Дискретная математика и математическая логика

Швецова Анна, Ермилов Антон

11 февраля 2017 г.

Содержание

1. Введение в математическую логику

1.1. Пропозициональная логика (исчисление высказываний)

1.1.1. Определения и базовые понятия

***Определение* 1.1.1.**
Алфавит – конечное множество элементов, которые называются символами.
Например, $\Sigma = \{0, 1\}$, где Σ – алфавит

***Определение* 1.1.2.**
Строка (слово) – конечная последовательность символов.
Σ^n – множество всех слов длины n. Более формально: $\Sigma^n = \Sigma \times \Sigma \times \cdots \times \Sigma$, где Σ повторено n раз.
Σ^* – множество всех строк, включая пустую. Более формально, $\Sigma^* = \bigcup\limits_{n \geqslant 0} \Sigma^n$
$\Sigma^0 = \{\Lambda\}$ – пустое слово.

***Определение* 1.1.3.** Булева функция: $\{0, 1\}^n \to \{0, 1\}$

Пример 1. Функция голосования

$$\mathrm{Maj}_n(x_1, \ldots, x_n) = \begin{cases} 1 & x_1 + x_2 + \ldots + x_n \geqslant \frac{n}{2} \\ 0 & \text{иначе} \end{cases}$$

Пример 2. Функция чётности

$$\mathrm{Partity}(x_1, \ldots, x_n) = \begin{cases} 1 & 2 \mid x_1 + x_2 + \ldots + x_n \\ 0 & \text{иначе} \end{cases}$$

***Определение* 1.1.4.**
$\Gamma = \{x_1, x_2, \ldots, x_n\}$ — множество пропозициональных переменных.

***Определение* 1.1.5.**
Пропозициональной называется переменная, вместо которой можно подставить 0 или 1.

***Определение* 1.1.6** (Пропозициональная формула)**.**

1. Пропозициональная переменная
2. Если φ – это пропозициональная формула, то «(φ)», «$\neg\varphi$»– тоже пропозиииональные формулы.
3. Если φ и ψ – пропозициональные формулы, то «$\varphi \vee \psi$», «$\varphi \wedge \psi$», «$\psi \to \varphi$» – тоже пропозициональные формулы.

Множество пропозициональных формул — это минимальное множество строк, которые обладают свойствами 1-3.

***Определение* 1.1.7** (Интерпретация формул)**.**
$I : \Gamma \to \{0, 1\}$

Формула, содержащая в себе n переменных, задаёт некоторую булеву функцию $\{0, 1\}^n \to \{0, 1\}$

1.1.2. Представление в виде КНФ и ДНФ. Понятие эквивыполнимости

***Определение* 1.1.8.**
Литерал – либо переменная, либо её отрицание. $l = x$ или $l = \neg x$

***Определение* 1.1.9.**
Конъюнкт – конъюнкция нескольких литералов. $c = l_1 \wedge l_2 \wedge \ldots \wedge l_k, k \geqslant 0$

***Определение* 1.1.10.**
Дизъюнкт – дизъюнкция некольких литералов. $d = l_1 \vee l_2 \vee \ldots \vee l_k, k \geqslant 0$

***Определение* 1.1.11** (Дизъюнктивная нормальная форма)**.**
ДНФ – дизъюнкция нескольких конъюнктов. $c_1 \vee c_2 \vee \ldots \vee c_k, k \geqslant 0$

***Определение* 1.1.12** (Конъюнктивная нормальная форма)**.**
КНФ – конъюнкция нескольких дизъюнктов. $d_1 \wedge d_2 \wedge \ldots \wedge d_k, k \geqslant 0$

Теорема 1.1.1.
Любая булева функция представима как в виде ДНФ, так и в КНФ.

Доказательство.
Рассмотрим пары $x \in \{0,1\}^n, f(x)$ Для каждого x такого, что $f(x) = 1$, запишем набор значений x в виде конъюнкта. Тогда дизъюнкция таких конъюнктов будет давать ДНФ т.к. каждый набор, для которого $f(x) = 1$, выполнится по своему конъюнкту, а те x, для которых $f(x) = 0$, формула не выполнится, т.к. конъюнкты заданы строго.

Аналогично для КНФ возьмем каждый $x \in \{0,1\}^n : f(x) = 0$, запишем все инвертированные значения x в виде дизъюнкта. Из получившихся дизъюнктов получим КНФ. □

***Определение* 1.1.13.**
Две функции эквивалентны, если задают одинаковую булеву функцию.

Пример эквивалентности.

