\mathbf{m}(x)/\mathbf{a}(x) \le O(1)$ для всех слов x). Более того, в 2010 году Lempp, Miller, Ng и Turetsky (работа не опубликована, мы благодарим J. Miller'а за возможность ознакомиться с ней) показали, что для любой последовательности ω отношение $\mathbf{m}(\omega_1 \ldots \omega_n)/\mathbf{a}(\omega_1 \ldots \omega_n)$ стремится к нулю при $n \to \infty$.

В этой главе мы покажем (теорема 4.2 в части 4.2), что верхнюю грань тут нельзя заменить на сумму: предположение о том, что сумма $\mathbf{m}(x)/\mathbf{a}(x)$ по всем началам любой бесконечной последовательности конечна, неверно. Таким образом, первое определение случайности относительно непрерывной полумеры (в котором используется сумма) отличается от второго (в котором используется супремум): если пользоваться первым определением, не все бесконечные последовательности случайны относительно $\mathbf{a}$.

Было бы интересно лучше понять, для каких бесконечных слов сумма

$$\sum_{x\sqsubset\omega} \mathbf{m}(x)/\mathbf{a}(x)$$

бесконечна. Связаны ли они как-нибудь с K-тривиальными словами (такими, что $\mathbf{m}(x)$ равно $\mathbf{m}(l(x))$ с точностью до ограниченного множителя)? Мы не знаем ответа. Мы покажем лишь (см. раздел 4.5), что существует такое перечислимое множество, что его характеристическая функция обладает этим свойством.

Результат про сумму $\mathbf{m}(x)/\mathbf{a}(x)$ имеет вычислительную природу: если мы разрешим $\mathbf{a}(x)$ использовать оракул для проблемы остановки, то сумма станет конечной. Покажем это.

Предложение **4.1**. *Пусть* $\mathbf{a}' = \mathbf{a}^{\mathbf{0}'}$ *— непрерывная априорная вероятность, использующая оракул* $\mathbf{0}'$. *Тогда сумма* $\sum_{x\sqsubset\omega} \frac{\mathbf{m}(x)}{\mathbf{a}'(x)}$ *ограниченна константой* (*не зависящей от* ω) *для всех* ω.

Доказательство. Достаточно построить такую $\mathbf{0}'$-вычислимую *меру* a', что

$$\sum_{x\sqsubset\omega} \frac{\mathbf{m}(x)}{a'(x)} \le 1$$

для всех ω. (После этого можно заметить, что $\mathbf{a}'(x)$ не меньше чем a'.) Явно

предъявим такую меру. Начнем с того, что посчитаем априорные вероятности всех слов x, начинающихся с нуля и единицы:

$$M_0 = \sum_u \mathbf{m}(0u); \quad M_1 = \sum_u \mathbf{m}(1u).$$

(Заметим, что $M_0 + M_1 + \mathbf{m}(\Lambda) \le 1$, где Λ — пустое слово, т.е корень дерева.) Разобьем 1 на $a'(0)$ и $a'(1)$ в соответствующей пропорции, т.е.

$$a'(0) = \frac{M_0}{M_0 + M_1}, \quad a'(1) = \frac{M_1}{M_0 + M_1}.$$

Продолжим тем же образом, разбивая $a'(0)$ на $a'(00)$ и $a'(01)$ в отношении $M_{00} : M_{01}$, и так далее. Здесь M_z определенно для любого слова z и равно $\sum_u \mathbf{m}(zu)$.

Числа M_z перечислимы снизу и поэтому $\mathbf{0'}$-вычислимы (и строго положительны), а значит, мера a' определена и $\mathbf{0'}$-вычислима. Осталось проверить, что она достаточно велика, чтобы сумма из формулировки была ограничена единицей.

Достаточно доказать эту оценку для конечных сумм (т.е. для вершин высоты N: если все частичные суммы ряда не больше 1, то и вся сумма не больше). Доказывать утверждение мы будем индукцией по N. Предположим, что это утверждение верно для левого и правого поддеревьев с соответствующим масштабированием. (Сумма кончается на том же уровне N, а значит высота дерева на 1 меньше и можно применить предположение индукции. База индукции очевидна, в корне отношение $\mathbf{m}(\Lambda)/a'$ не больше 1 по очевидным причинам.) Сумма $m(x)$ в левом поддереве ограничена (на самом деле даже равна) M_0, вместо 1 в исходном дереве; в правом поддереве сумма ограничена M_1. С другой стороны, значения a' в корнях этих деревьев (т.е. $a'(0)$ и $a'(1)$) тоже меньше. Поэтому предположение индукции утверждает, что для любого пути в левом поддереве сумма $\mathbf{m}(x)/a'(x)$ ограничена величиной $M_0/a'(0)$, а для каждого пути в правом поддереве эта сумма ограничена величиной $M_1/a'(1)$. Поэтому осталось только показать, что

$$\frac{M_0}{a'(0)} + \mathbf{m}(\Lambda) \le 1, \qquad \frac{M_1}{a'(1)} + \mathbf{m}(\Lambda) \le 1.$$

Вспомним, что мы определяли $a'(0)$ и $a'(1)$ так, чтобы они были пропорциональны соответственно M_0 и M_1 и давали в сумме единицу. Таким образом, оба отношения в последней формуле равны $M_0 + M_1$ и остается вспомнить, что $M_0 + M_1 + \mathbf{m}(\Lambda)$ в точности равно сумме $\mathbf{m}(x)$ по всем словам x и, следовательно, ограничено единицей. □

4.2. Основной результат и набросок доказательства: игровая техника

Теорема 4.2. *Существует такая последовательность из нулей и единиц ω, что*

$$\sum_{x \sqsubset \omega} \frac{\mathbf{m}(x)}{\mathbf{a}(x)} = \infty.$$

Это основной результат, который будет доказан в этой главе. Доказательство использует (теперь довольно стандартную) игровую технику. В этой части мы опишем некоторую бесконечную игру и покажем, как основной результат следует из существования вычислимой выигрышной стратегии в этой игре для одного из игроков (названного «математиком» или сокращённо **M**). Далее в разделе 4.3 мы сведем игру к конечной (точнее говоря, к классу конечных игр) и покажем, что если все такие игры имеют равномерно вычислимую выигрышную стратегию для **M**, то бесконечная игра имеет вычислимую выигрышную стратегию. Наконец, в разделе 4.4 мы построим (индуктивно) выигрышную стратегию для конечной игры. Это наиболее технически-сложная часть доказательства, где оценка потребует интегрирования некоторого не очень сложного выражения.

Опишем бесконечную игру с полной информацией между двумя игроками Математиком (**M**) и её Оппонентом (**A**). Мы будем говорить об обоих игроках в женском роде; обозначения **M** и **A** можно считать сокращениями от «Mathematician» и «Adversary», они также соответствуют обозначениям для строимых ими функций m и a. Эта игра проходит на бесконечном двоичном дереве.

Математик увеличивает свои рациональные веса (мы будем называть их *m-весами*) на вершинах дерева (= двоичных словах). Вначале все веса равны нулю. На каждом своем ходу **M** может увеличить конечное множество весов, но не может уменьшить их. Сумма весов, использованных **M**, не должна превышать 1. (Мы можем считать, что **M** мгновенно проигрывает, если использует слишком большой суммарный вес.) Текущий вес **M** вершины x мы будем обозначать $m(x)$, таким образом, условие переписывается в виде $\sum_x m(x) \le 1$ в любой момент игры (иначе **M** уже проиграла).

Оппонент также увеличивает свои неотрицательные рациональные веса (мы будем называть их *a-весами*) на вершинах дерева, кроме корня. Изначально они также равны нулю, кроме корня, где вес равен 1. Однако усло-

вие на веса, которые она использует, другое: для любой вершины x должно быть выполнено неравенство $a(x0) + a(x1) \le a(x)$. Неформально говоря, можно интерпретировать $a(x)$ как поток, который приходит в вершину x. Поток размера 1 поступает в корень. Из корня некоторые части потока (а именно $a(0)$ и $a(1)$) передаются левому и правом сыну соответственно (в то время, как оставшаяся часть $1 - a(0) - a(1)$ зарезервирована на будущее). На следующем уровне, скажем в вершине 0, входящий поток $a(0)$ разбивается на $a(00)$, $a(01)$ и (неотрицательный) резерв $a(0) - a(00) - a(01)$, и так далее. По мере развития игры входящий поток (из предка) растет и он может быть передан потомкам или использован в качестве резерва. Если **A** нарушает это ограничение (неравенство $a(x0) + a(x1) \le a(x)$ для какого-то x), она сразу проигрывает.