1. $x \to y \sim \neg x \vee y$
2. $\neg(x \vee y) \sim \neg x \wedge \neg y$ (Правило де Моргана)
3. $\neg(x \wedge y) \sim \neg x \vee \neg y$ (Правило де Моргана)
4. $x \wedge (y \vee x) \sim x \wedge y \vee x \wedge z$

Алгоритм приведения формулы в ДНФ.

1. Избавляемся от импликаций
2. Проносим все отрицания к переменным
3. Раскрываем по дистрибутивности

Алгоритм приведения формулы в КНФ.

1. Навесить отрицание
2. Привести в ДНФ
3. Ещё раз отрицание. По правилу де Моргана получим конъюнкцию.

***Определение* 1.1.14.**
Формула выполнима, если ∃ интерпретация, в которой её значение – истина. В противном случае формула называется невыполнимой или противоречивой.

***Определение* 1.1.15.**
Тавтология – истина при всех интерпретациях (её отрицание невыполнимо).

Пример. $x \to x$

Замечание.
В ДНФ очень просто проверить формулу на выполнимость: достаточно просто посмотреть, существует ли такой конъюнкт, что в нём одновременно не присутствуют отрицание и утверждение одной и той же переменной.

***Определение* 1.1.16.**
Две формулы φ и ψ называются эквивыполнимыми, если φ выполнима тогда и только тогда, когда ψ выполнима.

1.1.3. Теорема о приведении формулы в эквивыполнимый ей КНФ

Теорема 1.1.2.
Существует эффективный алгоритм, позволяющий приводить формулу в эквивыполнимый ей КНФ.

Доказательство.
Доказательство легко следует из построения. Чтобы показать основной принцип построения, приведем пример:

Пример.

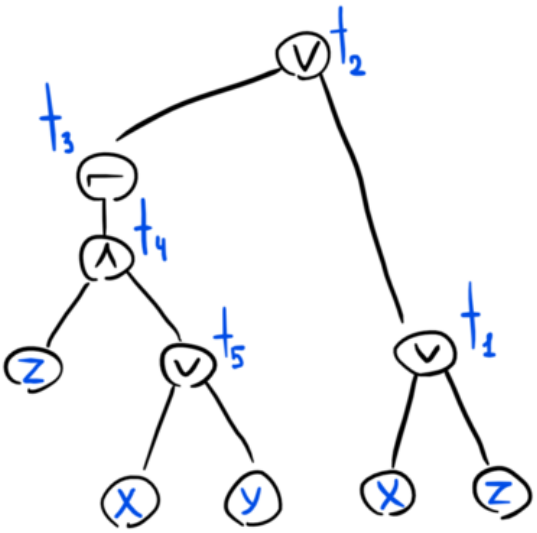


Рис. 1: Дерево для формулы $\neg(z \wedge (x \vee y)) \vee (x \vee z)$

Обозначим каждую вершину не-букву за дополнительную переменную t_i. Заведем систему

уравнений:

$$\begin{cases} t_2 = t_1 \vee t_3 \\ t_1 = x \vee z \\ t_3 = \neg t_4 \\ t_4 = z \wedge t_5 \\ t_5 = x \vee y \\ t_2 = 1 \end{cases}$$

Очевидно, что если полученная система имеет решение, то и исходная формула имеет решение и наоборот. Заметим, что «$x = y$» $\sim$ «$(x \wedge y) \vee (\neg x \wedge \neg y)$», тогда, заменив знаки «=» на эквивалентную им формулу и знак системы на конъюнкцию между выражениями системы, получим формулу в КНФ эквивыполнимую исходной. □

1.1.4. Метод резолюций

Утверждение 1.1.3.

Пусть есть дизъюнкты d_1,d_2,…,d_m, тогда, если есть 2 дизъюнкта, конфликтующие ровно по одной переменной x, из них может быть выделена ещё одна дизъюнкция: из посылок $x \vee C$ и $\neg x \vee B$ получается заключение $C \vee B$. В дальнейшем будет использоваться следующая форма записи:

$$\frac{x \vee C,\ \neg x \vee B}{C \vee B}$$

Теперь, если есть интерпретация, выполняющая обе посылки, то она будет выполнять и заключение.

Доказательство.

Действительно, посмотрим на то выражение, которое у нас было: $(x \vee C) \wedge (\neg x \vee B)$ Пусть $x = 1$, тогда первый дизъюнкт выполняется сразу, а во втором B – истина т.к. $\neg x = 0$.
Пусть $x = 0$, тогда в первом дизъюнкте истина – C, а во втором выполняется $\neg x$.
Мы хотели доказать, что при наличии интерпретации, выполняющей выражение выше, выражение $C \vee B$ тоже выполнится. Из сказанного выше заметим, что это действительно так. □

***Определение* 1.1.17.**

Резолюционное опровержение – это последовательность дизъюнктов d_1,d_2,…,d_n,□, где d_i – это либо дизъюнкт формулы, которую мы опровергаем, либо дизъюнкт, полученный по правилу резолюции, а □ – пустой дизъюнкт (невыполним).

Теорема 1.1.4 (о корректности метода резолюций)**.**

Если у формулы φ в КНФ есть резолюционное опровержение, то формула невыполнима.

Доказательство.

Пусть φ выполнима, I – выполняющая интерпретация. Тогда по допущенному утверждению, I выполняет все дизъюнкты из опровержения, но пустой дизъюнкт выполнить невозможно. Противоречие. □

Теорема 1.1.5 (о полноте резолюций)**.**

Если формула невыполнима, то для неё существует резолюционное опровержение.

Доказательство. Индукция по числу переменных n.

База:

$$n = 1,\ \frac{x,\ \neg x}{\square}$$

Переход $n \to n+1$:
Пусть теперь есть некоторая формула $\varphi(x_1,x_2,\ldots,x_{n+1})$ в КНФ. Разделим все её дизъюнкты на 3 группы:

1. S_1 – содержащие x_1
2. S_2 – содержащие $\neg x_1$
3. S_3 – не содержащие ни x_1, ни $\neg x_1$.

Рассмотрим $\varphi[x_1 := 0]$. Все дизъюнкты S_2 выполнились, все дизъюнкты в S_1 избавляются от x_1, S_3 никак не изменились. По индукционному предположению у $\varphi[x_1 := 0]$ есть резолюционное опровержение, выпишем его. Заметим, что в получившемся опровержении нет конфликтов по переменной x_1 и вернем $x_1 = 0$ обратно во все дизъюнкты опровержения, принадлежащие S_2 (и результирующие от них). Теперь конечный дизъюнкт такого опровержения либо $\square$, либо x_1. Если $\square$, то утверждение доказано. Если же x_1, то продолжим дальше.

Рассмотрим теперь $\varphi[x_1 := 1]$. Аналогично случаю выше, S_2 выполнятся, S_1 избавятся от $\neg x_1$, S_3 останется без изменений. По индукции, у получившейся формулы есть опровержение. Выпишем его, добавим в необходимые дизъюнкты $\neg x_1$. Получим конечным дизъюнктом либо $\square$, либо $\neg x_1$. Если $\square$, то утверждение доказано, если $\neg x_1$, то из случая выше наше опровержение также содержит x_1. $\frac{x_1,\ \neg x_1}{\square}$. $\square$

Следствие Алгоритм для проверки выполнимости формулы в 2-КНФ.

Формулой в 2-КНФ будем называть формулу в КНФ, в каждом дизъюнкте которой не более 2-х литералов.

Алгоритм заключается в следующем:

1. Пока из текущего набора дизъюнктов можно получить новые:
   (a) Перебрать все пары дизъюнктов.
   (b) Применить резолюцию к тем парам, с которыми можно так делать.
   (c) Если в результате появился какой-то дизъюнкт, которого до этого не было в нашем опровержении, приписать его в конец
   (d) Если этим дизъюнктом оказался $\square$, то данная нам формула невыполнима
2. Если в процессе выполнения операций выше мы не обнаружили, что формула невыполнима, значит, она выполнима.

Найдем теперь оценку на количество дизъюнктов в опровержении. Заметим, что при применении резолюции, мы в худшем случае создадим новый дизъюнкт из не более чем 2-х литералов. Значит, оценка сводится к количеству возможных дизъюнктов из не более чем 2-х литералов. Таких будет $\leqslant 2n * 2n$, где n – количество переменных. (Мы пытаемся таким образом поставить каждую переменную в пару с каждой; множители 2 появились из-за того, что каждая переменная может быть представлена как своим утверждением, так и отрицанием).

1.1.5. Дерево поиска противоречий

***Определение* 1.1.18.**
Пусть имеется некоторая формула φ. Тогда дерево противоречий для неё – это корневое бинарное дерево, каждая вершина, кроме листьев, которого помечена переменной. Тогда от вершины, помеченной переменной x, будет исходить 2 ребра, в одном из которых $x = 0$, а в другом $x = 1$. Каждый же лист такого дерева помечен дизъюнктом функции φ, который опровергается подстановкой переменных пройденных на пути от корня до листа. Таким образом, проход по ребру такого дерева эквивалентен подстановке переменной, а каждый лист такого дерева содержит в себе дизъюнкт, который такая подстановка опровергает.