Можно считать, что ходы игроков чередуются, однако это не очень важно: исход (бесконечной) партии в такой игре определяется предельной позицией и задержка некоторых ходов никогда не мешает (и, более того, может даже упростить игру, потому что у второго игрока меньше информации). Мы скажем, что **M** выигрывает в данной партии, если есть бесконечный путь в дереве (двоичная последовательность) ω, для которого

$$\sum_{x \sqsubset \omega} \frac{m(x)}{a(x)} = \infty,$$

где $m(x)$ и $a(x)$ — это предельные значения весов **M** и **A** соответственно. Нужно договориться только, что делать, если некоторые значения в знаменателе равны нулю. Это несущественно, поскольку Оппонент может легко сделать все свои веса неотрицательными. Однако для простоты можно считать, что $m/0 = \infty$, если $m \ne 0$ и $0/0 = 0$.

Теперь игра полностью определена. Поскольку все ходы — это конечные объекты, можно говорить про вычислимые стратегии. Следующая лемма — основной шаг в доказательстве теоремы 4.2.

***Лемма* 4.3.** *У **M** есть вычислимая выигрышная стратегия в этой игре.*

Лемма будет доказана в следующих двух разделах. Этот раздел мы закончим объяснением того, как утверждение теоремы 4.2 следует из леммы. Это стандартное рассуждение, которое используется в большинстве доказательств игровым методом. Пусть Оппонент игнорирует наши (Математика) ходы и просто перечисляет снизу значения непрерывной априорной вероятности $\mathbf{a}(x)$. (Они перечислимы снизу, дополнительно нужно позаботиться о

том, чтобы неравенство $a(x) \geq a(x0) + a(x1)$ выполнялось не только для предельных значений, но и для всех промежуточных, но это может быть сделано наиболее прямолинейным способом — Оппонент может увеличивать $a(x)$ вдоль всего пути вплоть до корня, если это необходимо.)

Действия **A** вычислимы. Пусть **M** использует свою вычислимую выигрышную стратегию против такого оппонента. Тогда поведение **M** вычислимо. Таким образом, предельные значения $m(x)$ — это перечислимая снизу функция, условия выигрыша утверждают, что для некоторой последовательности ω сумма $\sum_{x \sqsubset \omega} m(x)/\mathbf{a}(x)$ бесконечна. Осталось заметить, что дискретная априорная вероятность $\mathbf{m}(x)$ является оценкой сверху (с точностью до мультипликативной константы) для любой перечислимой снизу функции $m(x)$, удовлетворяющей условию $\sum_x m(x) < 1$.

4.3. Сведение к конечным играм

Чтобы построить выигрышную стратегию для **M** в бесконечной игре, описанной выше, мы объединим стратегии для похожих конечных игр. Конечная игра определяется двумя параметрами N и k; где N — высота конечного двоичного дерева, на котором происходит игра, а k — это сумма, которую **M** должна достичь, чтобы победить в этой игре. Здесь N — положительное целое число, а $k \geq 1$ — рациональное.

Изначально все вершины (= все конечные слова длины не больше N) имеют нулевые a- и m-веса, кроме корня, у которого единичный a-вес: $a(\Lambda) = 1$. Игроки ходят по очереди. На каждом ходу каждый из игроков может увеличить свои веса (рациональные числа) на любых вершинах, но оба игрока должны соблюдать свои ограничения: сумма m-весов не должна превосходить 1; для всех x (кроме листьев) должно быть выполнено $a(x) \geq a(x0)+a(x1)$; значение $a(\Lambda)$ остается 1. Позиция считается *выигрышной* для **M**, если есть такой лист w, что сумма $\sum_{x \sqsubset w} m(x)/a(x)$ хотя бы k. В противном случае позиция считается выигрышной для **A**. Каждый из игроков должен находиться в выигрышной позиции после своего хода (иначе проигрывает бесконечную игру). Также возможно проиграть игру, нарушив свои ограничения.

Лемма **4.4.** *Для всех положительных рациональных k существует N и выигрышная стратегия для **M**, которая гарантирует победу **M** после конечного числа ходов, при этом количество ходов до выигрыша зависит только от k (и не*

зависит от ходов **A**). *Значение* N, *оценка на количество ходов и стратегия вычислимы по* k.

Доказательство этой леммы будет дано в следующем разделе. В этом разделе мы покажем как, используя стратегию для конечных игр, выиграть в бесконечной игре, описанной в прошлом разделе (и тем самым, сведем основной результат, теорему 4.2, к этой лемме). Начнем с нескольких простых наблюдений.

Во-первых, заметим, что если у **M** есть выигрышная стратегия для какого-то N, то у неё есть выигрышная стратегия и для всех бóльших N (она может просто игнорировать вершины на высоте больше N). Таким образом, фраза «существует N» может быть заменена на «для всех достаточно больших N».

Во-вторых, игру можно масштабировать, ограничив общий вес **M** какой-то константой M (вместо 1) и положив $a(\Lambda)$ равным A (также вместо 1). Если **M** могла достичь суммы k в исходной игре, то в новой игре, используя практически ту же стратегию, она может достичь суммы kM/A. Для этого она должна представить, что ходы **A** в A раз меньше, и умножать ходы, рекомендованные стратегией, на M.

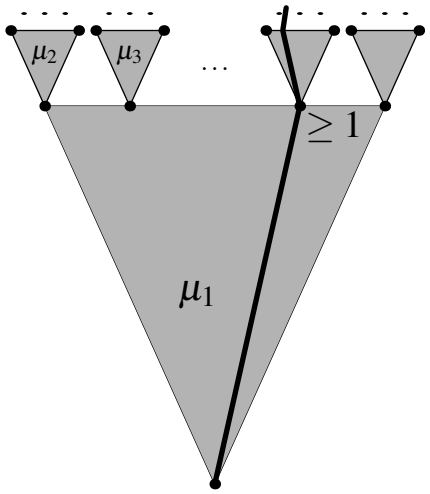

Рис. 4.1: Конечные поддеревья выбраны внутри бесконечного двоичного дерева. На этих поддеревьях **M** применяет выигрышную стратегию используя не более $\mu_1, \mu_2, \ldots$ в каждом, и строя пути с суммой 1 в каждом поддереве.

Поскольку k в лемме 4.4 произвольно, **M** может достичь сколь угодно большой суммы, даже если её веса ограничены сколь угодно малой константой $\mu > 0$ (известной заранее); размер дерева N зависит одновременно от того, какую сумму мы хотим достичь и того, насколько маленькую квоту μ хотим

использовать. Это наблюдение позволяет **M** использовать несколько стратегий параллельно на нескольких поддеревьях, предварительно выделив на каждое поддерево свою квоту $\mu_1, \mu_2, \mu_3, \ldots$, с условием $\sum \mu_i \leq 1$ для какой-нибудь вычислимой последовательности квот, например, $\mu_i = 2^{-i}$. Мы хотим достичь суммы 1 в каждом поддереве. Это возможно: можно считать, что потоки в каждом поддереве создаются Оппонентом независимо, если общий поток из корня не больше 1, то он не больше 1 и в каждой вершине, включая корни поддеревьев. (Заметим, что использование $a(\Lambda) < 1$ в корне вместо 1 делает игру сложнее для Оппонента, поэтому **M** выиграет в каждом поддереве, если будет играть так, как будто Оппонент может использовать всю меру 1.) На рис. 4.1 показано, как нужно выбирать поддеревья.

Зная μ_1, мы выбираем высоту первого поддерева; зная число листьев первого поддерева и соответствующие μ_i, мы можем выбрать достаточную высоту для поддеревьев второго уровня (можно выбрать одну общую высоту, чтобы картинка выглядела красивее); далее, зная количество листьев у всех них, мы смотрим на соответствующие μ_i и выбираем высоту для поддеревьев третьего уровня и т.д. Игры играются (и выигрываются) независимо для каждого поддерева. В каждом поддереве есть путь с суммой $\sum m(x)/a(x) \geq 1$, и мы можем собрать такие пути в один бесконечный начинающийся из корня. Сумма по нему оказывается бесконечной.

4.4. Выигрышная стратегия в конечной игре

В этом разделе мы докажем лемму 4.4 и тем самым закончим доказательство основного результата, теоремы 4.2. Как мы видели, выигрышная стратегия для Математика должна существенно использовать то, что игра содержит много раундов: если **M** делает только один шаг и останавливается, Оппонент может выиграть, разделяя в каждой вершине поток пропорциональна весам поддеревьев (см. Предложение 4.1).

Покажем сначала, что существует $k > 1$, для которого утверждение верно (для некоторого N). Сначала опишем выигрышную стратегию **M**, считая выигрышем существование пути просто с суммой больше 1. Для этого достаточно дерева высоты 2 (и даже некоторой его части).

M начинает с того, что увеличивает веса до $\frac{1}{4}$ на вершинах 0 и 00 (рис. 4.2). Теперь **A** должна решить, сколько потока она хочет отправить на 0 и 00. Возможны следующие варианты:

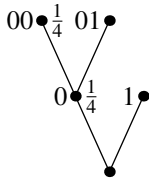

Рис. 4.2: Первый ход **M**.