Пример 1. $\varphi = y \wedge (x \vee \neg y \vee z) \wedge (x \vee \neg z) \wedge (\neg x \vee t) \wedge \neg t$ TODO Картинка

Пример 2. TODO Картинка

Теорема 1.1.6.
По любому дереву противоречий формулы φ в КНФ можно построить резолюционное опровержение, размер которого не превосходит числа вершин в дереве противоречий.

Доказательство.
Будем считать, что во внутренних вершинах дерева никакой дизъюнкции не опровергается (то есть любая вершина, для которой уже какой-то дизъюнкт не выполнился – лист). Тогда индукция по размеру дерева:

База: $\square$, дерево из одной вершины. Резолюционное опровержение очевидно. Переход: $\leqslant S$ вершин $\to S + 1$.
Рассмотрим текущее дерево из $S + 1$ вершины. Найдем в нем 2 самых глубоких листа с общим родителем. Раз это листы, то детей у них не будет, значит, в них опровергаются какие-то дизъюнкты. Пусть для одного листа опровергаемый дизъюнкт – $x \vee A$, а для другого – $\neg x \vee B$, где x – переменная, записанная в их родителе. Тогда заменим листья на их резолюцию $A \vee B$, уменьшив таким образом количество вершин в графе на 2. По предположению индукции, для дерева с количеством вершин $\leqslant S$, все условия выполняются, то есть для полученного дерева существует резолюционное опровержение длины $\leqslant S$. Теперь заметим, что мы только что сопоставили каждой из выбранных вершин-листьев по дизъюнкту, значит, длина опровержения для исходного дерева $\leqslant S + 1$ $\square$

Пример. TODO Картинка

1.2. Более сложные способы задания булевых функций

Пусть есть некоторая функция $f : \{0, 1\}^n \to \{0, 1\}$. Существует множество способов её записать. Перечислим некоторые из них в порядке невозрастания длины:

1. Таблица истинностей
2. Деревья решений
3. Формулы в ДНФ и КНФ
4. Пропозициональные формулы
5. Ветвящаяся программа (branching program или ordered binary decision diagrams)
6. Схемы из функциональных элементов

1.2.1. Дерево решений

***Определение* 1.2.1.**

TODO Картинка Дерево решений – это бинарное дерево, каждая внутренняя вершина которого помечена переменной. Из каждой внутренней вершины исходит по 2 ребра, одно из которых помечено 0, а другое – 1, эти пометки соответствуют значениям переменной в вершине, которые мы присваиваем ей во время спуска по дереву. (Да, дерево решений действительно очень похоже на дерево поиска противоречий). Листья же такого дерева помечены значениями из множества $\{0, 1\}$.

Тогда, для того, чтобы вычислить значение функции на таком дереве, имея набор аргументов $\{x_1,x_2,\ldots,x_n\}$, необходимо спускаться вниз по ребрам, соответствующим значениям аргументов, до тех пор, пока не будет достигнут лист. Значение в листе и есть значение функции.

Свойства.

1. Глубина дерева d – максимальное расстояние по ребрам от корня до листа.
2. Минимальная по всем возможным деревьям глубина дерева – запросовая сложность этой функции (query complexity).
3. Размер дерева – количество вершин в нем.

Пример 1.

$f = x_1+x_2+\ldots+x_n$`mod 2`. У такой функции дерево решений – полное бинарное дерево глубины n. Любое её дерево решений велико (точнее, максимально) как по размеру, так и по глубине.

Пример 2.

$f = x_1\wedge x_2\wedge\ldots\wedge x_n$. В отличие от предыдущего примера, размер дерева решений для такой функции в разы меньше (попробуйте посчитать, чему он равен), но глубина всё ещё n.
Давайте поймём, что глубина действительно n. В таких случаях можно применить метод противника, который заключается в следующем.

Пусть есть набор переменных $\{x_1,x_2,\ldots,x_n\}$ и функция f. Два человека играют в игру. Один из них пытается отгадать значение заданной функции на наборе переменных и в свой ход может узнать значение любой из них, но только одной. Его задача – минимизировать количество запросов. Второй же человек пытается отвечать на запросы первого игрока таким образом, чтобы тот как можно дольше не знал ответа. То есть в каждый момент времени второй игрок придумывает максимально выгодное для него значение переменной, которую у него запросили. Максимальное количество запросов, которое в такой игре придется сделать первому игроку, и есть глубина дерева решений.