- Поток на 0 маленький: $a(0) < \frac{1}{2}$. Тогда $a(00)$ по очевидным причинам меньше $\frac{1}{2}$, и
$$\frac{m(\Lambda)}{a(\Lambda)} + \frac{m(0)}{a(0)} + \frac{m(00)}{a(00)} > 0 + \frac{1}{2} + \frac{1}{2} > 1.$$
Этот ход не может приводить в выигрышную позицию для **A** (т.е. Оппонент проиграл).

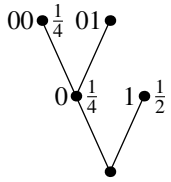

Рис. 4.3: Выигрывающий ход **M** во втором случае.

- Поток через 0 большой: $a(0) > \frac{1}{2}$. В этом случае **A** может быть в выигрышной позиции. Однако, **M** по-прежнему может выиграть. В самом деле, $a(1) \leq 1 - a(0)$ меньше $\frac{1}{2}$ и навсегда останется меньше 1/2. А значит, **M** может положить вес $\frac{1}{2}$ на вершину 1 (рис. 4.3), делая сумму по этому пути больше чем 1, и **A** не может ничего сделать.

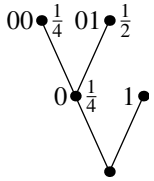

Рис. 4.4: Выигрывающий ход **M** в третьем случае.

- Промежуточный случай: $a(0) = \frac{1}{2}$. В этом случае, $a(00)$ также должно быть равно $\frac{1}{2}$, иначе сумма на пути 00 будет по-прежнему больше 1 и **A** не находится в выигрышной позиции. Однако, если $a(00) = a(0) = 1/2$, то **M** может положить свой вес $\frac{1}{2}$ в вершину 01, и **A** не может сделать

поток больше 1/2 через 01 (т.к. 1/2 уже направлена в 00). Итого по пути 01:

$$\frac{m(\Lambda)}{a(\Lambda)} + \frac{m(0)}{a(0)} + \frac{m(01)}{a(01)} \geq 0 + \frac{1}{4} + \frac{1/2}{1/2} = \frac{5}{4} > 1.$$

Более аккуратный анализ показывает, что используя эту идею можно построить выигрышную стратегию для **M** при $k = 17/16$. Но нам необходимо сколь угодно большое k, поэтому мы не будем углубляться в детали здесь, а проведем подробный анализ сразу для более общей конструкции.

Выигрышная стратегия для произвольного k будет построена индуктивно: мы предположим, что у **M** есть выигрышная стратегия для некоторого k и затем используем эту стратегию, чтобы построить выигрышную стратегию для некоторого $k' = k + \varepsilon$, где $\varepsilon > 0$. Приращение ε зависит от k и довольно маленькое, но ограничено снизу некоторой строго положительной непрерывной функцией $f(k)$.

Повторяя эту конструкцию, мы получаем выигрышные стратегии для $k = k_i$, причем $k_1 = 1$ (для $k = 1$ выигрывающая стратегия тривиальна) и

$$k_{i+1} \geq k_i + f(k_i).$$

Заметим, что $k_i \to \infty$ при $i \to \infty$; в самом деле, если $k_i \to K$ для какого-то конечного K, то $k_{i+1} \geq k_i + f(k_i) \to K + f(K)$. Противоречие.

Чтобы объяснить эту конструкцию, давайте сначала более подробно посмотрим на простой пример, описанный выше (как получить сумму больше 1). Сделав первый ход, **M** сохраняет резерв, который позже может быть использован на вершине 1. Эта возможность создает угрозу для **A** и не дает ей пустить слишком большой поток через вершину 0. Примерно такая же угроза будет использоваться в сложной конструкции. Опять вершина 1 будет использоваться для угрозы такого типа. Если **A** направит слишком большой поток налево (в вершину 0), **M**, обнаружив это, использует всю свою оставшуюся меру, чтобы выиграть в правом поддереве.

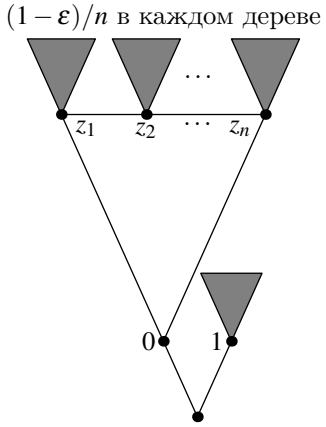

Рис. 4.5: Дерево для построения $(k + \varepsilon)$-стратегии.

Однако, стратегия в этот раз несколько более сложная. Два основных улучшения по сравнению с предыдущей состоят в следующем: во-первых, вместо

того, чтобы просто помещать какую-то меру в вершину, **M** использует масштабированную k-стратегию в дереве, корнем которого является эта вершина, что позволяет ей использовать свою меру более эффективно (с множителем k). Этот трюк используется и в левом поддереве, и при реализации угрозы в вершине 1. (Поддеревья, где используется k-стратегия, показаны серым на рис. 4.5.) Во-вторых, в левом поддереве **M** использует последовательно n вершин $z_1, \dots, z_n$ (и n соответствующих поддеревьев) для достаточно большого n. (Далее мы подробно обсудим, как выбрать n.)

Давайте опишем $(k+\varepsilon)$-стратегию более подробно.

На первом ходу **M** увеличивает до ε вес на вершине 0 (и **M** никогда больше не будет добавлять вес в эту вершину, поэтому вершина 0 всегда будет иметь вес ε).[2]

После этого у **M** осталось в резерве $1-\varepsilon$ веса. Она разделяет его на n равных частей, каждая по $(1-\varepsilon)/n$. Эти части используются последовательно в поддеревьях с корнями $z_1, \dots, z_n$ (рис. 4.5). В этих поддеревьях **M** использует масштабированную k-стратегию с коэффициентом подобия $(1-\varepsilon)/n$. Таким образом **M** заставляет **A** отправить большой поток в эти n поддеревьев или проиграть k-игру в одном из этих поддеревьев (и, таким образом, $(k+\varepsilon)$-игру во всем дереве, если параметры выбраны правильно).

Угроза на вершине 1 используется следующим образом: если в какой-то момент (после i игр для какого-то i) поток, отправленный **A** в вершину 0, слишком большой, то **M** изменяет свою стратегию и использует весь оставшийся вес (равный $(1-\varepsilon)(1-\frac{i}{n})$) для k-стратегии в поддереве с корнем 1 (И выигрывает, если параметры подобраны правильно).

Теперь осталось формализовать понятия «большой поток» и «слишком большой поток», выбрав некоторые границы. Предположим, что после i-го раунда **A** выделяет поток d в вершину 0. Тогда она может использовать только $1-d$ для игры в вершине 1. Используя k-стратегию с резервом $(1-\varepsilon)(1-\frac{i}{n})$, **M** может достигнуть суммы (по некоторому пути в правом поддереве)

$$\frac{k(1-\varepsilon)(1-\frac{i}{n})}{1-d},$$

[2] В этой стратегии вес вершины 0 равен приращению k, которое мы хотим получить; в этом нет глубокого смысла, можно использовать и другую константу, но использование одной константы несколько упрощает выкладки.

таким образом d_i находится из равенства

$$\frac{k(1-\varepsilon)(1-\frac{i}{n})}{1-d_i} = k+\varepsilon. \qquad (*)$$

Если поток в левое поддерево 0 не меньше d_i, **M** прекращает играть на левом поддереве и выигрывает игру, начав играть в правом поддереве согласно k-стратегии и используя весь свой оставшийся вес.

Что случается, если **A** не пересекает границу d_i? Тогда вершина 0 добавляет не меньше ε/d_i к сумме i-й игры, и, чтобы выиграть в большой игре, **M** достаточно получить сумму $(k+\varepsilon)-\varepsilon/d_i$ в i-й игре. Это можно сдедать, используя (масштабированную) k-стратегию с весом $(1-\varepsilon)/n$, если **A** отправит поток, меньший a_i, в поддерево с корнем z_i, где a_i определяется из равенства

$$k+\varepsilon-\frac{\varepsilon}{d_i} = k\frac{(1-\varepsilon)/n}{a_i} \qquad (**)$$

Нам осталось доказать, что для некоторого ε (зависящего от k) и для достаточно большого n значения a_i, определенные исходя из (**) (где d_i определены исходя из (*)) удовлетворяют неравенству

$$\sum_{i=1}^{n} a_i > 1.$$

В этом случае **A** не может отправить поток a_i в каждое из соответствующих поддеревьев (в поддеревья с корнями z_i для всех i) и проиграет игру.

Последняя сумма может быть переписана в следующем виде:

$$\sum_{i=1}^{n} a_i = \frac{1}{n}\sum_{i=1}^{n} \frac{kd_i(1-\varepsilon)}{d_i(k+\varepsilon)-\varepsilon}.$$

Заметим, что d_i зависит только от k, ε и $u=\frac{i}{n}$, таким образом, сумма может рассматриваться как частичная сумма для интеграла

$$\int_0^1 k(1-\varepsilon)\frac{d(u)}{(k+\varepsilon)d(u)-\varepsilon}du$$

где

$$d(u) = 1-\frac{k}{k+\varepsilon}(1-\varepsilon)(1-u).$$

Заметим, что мы интегрируем дробно-рациональную функцию вида $(Au + B)/(Cu + D)$, поэтому этот интеграл можно явно взять и получить

$$\int_0^1 \frac{(1-\varepsilon)(ku - ku\varepsilon + k\varepsilon + \varepsilon)}{(k+\varepsilon)(u+\varepsilon-u\varepsilon)}du = \\ = \frac{k(u+\varepsilon-\varepsilon u) + \varepsilon \log|u(\varepsilon-1)-\varepsilon|}{k+\varepsilon}\Bigg|_0^1 = \\ = \frac{k(1-\varepsilon) + \varepsilon \cdot \log(1/\varepsilon)}{k+\varepsilon}.$$

Заметим, что при $\varepsilon \to 0$ это выражение можно переписать как

$$1 + \varepsilon \cdot (\log(1/\varepsilon) - k - 1)/k + O(\varepsilon^2),$$

при достаточно маленьком $\varepsilon > 0$ этот интеграл будет больше чем $1 + \varepsilon$, и мы можем выбрать достаточно большое n, при котором риманова сумма будет больше 1. Легко получить положительную нижнюю границу для ε (как функцию от k) и эффективно найти n.

Это рассуждение заканчивает доказательство леммы 4.4 и, тем самым, доказательство основного результата, теоремы 4.2.

Замечание. Повторим основной аргумент доказательства: в начале, пока у **M** есть большой резерв, **A** не может отправить большой поток в вершину 0, поэтому вес ε, размещенный в этой вершине, берется с большим множителем. Если бы угрозы в вершине 1 не было, то **A** могла бы отправить весь поток в левое поддерево и у **M** ничего бы не получилось.

4.5. Усиление: перечислимое множество, характеристическая последовательность которого имеет бесконечную сумму

Некоторое изменение предыдущей конструкции позволяет построить последовательность ω, для которой $\sum_{x \sqsubset \omega} \mathbf{m}(x)/\mathbf{a}(x) = \infty$ и которая обладает некоторыми дополнительными свойствами.

***Теорема* 4.5**. *Существует такое перечислимое множество X, что для его характеристической последовательности ω_X (где $\omega_i = 1$ при $i \in X$ и $\omega_i = 0$ при $i \notin X$) сумма $\sum_{x \sqsubset \omega_X} \mathbf{m}(x)/\mathbf{a}(x)$ бесконечна.*

Доказательство. Начнем с модификации конечной игры в лемме 4.4. Потребуем, чтобы **M** на каждом ходу не только находилась в выигрышной позиции, но также явно выбирала одну из вершин дерева, где сумма не меньше k (сумма подсчитывается так, как описано в условиях на выигрышную позицию). В случае, если есть несколько листьев, где (текущая) сумма достигает k, **M** может выбрать любой из них. В ходе игры **M** может изменять выбранный узел, но мы дополнительно требуем монотонности: выбранные вершины должны быть покоординатно неубывающей последовательностью слов (для любой вершины двоичного дерева можно рассмотреть слово из нулей и единиц, ведущее в эту вершину и добавить бесконечно много нулей; когда изменяется выбранная вершина, новая последовательность должна получаться из старой несколькими заменами вида $0 \to 1$).

Это требование может быть удовлетворено с помощью небольшого изменения выигрышной стратегии для игры с $k = (k + \varepsilon)$. Вспомним, что выигрышная стратегия для $(k + \varepsilon)$ использует k-стратегию для вершин $z_1, \ldots, z_n$, и, возможно, вершины с угрозой. Используя индукцию, мы можем считать, что k-стратегия удовлетворяет условию монотонности. Этого еще не достаточно: стратегия **M** должна удовлетворять условию монотонности также при переходе от k-стратегии в поддереве с корнем z_i к k-стратегии в поддереве с корнем z_{i+1} (или при переходе к вершине угрозы). Чтобы это достичь, нужно несколько изменить стратегию **M**. Во-первых, нужно выбирать $z_1, \ldots, z_n$ так, чтобы покоординатно выполнялось $z_1 \le z_2 \le \ldots \le z_n$. Более того, во время игры в поддереве с корнем z_1, **M** ставит некоторые биты в 1 (в соответствии с выигрышной k-стратегией в поддереве). Эти биты не могут быть возвращены в ноль, но это не страшно: можно добавить достаточное количество единиц после z_2 (чтобы покрыть все биты, которые были изменены при игре в поддереве с корнем z_1), и использовать поддерево с корнем в этой вершине, сделать тот же трюк с z_3, и т.д. (см. рис. 4.6). И, наконец, тот же трюк можно сделать с вершиной угрозы.

Бесконечная игра из леммы 4.3 также может быть изменена. А именно, **M** должна после каждого своего хода поддерживать *текущую ветвь*: бесконечный путь в двоичном дереве, который содержит лишь конечное количество единиц (и поэтому является конечным объектом и может указываться **M** явно). Текущая ветвь может изменяться во время игры, но только монотонно. Другими словами, если прошлая ветвь шла направо на каком-то уровне, тогда и текущая тоже должна идти направо на этом уровне. Это требование

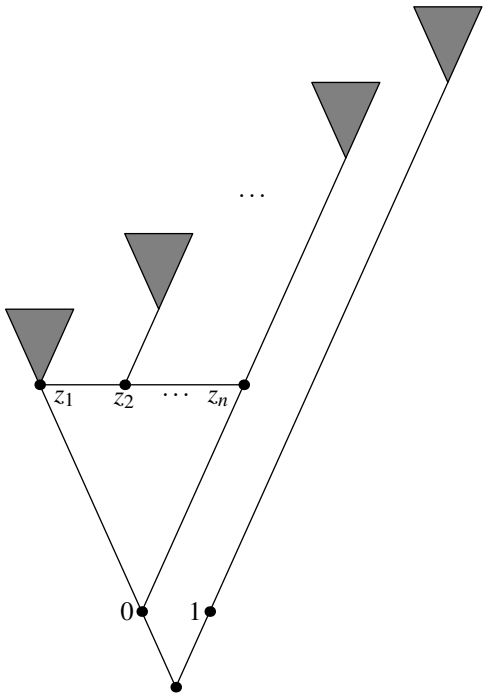

Рис. 4.6: Сохранение монотонности на шаге индукции

монотонности гарантирует существование предельной ветви и **M** выигрывает (бесконечную) игру, если сумма бесконечна вдоль этой ветви.

Мы утверждаем, что у **M** есть выигрышная стратегия в этой игре. Если это так, мы можем легко построить перечислимое множество, которое требуется в теореме 4.5. Мы используем вычислимую выигрышную стратегию против «слепого» Оппонента, который перечисляет снизу непрерывную априорную вероятность. Тогда поведение Оппонента вычислимо, поведение стратегии тоже вычислимо и предельная ветвь очевидно является характеристическим бесконечным словом некоторого перечислимого множества.

Осталось объяснить как можно объединить стратегии для конечных (модифицированных) игр, получив выигрышную стратегию для бесконечной игры. Мы не можем использовать стратегии параллельно, как мы это делали раньше, поскольку переход от одной игры к другой будет нарушать условие монотонности. Вместо этого мы начинаем играть в подигру с корнем в корне дерева. Эта стратегия делает ход, в частности, выбирает некоторый лист поддерева (текущий кандидат). После этого мы начинаем играть в поддереве, которое имеет корнем помеченный лист поддерева. Эта стратегия также выбирает своего кандидата и т.д. (рис. 4.7.)

В какой-то момент одна из стратегий может изменить свой отмеченный

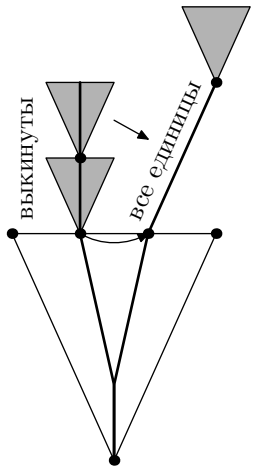

Рис. 4.7: Поддеревья, где игры уже начались. Когда изменяется отмеченная вершина, все поддеревья над ней выкидываются и новые поддеревья выбираются так, чтобы на позициях, которые могли быть использованы в ходе старых игр, были только единицы.

лист. Тогда все игры, которые проходили в потомках этого (теперь не отмеченного) листа, бесполезны и должны быть выброшены; мы начинаем игру над новой отмеченной вершиной. Чтобы удовлетворить условие монотонности, мы начинаем новую игру на достаточно большом уровне и заполняем промежуток между новой отмеченной вершиной и этим уровнем единицами. Это гарантирует, что все уже отмеченные единицы останутся единицами (мы предполагаем, что в любой момент началось только конечное число игр и текущая ветвь имеет лишь конечное число единиц).