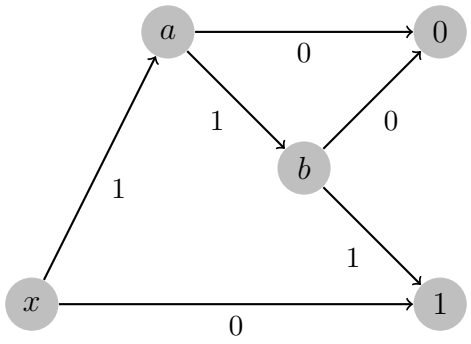
Посмотрим теперь, что будет происходить, если наша функция – AND (Пример 2). Тогда второму игроку выгоднее всего каждый раз отвечать «1», а когда неизвестная переменная останется всего одна, выбирать любое значение. Тогда значение такой функции можно будет узнать только узнав значения всех переменных. Отсюда и получаем, что глубина дерева решений не меньше n.

1.2.2. Ветвящаяся программа

***Определение* 1.2.2.**

Ветвящаяся программа представляет собой ориентированный ациклический граф (directed acyclic graph). В таких графах всегда есть вершины, в которые нет ребер (будем называть их входами) и вершины, из которых не исходит ребер (будем называть их выходами). Таким образом, ветвящаяся программа – это орграф без циклов с 1 входом и 2-мя выходами. При этом

Рис. 2: Пример ветвящейся программы, построенной по формуле $\neg x \vee (a \wedge b)$



один выход помечен «1», а другой – «0». Каждая вершина кроме выходов помечена переменной, из каждой вершины исходит ровно 2 ребра, проход по одному из которых эквивалентен присвоению переменной в исходной вершине значения нуля, а другой – единицы. (Можно сказать, что ветвящаяся программа – это просто такое сжатое дерево решений).

***Определение* 1.2.3.**
Размер ветвящейся программы – это число вершин, не являющихся выходами.

Теорема 1.2.1.
Для любой формулы размера S существует ветвящаяся программа размера $\leqslant S$, которая считает ту же функцию.

Доказательство.
Доказательство по индукции.

База: формула из одного элемента. TODO Картинка
Переход: пусть есть некоторые формулы φ и ψ, которым соответствуют ветвящиеся схемы A и B. Тогда попробуем получить из них всевозможные комбинации. Пусть нам нужна схема для формулы $\neg\varphi$. Чтобы её получить, достаточно всего лишь поменять значения на выходах схемы A. На примере операций $\vee$ и $\wedge$ покажем, как строить схемы $\psi \vee \varphi$, $\psi \wedge \varphi$, чтобы соблюдались условия теоремы: см. Рис. 2. □

1.2.3. Схемы из функциональных элементов

***Определение* 1.2.4.**
Базис представляет собой конечный набор булевых функций, через которые можно выразить любую булеву функцию. Таким образом, любая булева функция будет представлена композицией функций базиса. Иначе говоря, элемент базиса представляет собой некоторую функцию $f : \{0,1\}^n \to \{0,1\}$, входы которой пронумерованы (пронумерованы $\Leftrightarrow$ порядок аргументов имеет значение).

***Определение* 1.2.5.**
Схема из функциональных элементов базиса B – это ориентированный ациклический граф, входы которого помечены переменными. Каждая вершина, кроме входов, помечена каким-то элементом базиса. Если вершина помечена функцией арности k, то в неё входит ровно k ребер и они пронумерованы. TODO Картинка

Свойства.

1. Глубина схемы – максимальная длина пути от вершины-входа до вершины-выхода
2. Существует разбиение на уровни (уровнями будем называть множества вершин, находящихся на одной глубине). Причем, схемы, расположенные на одном уровне, можно выполнять параллельно.
3. $Size_B(f)$ – размер кратчайшей схемы базиса B, вычисляющей функцию f, где размер схемы – это количество вершин.

Теорема 1.2.2 (об эквивалентности базисов)**.**
Пусть A, B – 2 базиса. Тогда $\exists c : \forall f$ – булева функция $Size_A(f) \leqslant c \cdot Size_B(f)$. Иначе говоря, утверждается, что минимальные размеры схем, вычисляющих одну и ту же функцию, но построенные на разных базисах, будут равны с точностью до константы.