Легко видеть, что в пределе у нас по-прежнему получается ветвь с бесконечной суммой. В самом деле, для игры в корне отмеченная вершина может покоординатно только возрастать и изменется конечное число раз. Таким образом, какой-то из отмеченных листьев в корневой игре останется отмеченным навсегда. Игра, которая началась над этим листом, останется навсегда, хотя отмеченный лист в этой игре тоже может несколько раз измениться (монотонно). Таких изменений может быть лишь конечное число, и после последнего такого изменения второй отмеченный лист тоже никогда не изменится, а значит, игра над ним тоже не будет выброшена, и так далее.

Монотонность гарантирована и на позициях внутри поддеревьев, где проходят конечные игры (из-за монотонности конечных игр), и на позициях вне (мы используем только единицы на этих позициях)

□

Глава 5

Очень большие числа

5.1. Введение

В 1962 году Т. Радо (Tibor Radó [19]) предложил для каждого n рассмотреть наибольшее число единиц, которое может напечатать перед остановкой машина Тьюринга, имеющая не более n внутренних состояний. Алфавит машины состоит из двух символов (пробела и еще одного). Машина начинает работу на пустой ленте; когда она останавливается, подсчитывается общее количество символов, отличных от пробела. Он доказал, что эта функция растет быстрее любой вычислимой функции от n. Это же верно и для других функций с похожими определениями (максимальное время работы такой машины, максимальный сдвиг головки). Недостатком этих определений является зависимость от конкретной модели вычислений (машины Тьюринга): даже небольшие изменения модели приводят, вообще говоря, к другим (но тоже быстро растущим) функциям.

Более инвариантный подход возможен в терминах колмогоровской сложности: можно рассматривать наибольшее число, имеющее сложность не более n, то есть наибольшее число, которое получается в результате работы программы из не более чем n битов. Здесь предполагается, что выбран оптимальный способ программирования (декомпрессор). Несложно показать (см., например, [3, раздел 1.2]), что та же самая функция (с точностью до изменения аргумента на константу) получится, если рассматривать время работы (в какой-то модели вычислений) оптимального декомпрессора (интерпретатора оптимального языка программирования) на входах (программах) длины не более n.

Более точно, пусть D — оптимальный декомпрессор для простой колмо-

горовской сложности, $KS(\cdot)$ — соответствующая функция сложности,

$$B(n) = \max\{N \mid KS(N) \le n\}.$$

Другими словами, $B(n)$ — максимальное значение декомпрессора на аргументах длины не больше n, если входами декомпрессора считать двоичные слова, а выходами — натуральные числа. Определим теперь $BB(n)$ как максимальное время работы D на входах длины не больше n (для произвольной машины, вычисляющей D, измеренное в произвольной вычислительной модели). Тогда найдется такая константа c, что

$$B(n-c) \le BB(n) \le B(n+c)$$

при всех n (см. [3]). Такая точность — изменение константы в аргументе — является максимально возможной, так как функция $KS(N)$ определена с точностью до ограниченного слагаемого. Поэтому мы не будем различать функции $B(n)$ и $BB(n)$, и будем использовать обозначение $B(n)$ для любой из них.

Это же рассуждение практически без изменений переносится на случай префиксной сложности. Мы можем определить

$$BP(n) = \max\{N \mid KP(N) \le n\};$$

можно также рассмотреть максимальное время работы оптимального беспрефиксного декомпрессора [3, раздел 4.4] на входах длины не больше n, и получится эквивалентное определение — с точностью до изменения аргумента на константу.

Исходя из этого, в дальнейшем мы рассматриваем только функции B и BP, определенные в терминах сложности, и не говорим о времени вычисления. Оказывается, что функции B и BP различаются. В этой главе мы сравниваем скорости их роста, а также скорость роста промежуточной функции BP', определяемой в терминах априорной вероятности.

Напомним, что априорная вероятность $\mathbf{m}(k)$ числа k определяется как k-й член максимального (с точностью до константы) сходящегося перечислимого снизу ряда с неотрицательными членами; как показал Левин (который и ввел понятия априорной вероятности и префиксной сложности),

$$\mathbf{m}(k) = 2^{-KP(k)+O(1)}$$

(доказательства можно найти, например, в [3, глава 4]). Теперь можно рассмотреть регулятор сходимости этого ряда, то есть для данного n рассмотреть

то наименьшее N, для которого

$$\sum_{k>N} \mathbf{m}(k) < 2^{-n}.$$

Обозначим это N через $BP'(n)$. Разница с определением $BP(n)$ состоит в том, что в $BP(n)$ мы ищем место, начиная с которого члены ряда $2^{-KP(n)}$ малы (меньше 2^{-n}), а в $BP'(n)$ нам важно место, с которого мал хвост ряда. Очевидно,

$$BP(n) \le BP'(n);$$

точнее, следовало бы написать $BP(n) \le BP'(n + O(1))$, поскольку обе величины определены с точностью до изменения аргумента на константу (напомним, что априорная вероятность определена с точностью до ограниченного множителя, а префиксная сложность — с точностью до ограниченного слагаемого).

В этой главе мы оцениваем, насколько отличаются эти три величины. Теорема 5.1 показывает, что они близки друг к другу. Из неё, в частности, вытекает, что все три различаются не более чем на $(1 + \varepsilon)\log n$ в аргументе. Теорема 5.2 показывает, что оценка теоремы 5.1 близка к точной и, более того, большой зазор может быть как между BP и BP', так и между BP' и B.

***Теорема* 5.1**. (i) *Существует такое c, что $BP(n) \le BP'(n+c)$ и $BP'(n) \le B(n+c)$ при всех n.*

(ii) *Существует такое c, что $B(n) \le BP(n + KP(n) + c)$ при всех n.*

(iii) *Пусть (x_n, y_n) — последовательность пар натуральных чисел, для которых $x_n \le y_n$, последовательность x_n перечислима снизу, а последовательность y_n перечислима сверху. Пусть при этом $\sum_n 2^{x_n - y_n} < +\infty$. Тогда существует такое c, что $B(x_n) \le BP(y_n + c)$ при всех n.*

Утверждения пунктов (i) и (ii) простые, а пункт (iii), как мы покажем, можно рассматривать как более симметричную переформулировку утверждения (ii). В нем используется понятие перечислимости снизу и сверху. Напомним (подробнее см., например, [3]), что перечислимость снизу последовательности натуральных чисел y_n означает, что y_n можно представить как предельные значения неубывающей по k вычислимой всюду определенной функции двух аргументов: $y_n = \lim_k y(n,k)$. Аналогично определена перечислимость сверху. Утверждение (ii) является частным случаем (iii) при $x_n = n$, $y_n = n + KP(n)$. Можно написать и симметричное неравенство: взяв

$x_n = n - KP(n)$ и $y_n = n$, мы получим, что $B(n - KP(n)) \le BP(n + c)$ при некотором c и всех n. Утверждение с $(1 + \varepsilon)\log n$ получится, если заметить, что ряд $\sum 2^{-(1+\varepsilon)\log n} = \sum(1/n^{1+\varepsilon})$ сходится при $\varepsilon > 0$ (и вычислим — без ограничения общности можно считать ε вычислимым, уменьшив его при необходимости).

В утверждениях (ii) и (iii) есть некоторая асимметрия: почему мы добавляем c в правую часть, вместо того, чтобы вычитать её из левой? Другими словами, можно рассмотреть и симметричные им утверждения:

(ii′) Существует такое c, что $B(n - KP(n) - c) \le BP(n)$ при всех n;

(iii′) в условиях (iii) существует такое c, что $B(x_n - c) \le BP(y_n)$ при всех n.

Эти утверждения тоже верны (и легко следуют из теоремы, мы вернемся к этому после её доказательства).

Теперь покажем, что если ряд $\sum_n 2^{x_n - y_n}$ расходится, то утверждение (iii) перестает быть верным, причем разрыв может быть больши́м для любого из двух неравенств.

***Теорема* 5.2.** *Пусть (x_n, y_n) — последовательность различных пар натуральных чисел, причем $x_n \le y_n$, последовательность x_n перечислима снизу, а последовательность y_n перечислима сверху. Пусть при этом $\sum 2^{x_n - y_n} = +\infty$. Тогда:*

(i) *найдется n, при котором $B(x_n) > BP'(y_n)$;*

(ii) *найдется n, при котором $BP'(x_n) > BP(y_n)$.*

В формулировке этой теоремы (в отличие от предыдущей) нет константы c, однако её можно добавить с любой стороны (или даже с обеих: от изменения всех x_n или всех y_n на константу сходимость ряда не изменится, а функции B, BP и BP' монотонны). Например, можно утверждать, что *для любого c найдется n, при котором $B(x_n) > BP'(y_n + c)$*, или *для любого c найдется n, при котором $B(x_n - c) > BP'(y_n)$*, и так далее.