Доказательство.
Пусть $A = \{g_1^{(k_1)}, g_2^{(k_2)}, \ldots, g_n^{(k_n)}\}$, где $g_i^{(k_i)}$ – k_i-арная функция. Тогда, так как для любого базиса верно, что с помощью его элементов можно получить любую булеву функцию, то и с помощью элементов B можно получить любую $g_i^{(k_i)}$. Пусть $C_1, C_2, \ldots, C_n$ – схемы, состоящие из элементов базиса B и $C_i \sim g_i^{(k_i)}$. Тогда возьмем схему базиса A любой функции f и заменим в ней все элементы базиса A эквивалентными схемами C. Таким образом, мы научились строить схему по нужному нам базису, имея готовую схему по любому другому базису. Заметим теперь, что поскольку исходные наборы базисов константны, и любая функция базиса представима в другом базисе с использованием конечного числа вершин, то мы изменили размер схемы в худшем случае в $\max(|C_i|)$ раз. □

Теорема 1.2.3 (о размере кратчайшей схемы)**.**
$\forall$ базиса B $\exists c = \text{const} > 1 : \forall\ n\ \exists\ f : \{0,1\}^n \to \{0,1\} : Size_B(f) \geqslant \frac{2^n}{cn}$. Более того, такое $\frac{2^n}{cn}$ часто можно даже явно предъявить. (Если отбросить все числа, можно заметить, что это утверждение о том, что не существует базиса такого, что размер минимальной схемы для любой функции на этом базисе можно ограничить сверху каким-то числом)

Доказательство.
Для начала поймем, что раз в зависимости от базиса размер схемы меняется всего лишь на константу, то нам достаточно доказать утверждение для какого-то конкретного базиса. Тогда пусть $B = \{\downarrow_2\}$ ($\downarrow$ – стрелка Пирса, $x \downarrow y = \neg(x \vee y)$).

Научимся теперь представлять схему в виде битовой строки. Пусть в нашей схеме n переменных и S вершин. Тогда если пронумеровать все вершины, то каждый номер будет представим в виде битовой строчки длины не более $\log S$. Тогда будем кодировать все вершины, кроме переменных (входов), парой номеров вершин, из которых в текущую вершину ведет ребро (парой т.к. стрелка Пирса имеет арность 2). Таким образом, схему из S вершин можно представить в виде битовой строчки длины не более $dS \log S$, где d – какая-то константа (где-то нам придется умножить на 2, так как каждую вершину мы кодируем парой, где-то ещё что-то, но по смыслу это просто какая-то константа).

Посчитаем теперь число булевых функций такого вида: $f : \{0,1\}^n \to \{0,1\}$ (помним, у нас всего n переменных на входе). Таких функций будет всего 2^{2^n} (2^n возможных бинарных строчек длины n, а ещё каждую такую строчку нам нужно сопоставить 1 или 0 и посчитать количество таких сопоставлений; получается 2^{2^n}).

Теперь посчитаем количество булевых функций размера $S \leqslant \frac{2^n}{cn}$ (по вершинам), $(d < c)$. Из написанного выше, каждую такую функцию можно представить в виде бин. строки длины $d\frac{2^n}{cn} \log \frac{2^n}{cn} \leqslant d\frac{2^n n}{cn} = d\frac{2^n}{c}$. Отсюда мы получим, что число таких функций, чьи бин. строки имеют длину $\leqslant d\frac{2^n}{c}$, не более чем $2^{d\frac{2^n}{c}} < 2^{2^n}$. Значит, функция такая, что её размер больше заявленного числа, всегда существует. □

1.3. Предикатная логика

1.3.1. Предикатные формулы (логика первого порядка)

***Определение* 1.3.1.**
Будем обозначать множество всех натуральных чисел с нулем как $\mathcal{N}$

***Определение* 1.3.2.**
k-местный предикат на множестве M:

$$p : M^k \to \{0, 1\}, k \in \mathcal{N}$$

k-местная функция на множестве M:

$$f : M^k \to M, k \in \mathcal{N}$$

***Определение* 1.3.3.** Сигнатура – совокупность некоторых предикатов P и функций F. В таком случае пишут, что (P, F) – сигнатура, где $P = \{p_1^{(k_1)}, p_2^{(k_2)} \ldots\}$ – множество предикатных символов, а $F = \{f_1^{(l_1)}, f_2^{(l_2)} \ldots\}$ – множество функциональных символов.

Пример. $P = \{=_2, <_2\}, F = \{+_2, \times_2, 0_0, 1_0\}, (P, F)$ – сигнатура

***Определение* 1.3.4** (Терм)**.**
$\Pi = \{x_1, x_2, \ldots, x_n\}$ – множество предметных переменных, тогда:

1. Предметная переменная – это терм
2. $f^k \in F, t_1, t_2, \ldots, t_k$ – термы, тогда $f(t_1, t_2, \ldots, t_k)$ тоже терм
3. Множество термов – это минимальное множество строк, удовлетворяющее свойствам 1-2.

Пример. Пусть $F = \{f^{(1)}, g^{(2)}, c^{(0)}\}$ – некоторое множество функций, тогда x; $f(x)$; $g(x, f(x))$; $g(c, f(x))$; $g(g(c, c), f(c))$ – термы.

***Определение* 1.3.5** (Элементарная/атомарная формула)**.**
Если $p^{(k)} \in P$, $t_1, t_2, \ldots, t_k$ – термы, то $p(t_1, t_2, \ldots, t_k)$ – атомарная формула.

Пример. Пусть $P = \{p^{(2)}\}$, тогда $p(f(x), c)$; $p(g(c, f(x)), g(x, x))$ – атомарные формулы.

***Определение* 1.3.6** (Предикатные формулы/формулы первого порядка)**.**

1. атомарная формула – предикатная формула
2. φ, ψ – предикатные формулы, тогда (φ), $\neg\varphi$, $\psi \vee \varphi$, $\psi \wedge \varphi$, $\psi \to \varphi$ ... – тоже предикатные формулы.
3. φ – формула, $x \in \Pi$, тогда $\exists x\varphi$ и $\forall x\varphi$ – тоже формулы.
4. Множество формул – это минимальное множество строк, удовлетворяющих свойствам 1-3.

Пример. $p(f(x), g(f(c), x)) \to p(f(c), c) \to \exists y \forall z (p(y, y) \vee p(z, z))$

***Определение* 1.3.7** (свободные и связанные переменные)**.**
Связанная переменная – это переменная, которая находится в области действия квантора. Свободной переменной называется несвязанная переменная. Например, в выражении

$$[\forall x \ \exists y \ p(x, y)] \vee q(f(x))$$

переменная x внутри квадратных скобок связана, а вне – нет. Переменная y является связанной.

***Определение* 1.3.8.**
Интерпретация I сигнатуры (P, F)

1. Носитель интерпретации – множество M
2. $\forall p^{(k)} \in P$ сопоставляется $M^k \to \{0, 1\}$
3. $\forall f^{(l)} \in F$ сопоставляется $M^l \to M$

Пример. (Великая теорема Ферма для степени 3)

$F = \{+, \times, 0, 1, 2\}$

$P = \{=, <\}$

$\forall x\ \forall y\ \forall\ z\ (x^3 + y^3 = z^3) \to [(x = 0) \vee (y = 0) \vee (z = 0) \vee (x < 0) \vee (y < 0) \vee (z < 0)]$

***Определение* 1.3.9.** (Оценка)
$\alpha : \Pi \to M$ – оценка.

1. $x_i \in \Pi$ – переменная, тогда $\alpha(x_i)$ – её значение.
2. $f^{(k)}(t_1,t_2,\ldots,t_k) \in F$ – некоторая функция, тогда $\alpha(f)$ – результат после подстановки значений всех $\alpha(t_i)$.
3. $p^{(l)}(t_1,t_2,\ldots,t_l) \in P$ – некоторый предикат, тогда $\alpha(p)$ – результат после подстановки значений всех $\alpha(t_i)$

Тогда если φ – некоторая предикатная формула, то будем обозначать $[\varphi]_{I,\alpha}$ её значение в интерпретации I при оценке α Тогда:

$$[\forall x\ \varphi]_{I,\alpha} = \bigwedge_{a \in M} [\varphi]_{I,\alpha^{[x \leftarrow a]}}$$

$$[\exists x\ \varphi]_{I,\alpha} = \bigvee_{a \in M} [\varphi]_{I,\alpha^{[x \leftarrow a]}}$$

Пусть I – интерпретация с носителем M. Если формула φ содержит k свободных переменных, то φ задаёт некоторое отображение(предикат) $M^k \to \{0, 1\}$

***Определение* 1.3.10.**
Пусть $p : M^k \to \{0, 1\}$, (P, F) – сигнатура, I – интерпретация относительно множества M, тогда будем говорить, что p выразим, если его можно задать формулой сигнатуры (P, F).

1.3.2. Арифметика

***Определение* 1.3.11** («Арифметика»)**.**

$P = \{=\}$

$F = \{+, \times\}$

Носитель интерпретации – $\mathcal{N} = \{0, 1, 2, \ldots\}$

Примеры.