Из теорем 5.1 и 5.2 следует, что если a_n — перечислимая сверху неотрицательная последовательность натуральных чисел, то любое из утверждений

- $BP(n) \le BP'(n + a_n + c)$ для некоторого c и для всех n;
- $BP'(n) \le B(n + a_n + c)$ для некоторого c и для всех n;
- $BP(n) \le B(n + a_n + c)$ для некоторого c и для всех n;
- $BP(n - a_n) \le BP'(n + c)$ для некоторого c и для всех n;

- $BP'(n - a_n) \le B(n + c)$ для некоторого c и для всех n;
- $BP(n - a_n) \le B(n + c)$ для некоторого c и для всех n

равносильно сходимости ряда $\sum 2^{-a_n}$ (что, в свою очередь, равносильно неравенству $a_n \ge KP(n) - O(1)$, см. [3]).

Можно сказать, что в теоремах 5.1 и 5.2 мы оцениваем разницу между *обратными к B, BP и BP' функциями*, причем оценка формулируется в терминах значений этих обратных функций. Скажем, мы видели, что они отличаются не более чем на логарифм значения любой из них с множителем $(1 + \varepsilon)$, но выбросить ε здесь нельзя: без него *оба* неравенства могут давать большой разрыв. Из верхней оценки следует, что этот большой разрыв достигается в разных местах (иначе суммарный разрыв превысил бы верхнюю оценку).

Утверждение (ii) теоремы 5.2 было доказано П. Гачем [15] в частном случае $x_n = n - a_n, y_n = n$ (что, по существу, эквивалентно нашей формулировке), так что основным результатом этой главы является утверждение (i) теоремы 5.2. Тем не менее для полноты (и удобства читателя) мы приведем доказательства всех утверждений.

Какие еще варианты определений можно рассмотреть? Можно было бы рассмотреть максимальное N, при котором $KS(N|n) \le n$ или $KP(N|n) \le n$. Но это не даст ничего нового: очевидно, что эти сложности не больше соответствующих безусловных сложностей; с другой стороны, если $KS(N|n) = n$, то эта же программа длины n для числа N при известном n может рассматриваться как беспрефиксная программа для N при известном n, поскольку при известном n мы знаем, где она заканчивается (зная её длину). Кроме того, её можно рассматривать и как безусловное описание N, так как в качестве условия используется длина программы, известная и так. И вообще $KP(x|\, KS(x)) = KS(x|\, KS(x)) = KS(x)$ с точностью до ограниченного слагаемого, см. [3, раздел 4.7].

Наконец, отметим, что можно определять функцию BP' и как регулятор сходимости вычислимых рядов с неотрицательными членами и случайными пределами.

***Теорема* 5.3**. *Пусть $\sum a_n$ — сходящийся вычислимый ряд с рациональными членами и случайным по Мартин-Лёфу пределом, и $N(\varepsilon)$ — его регулятор сходимости, то есть минимальное N, для которого $\sum_{n>N} a_n < \varepsilon$. Тогда*

$$BP'(n - c) \le N(2^{-n}) \le BP'(n + c).$$

В одном направлении (первое неравенство) это было показано в [11, Th. 19], в другую сторону это следует из максимальности априорной вероятности (она мажорирует любой вычислимый ряд). В [11] также показано, что если для какого-то вычислимого ряда с неотрицательными членами выполнено неравенство $BP(n-c) \le N(2^{-n})$ (для некоторого c и всех n), то аналогичное неравенство выполнено и для BP' (с другим c).

5.2. Верхние оценки

В этом разделе мы докажем теорему 5.1.

(i) Неравенство $BP(n) \le BP'(n+c)$ непосредственно следует из определения — и даже без константы c, если мы согласованно выбираем функции априорной вероятности и префиксной сложности: $\mathbf{m}(n) = 2^{-KP(n)}$.

Докажем теперь неравенство $BP'(n) \le B(n+c)$ для некоторого c и для всех n. Для этого укажем алгоритм, который по данному n перечисляет не более чем 2^n чисел (в порядке возрастания), и последнее из них оказывается больше $BP'(n)$. Затем n-битовую запись порядкового номера последнего числа можно взять в качестве описания этого последнего числа, что дает требуемую оценку. Сам алгоритм действует так: мы параллельно перечисляем снизу все значения $\mathbf{m}(n)$; если хвост текущего приближения к ряду $\sum \mathbf{m}(n)$ справа от последнего перечисленного числа становится больше 2^{-n} (то есть текущий кандидат в $BP'(n)$ превысил последнее перечисленное число), то мы указываем новое число, которое больше всех ненулевых (на данный момент) членов ряда $\sum \mathbf{m}(n)$. Ясно, что такое может случиться не более 2^n раз, так как на каждом шаге мы отрезаем новый (не пересекающийся с прежними) кусок ряда с суммой не менее 2^{-n}.

(ii) Хорошо известно, что $KP(x) \le KS(x) + KP(KS(x)) + O(1)$ (см., например, [3, раздел 4.6]). Небольшое обобщение: если $KS(x) \le n$, то $KP(x) \le n + KP(n) + O(1)$. В самом деле, программу для x длины не более n можно дополнить (слева) самоограниченной группой вида 0^k1, чтобы её длина стала ровно $n+2$, а до этого еще можно дописать самоограниченное описание числа n. Из этого обобщения непосредственно следует, что $B(n) \le BP(n + KP(n) + c)$ при некотором c и всех n.

(iii) Выведем это утверждение из (ii). Для этого сначала покажем, что без ограничения общности можно считать x_n и y_n вычислимыми.

Пусть x_n и y_n перечислимы снизу и сверху соответственно. Рассмотрим мо-

нотонно приближающие их пары (x_n^i, y_n^i). Выкинем повторяющиеся пары из этого перечисления. Таким образом, каждая оригинальная пара (x_n, y_n) даст конечное множество пар в новой вычислимой последовательности, которую мы будем обозначать $(\tilde{x}_i, \tilde{y}_i)$. Эта последовательность, очевидно, содержит все исходные пары x_n, y_n. При этом сумма $\sum_i 2^{\tilde{x}_i - \tilde{y}_i}$ отличается от $\sum_n 2^{x_n - y_n}$ не более чем вдвое, поскольку на каждом шаге приближения к окончательному значению для x_n и y_n член ряда удваивается, поэтому для новой последовательности сумма тоже конечна. Таким образом мы свели общий случай к случаю вычислимой последовательности пар.

Теперь предположим, что последовательность (x_i, y_i) вычислима. Рассмотрим функцию

$$f(n) = \min\{y_i \mid x_i = n\} - n.$$

Эта функция (бесконечная, если n не встречается среди x_i) перечислима сверху, при этом

$$\sum_n 2^{-f(n)} < +\infty.$$

Поэтому $f(n) > KP(n) - O(1)$ (максимальность априорной вероятности). Таким образом,

$$x_i + KP(x_i) < y_i + O(1)$$

и

$$B(x_i) \le BP(x_i + KP(x_i) + O(1)) \le BP(y_i + O(1))$$

в силу монотонности функции BP, что и требовалось доказать.

Несложно доказать и упоминавшиеся в предыдущем разделе симметричные утверждения

(ii′) $B(n - KP(n) - c) \le BP(n)$ при некотором c и всех n;

(iii′) в указанных условиях на x_n и y_n выполнено неравенство $B(x_n - c) \le BP(y_n)$ (при некотором c и при всех n).

Чтобы доказать (ii′), воспользуемся (ii) для меньшего значения аргумента:

$$B(n - KP(n) - e) \le BP(n - KP(n) - e + KP(n - KP(n) - e) + c)$$

для некоторого c и для любых n и e. Нам надо теперь подобрать значение e так, чтобы аргумент BP в правой части был не больше n при всех n:

$$n - KP(n) - e + KP(n - KP(n) - e) + c \le n.$$

Это возможно: вспомним, что

$$KP(n - KP(n) - e) \le KP(n, KP(n)) + KP(e) + O(1) \le KP(n) + KP(e) + O(1),$$

и при больших e значение $e - KP(e)$ может быть сколь угодно большим и перевесит имеющиеся константы.

Далее можно вывести (iii′) из (ii′) тем же способом, что мы выводили (iii) из (ii), нужно только группировать слагаемые с одним и тем же y_i (вместо x_i).

5.3. Нижние оценки

В этом разделе приведено доказательство теоремы 5.2.

5.3.1. Доказательство утверждения (i)

Пусть дана последовательность различных пар (x_n, y_n) с $x_n \le y_n$. Мы знаем, что последовательность x_n перечислима снизу, последовательность y_n перечислима сверху и $\sum 2^{x_n - y_n} = +\infty$. Нужно доказать, что найдется n, при котором $B(x_n) > BP'(y_n)$.

Сначала сведем общий случай к частному случаю $x_n = n$, $y_n = n + a_n$ (где a_n — перечислимая сверху целочисленная последовательность, возможно, с бесконечными значениями).

Для сведения применим тот же прием, что и в предыдущем разделе. Заменим x_n, y_n их приближениями (x_n^i, y_n^i), соединив их в одну последовательность и вычеркнув повторения. Сумма ряда могла только увеличиться (добавились новые члены), все члены по-прежнему различны (повторения мы вычеркнули), и если $B(x_n^i) > BP'(y_n^i)$, для какого-то приближения (x_n^i, y_n^i), то и $B(x_n) > BP'(y_n)$ в силу монотонности функций B и BP'. Таким образом, можно считать без ограничения общности, что (x_n, y_n) — вычислимая последовательность.

Определим $a_n = \min\{y_i \mid x_i = n\} - n$. Эта функция перечислима сверху в силу вычислимости (x_i, y_i). Заметим, что

$$\sum_n 2^{-a_n} \ge \tfrac{1}{2} \sum_i 2^{x_i - y_i}.$$

В самом деле, достаточно сгруппировать в сумме в правой части неравенства члены $2^{x_i - y_i}$ с $x_i = n$ и оценить геометрическую прогрессию со знаменателем

1/2 удвоенным первым членом. Поэтому $\sum_n 2^{-a_n} = +\infty$, что завершает сведение: все пары $(n, n + a_n)$ встречаются среди (x_i, y_i).

Теперь воспользуемся следующей леммой: *Если a_n перечислимая сверху последовательность и $\sum_n 2^{-a_n} = +\infty$, то существует вычислимая последовательность $\tilde{a}_n \geq a_n$, для которой так же выполнено $\sum_n 2^{-\tilde{a}_n} = +\infty$.* Доказательство леммы: рассмотрим перечисление a_n. В какой-то момент сумма начала ряда $\sum_n a_n$ окажется больше 1. Зафиксируем это приближение на этом отрезке как $\tilde{a}_n$ и повторим то же действие для хвоста ряда. В силу этой леммы, можно считать, что a_n вычислимо.

Итак, мы рассматриваем случай $(x_n, y_n) = (n, n + a_n)$ для вычислимой последовательности a_n. Мы знаем, что $\sum 2^{-a_n} = +\infty$, и хотим доказать, что найдется n, при котором $B(n) > BP'(n+a_n)$. Другими словами мы хотим доказать, что существует такое u (и n), что $KS(u) \leq n$ и $\sum_{i>u} \mathbf{m}(i) < 2^{-n-a_n}$.

Для того, чтобы оценить $KS(u)$ сверху необходимо построить декомпрессор, относительно которого у u есть короткое определение. Однако построение декомпрессора дает оценку на KS только с точностью до аддитивной константы, поэтому мы хотим построить декомпрессор D со следующим свойством: для любого d существуют такие u и n, что

$$KS_D(u) \leq n - d \quad \text{и} \quad \sum_{i>u} \mathbf{m}(i) < 2^{-n-a_n},$$

где $KS_D(u)$ обозначает минимальную длину слова p, при котором $D(p) = u$.

Строить этот декомпрессор и последовательность мы будем игровым методом. Рассмотрим следующую игру. Алиса перечисляет некоторые множества $D_0, D_1, \ldots$, добавляя в них натуральные числа; мы считаем, что на каждом своем ходу она может добавить лишь конечное число элементов к конечному числу множеств, так что её ход является конечным объектом. Боб перечисляет снизу полумеру μ. Изначально все $\mu(i)$ равны нулю, затем Боб их увеличивает; на своем ходу может увеличить μ на любом конечном множестве элементов $\mathbf{N}$. Увеличения мы считаем рациональными, так что ход Боба — тоже конечный объект. В множестве D_n должно быть не более 2^n элементов (иначе Алиса проигрывает); Боб должен соблюдать условие $\sum \mu(x) \leqslant 1$ (иначе проигрывает). Если оба игрока соблюдали ограничения, то выигрыш определяется так: Алиса выигрывает, если для любого d найдутся n и u, при которых $u \in D_{n-d}$ и $\sum_{i>u} \mu(y) < 2^{-n-a_n}$; в противном случае выигрывает Боб. Заметим, что в условии выигрыша можно элиминировать квантор по u: для

всех d должно существовать n удовлетворяющее:

$$\sum_{i>\max(D_{n-d})} \mu(i) < 2^{-n-a_n}. \qquad (*)$$

Мы построим вычислимую выигрышную стратегию для Алисы. Несложно заметить, что этого достаточно для того, чтобы предъявить требуемый декомпрессор. В самом деле, применим эту стратегию против Боба, который перечисляет априорную вероятность $\mu(i) = \mathbf{m}(i)$ (и игнорирует ходы Алисы), в результате получим перечислимые множества D_i. Рассмотрим декомпрессор, отображающий слова длины k в элементы множества D_k (например, можно рассматривать слова длины k как двоичные записи всех чисел от 0 до 2^k и отображать слово, соответствующее i, в i-ый элемент D_k). Этот декомпрессор обладает требуемым свойством в силу того, что Алиса выигрывает.

Осталось описать выигрышную стратегию для Алисы. Она хочет для каждого d построить свое n. Делать она это будет независимо для всех d. Для каждого d она фиксирует свой интервал натурального ряда $[l_d, r_d]$ в котором будет строить требуемое n. Она выбирает эти интервалы так, чтобы не было коллизий среди $n - d$ (т.е. интервалы $[l_d - d, r_d - d]$ не пересекаются для различных d). Так же эти интервалы должны быть достаточно большими: сумма 2^{-a_n} по $n \in [l_d, r_d]$ должна быть больше, чем 2^{d+1} (дальше мы покажем, что этого достаточно, чтобы построить n в этом интервале). Такие $[l_d, r_d]$ можно выбрать вычислимо, так как a_n вычислимо и $\sum_n 2^{-a_n} = +\infty$.

Алиса будет строить D_{n-d} для $n \in [l_d, r_d]$ независимо для всех d. Как она это будет делать? Она выберет какое-нибудь n (например, l_d) и будет пытаться достичь $(*)$ добавляя в D_{n-d} достаточно большие числа. Более точно, она будет добавлять число такое, что μ-мера всех больших равна 0 (т.е. будет добавлять число k, такое что $\mu(k') = 0$ при всех $k' \geq k$), если $(*)$ нарушено. Затем Боб будет увеличивать μ и $(*)$ окажется опять нарушенным, Алиса добавит свой элемент и так далее. Через некоторое время (через 2^{n-d} итераций) множество D_{n-d} достигнет своего максимального размера. К этому моменту Боб был обязан потратить хотя бы $2^{n-d}2^{-n-a_n} = 2^{-d-a_n}$ меры (каждый раз отрезался хвост размера 2^{-n-a_n}). После этого, Алиса переходит к следующему n в $[l_d, r_d]$, Боб обязан опять потратить 2^{-d-a_n} или проиграть и так далее. Так как сумма 2^{-d-a_n} по всем $n \in [l_d, r_d]$ больше 1, то ему придется проиграть в какой-то момент. (Техническое замечание: мы требуем, чтобы после предельного перехода сумма хвоста меры β была строго меньше некоторой константы, для

этого не достаточно, чтобы в каждый конечный момент времени сумма была меньше (при предельном переходе строгое неравенство может перейти в нестрогое). Чтобы обойти это препятствие, Алиса будет делать ход в тот момент, когда суммарный хвост μ достиг половины от той константы, которую она пытается избежать. В результате появится множитель 2, из-за которого мы требуем, чтобы сумма 2^{-a_n} по $n \in [l_d, r_d]$ была больше, чем 2^{d+1}, а не 2^d).

Утверждение (i) доказано.

5.3.2. Доказательство утверждения (ii)

Мы хотим доказать (вслед за Гачем), что для любой последовательности различных пар (x_n, y_n), где $x_n \le y_n$, x_n перечислима снизу, y_n перечислима сверху и $\sum 2^{x_n - y_n} = +\infty$, найдется n, при котором $BP'(x_n) > BP(y_n)$.

Повторяя рассуждение из (i), в этот раз сгруппировав пары с одинаковым y_i, мы сводим общую задачу к случаю $(x_n, y_n) = (n - a_n, n)$, где a_n вычислимы и $\sum_n 2^{-a_n} = +\infty$. Теперь мы хотим доказать, что существует такое n, что выполнено $BP'(n - a_n) > BP(n)$. Другими словами, мы хотим доказать, что существуют такие n и u, что $\mathbf{m}(i) < 2^{-n}$ для всех $i \ge u$ и $\sum_{i \ge u} \mathbf{m}(i) > 2^{-n+a_n}$ (все члены ряда $\mathbf{m}$ малы после u, но сумма начиная с u велика).

Для того, чтобы доказать, что сумма $\mathbf{m}$-хвоста велика, достаточно построить какую-нибудь перечислимую меру, для которой сумма велика и затем воспользоваться максимальностью $\mathbf{m}$. Опять появляется константа (в этот раз мультипликативная) и утверждение, которое мы хотим доказать формулируется так: существует такая перечислимая снизу полумера α, что для всех d существуют n и u со следующим свойством:

$$\sum_{i \ge u} \alpha(i) > 2^{-n+a_n+d} \quad \text{и} \quad \mathbf{m}(i) < 2^{-n} \text{ для всех } i \ge u.$$

Чтобы построить искомую полумеру, снова применим игровой метод. Алиса и Боб перечисляют снизу полумеры α и β соответственно, каждая из них в сумме не должна превышать 1. Алиса выигрывает, если выполнено указанное свойство (с β вместо $\mathbf{m}$) и проигрывает иначе.

Сначала заметим, что Алисе достаточно научиться выигрывать в игре, где в условиях выигрыша зафиксировано d. В самом деле, в этом случае Алиса может выиграть d-игру использовав не больше 2^{-d} меры (для этого она должна применить $2d$-стратегию, умноженную на 2^{-d}). После этого она может играть для всех d параллельно, рассматривая каждый ход Боба как ход в каждой

из игр. Суммарно она потратит не больше единицы, и условие на выигрыш монотонно по мере, а значит мера, использованная Алисой в других подиграх, не усложняет ей выигрыш в фиксированной подыгре, и она выиграет для всех d.

Осталось явно описать выигрышную стратегию для Алисы в d-игре. Она будет увеличивать свою меру маленькими шагами достаточно далеко (дальше, чем вся ненулевая мера Боба), если она не находится в выигрышной позиции. Как только мера, которую она потратила превысит 2^{-n+a_n+d} (для некоторого n), Боб будет обязан положить хотя бы 2^{-n} на какое-то i. После этого, когда Алиса вновь потратит 2^{-n+a_n+d} Боб будет обязан построить другое i с мерой хотя бы 2^{-n} и так далее. Если Алиса делает достаточно маленькие шаги (более подробно мы это обсудим дальше) в случае, если она не в выигрышной позиции и потратит всю меру, то следующее свойство будет выполнено[1]:

> *Для всех n есть хотя бы 2^{n-a_n-d} точек, где мера Боба $\beta(\cdot)$ больше, чем 2^{-n}.*

Заметим, что ходы Алисы не зависят от n, зато Боб должен следить за выигрышем для всех n (то есть указывать точку с достаточно большой мерой для всех n).

Оценим, сколько суммарно меры Боб должен потратить. Выделенное нами свойство гарантирует, что Боб должен потратить хотя бы 2^{-a_n-d}, чтобы Алиса не выиграла для фиксированного n. Сумма таких слагаемых бесконечна (по нашему предположению), а значит в какой-то момент Боб не сможет увеличить свою меру, когда он должен это сделать. Однако есть тонкость: один и тот же ход Боба посчитан много раз (для разных n). Поэтому нам потребуется воспользоваться еще одной технической леммой (верной для любого ряда с неотрицательными членами):

$$\sum_j 2^{-j} \cdot |\{i : \beta(i) \geq 2^{-j}\}| \leqslant 2 \sum_n \beta(n).$$

(Доказательство леммы: Член $\beta(n)$ в правой части дает в левой части вклады 2^{-j} при всех j, для которых $2^{-j} \leq \beta(n)$, и сумма этих вкладов не превышает $2\beta(n)$.)

[1]Формально, Боб обязан отвечать только после того, как суммарная мера Алисы строго превысит 2^{-n+a_n+d}, но это добавит только множитель 2, который мы для простоты будем игнорировать

Чтобы закончить описание выигрышной стратегии, осталось объяснить, как выбрать размер маленького шага Алисы. Мы знаем, что $\sum_n 2^{-a_n-d}$ бесконечна, а значит есть конечная часть с достаточно большой суммой (нам хватит 4). Алиса может использовать в качестве своих шагов 2^{-s}, где s больше, чем $n + a_n + d$ для всех n из этой конечной части.

Литература

[1] Л. Биенвеню, П. Гач, М. Хойруп, К. Рохас, А. Шень, «Алгоритмические тесты и случайность относительно классов мер», *Алгоритмические вопросы алгебры и логики*, Сборник статей. К 80-летию со дня рождения академика Сергея Ивановича Адяна, Тр. МИАН. **274**, МАИК, М.,2011, с.41–102. См. также: `arXiv:1103.1529v2`.

[2] П. Гач, «О симметрии алгоритмической информации.», *Доклады АН СССР.* **218**:6 (1974), с. 1477–1480.

[3] Н. К. Верещагин, В. А. Успенский, А. Шень, «Колмогоровская сложность и алгоритмическая случайность», М.:МЦНМО. — 2013. 576 с.

[4] А. К. Звонкин, Л. А. Левин, «Сложность конечных объектов и обоснование понятий информации и случайности с помощью теории алгоритмов», *Успехи математических наук*, **25**:6(156) (1970), с. 83–124.

[5] А. Н. Колмогоров, «Три подхода к определению понятия “количество информации”», *Проблемы передачи информации*, **1**:1 (1965), с. 3–11.

[6] Л. А. Левин, «Законы сохранения (невозрастания) информации и вопросы обоснования теории вероятностей», *Проблемы передачи информации*, **10**:3 (1974), с. 30–35.

[7] Л. А. Левин, «Некоторые теоремы об алгоритмическом подходе к теории вероятностей и теории информации», *Диссертация на соискание ученой степени кандидата физико-математических наук*, МГУ, Москва. — 1971. — 53 стр.

[8] Л. А. Левин, «О понятии случайной последовательности», *Доклады АН СССР*, **212**:3 (1973), с. 548–550.

[9] Л. А. Левин, «Равномерные тесты случайности», *Доклады АН СССР*, **227**:1 (1976), с. 33–35.

[10] Archimedes, «The Sand Reckoner». In *The Works of Archimedes*, Dover, New York, 1953. (Архимед. Сочинения. Перевод Ю. Н. Веселовского, М.:Физматгиз, 1962.)

[11] L.Bienvenu, A.Shen, «Random Semicomputable Reals Revisited», In: Computation, Physics and Beyond - International Workshop on Theoretical Computer Science, WTCS 2012, Dedicated to Cristian S. Calude on the Occasion of His 60th Birthday. — Springer. — 2012. — **7160** . См. также `http://arxiv.org/pdf/1110.5028v1.pdf`

[12] G.J. Chaitin, «A theory of program size formally identical to information theory», *Journal of the ACM*, **22** (1975), no. 3, p. 329–340.

[13] C. S. Calude, A. Nies, L. Staiger, F. Stephan,«Universal recursively enumerable sets of strings», In: *Developments in Language Theory, 2008*, Lecture Notes in Computer Science, **5257** (2008), p. 170–182.

[14] P. Gács, «Exact expressions for some randomness tests», *Zeitschrift f. Math. Logik und Grundlagen d. Math.*, **26** (1979), p. 385–394.

[15] P. Gács, «On the relation between descriptional complexity and algorithmic probability», *Theoretical Computer Science*, **22** (1983), p. 71–93.

[16] P. Martin-Löf, «The definition of random sequences», *Information and Control*, v. 9 (1966), p. 602–619.

[17] Joseph S. Miller, Liang Yu, «On initial segment complexity and degrees of randomness», *Transaction of the AMS*, **360** (2008), issue 6, p. 3193-3210.

[18] A. A. Muchnik, I. Mezhirov, A. Shen, N. Vereshchagin, «Game interpretation of Kolmogorov complexity», `arxiv:1003.4712`

[19] T. Radó, «On non-computable functions.» *Bell System Technical Journal*, **41**(3), May 1962, 877–884.

[20] C. P. Schnorr, «Process complexity and effective random tests», *Journal of Computer and System Sciences*, 7 (1973), p. 376–388. Preliminary version: *Proc. 4th ACM Symp. on Theory of Computing* (STOC), 1972, p. 168–176.

[21] A. Shen, «Algorithmic Information Theory and Kolmogorov Complexity», *Lecture notes of an introductory course at Uppsala university*, available at `www.it.uu.se/research/publications/reports/2000-034`.

[22] A. Shen, N.K. Vereshchagin, V.A. Uspensky, «Kolmogorov complexity and algorithmic randomness» (english version) — Not published yet. Available on-line: http://www.lirmm.fr/~ashen/kolmbook-eng.pdf

Работы автора по теме диссертации

[23] Mikhail Andreev, «Busy Beavers and Kolmogorov Complexity», Pursuit of the Universal 12th Conference on Computability in Europe, CiE 2016, Paris, France, June 27 — July 1, 2016, Proceedings — 2016 — LNCS, Volume 9709 — p.195-204.

[24] Mikhail Andreev, Ilya Razenshteyn, Alexander Shen, «Not every domain of plain decompressor contains the domain of a prefix-free one», *Theoretical Computer Science*, **412** (Feb. 2011)

[25] Mikhail Andreev, Akim Kumok, «The Sum $2^{KM(x)-K(x)}$ Over All Prefixes x of Some Binary Sequence Can be Infinite», *Theory of Computing Systems*, **58**:3 (2016)