1. $x = 0 \qquad x + x = x$

2. $x > 0$ $\quad \neg(x = 0)$
3. $x = 1$ $\quad (x > 0) \wedge (x \times x = x)$
4. $x \geqslant y$ $\quad \exists z(y + z = x)$
5. $x > y$ $\quad (x \geqslant y) \wedge \neg(x = y)$
6. $x = 42$ $\quad \exists y((y = 1) \wedge (x = \underbrace{y + \ldots + y}_{\text{42 слагаемых}}))$
7. $x \vdots y$ $\quad \exists z(x = y \times z)$
8. x – простое число $\quad (x > 1) \wedge (\forall z \forall w((x = z \times w) \to ((z = 1) \vee (w = 1))))$
9. x – степень числа 2 $\quad \forall z(((x \vdots z) \wedge (z - \text{простое})) \to (z = 2))$
10. x – степень числа 4 $\quad \exists y((y - \text{степень } 2) \wedge (y \times y = x))$

***Определение* 1.3.12.**

Поставим любому числу x в соответствие строку, которую можно получить из двоичного представления числа $x+1$, из которого вычеркнули первую единицу. Тогда числу 0 соответствует пустая строка, 1 – строка «0», 8 – «001».

Будем обозначать такую строку, соответствующую числу x, как $\widetilde{x}$.

Продолжим серию примеров.

Примеры.

1. $\widetilde{x}$ состоит из одних нулей

   На самом деле, это то же самое, что и проверить, что $x + 1$ – степень двойки. Решение: $\exists z[(z = 1) \wedge ((x + z) - \text{степень двойки})]$

2. $|\widetilde{x}| = |\widetilde{y}|$, $\widetilde{x}$ и $\widetilde{y}$ имеют одинаковую длину.

   Это выполняется тогда и только тогда, когда $2^t \leqslant x + 1, y + 1 < 2^{t+1}$.
   Решение: $\exists z[(z - \text{степень двойки}) \wedge (x + 1 \geqslant z) \wedge (y + 1 \geqslant z) \wedge (x + 1 < z + z) \wedge (y + 1 < z + z)]$

3. $\widetilde{x} = \widetilde{y}\widetilde{z}$, то есть $\widetilde{x}$ является конкатенацией (склейкой) строк $\widetilde{y}$ и $\widetilde{z}$.
   Пусть $t = |\widetilde{z}|$. Утверждение другими словами: $x + 1 = (y + 1) \cdot 2^t + (z + 1) - 2^t = y \cdot 2^t + z + 1$.
   Объяснение: $(y + 1) \cdot 2^t$ – сдвинули $y + 1$ влево на нужную позицию, $(z + 1) - 2^t$ – вписали $z + 1$ в младшие биты, вычев старшую единицу числа $z + 1$.
   Решение: $\exists s[(\widetilde{s} - \text{состоит из нулей}) \wedge (|\widetilde{s}| = |\widetilde{z}|) \wedge ((x + 1) = y(s + 1) + (z + 1))]$, где s будет подобрано так, что $s + 1 = 2^t$.

4. $|\widetilde{x}| > |\widetilde{y}|$
   Решение: $\neg(|\widetilde{x}| = |\widetilde{y}|) \wedge (x > y)$.

5. $\widetilde{x}$ является началом $\widetilde{y}$.
   Решение: $\exists z(\widetilde{y} = \widetilde{x}\widetilde{z})$.

6. $\widetilde{x}$ является подстрокой $\widetilde{y}$.
   Решение: $\exists z[(\widetilde{z} - \text{начало } \widetilde{y}) \wedge (\widetilde{x} - \text{конец } \widetilde{z})]$.

7. Существует выразимый предикат $S(x, a, b)$ (β-функция Гёделя):

   (a) $\forall a, b \in \mathcal{N}\ S_{a,b} = \{x \in \mathcal{N} | S(a, b, x) = 1\}$ – конечное множество.

(b) Любое конечное множество можно задать в таком виде.

Иначе говоря, для любого конечного множества $X \subseteq \mathcal{N}$ $\exists a, b \in \mathcal{N} : S_{a,b} = \{x \in \mathcal{N} \mid S(a, b, x) = 1\} = X$

Доказательство.
Явно покажем такой предикат.
$S(x, a, b) = (\widetilde{a}\widetilde{x}\widetilde{a}$ является подстрокой $\widetilde{b}) \wedge (|\widetilde{a}| > |\widetilde{x}|)$ Докажем пункты выше:

(a) Очевидно, так как $\widetilde{b}$ – строка конечной длины.

(b) Пусть $X = \{x_1, x_2, \ldots, x_n\}$, $m = \max_i |\widetilde{x_i}|$. Тогда возьмем в качестве $\widetilde{a}$ строку «$10\overset{\text{нули}}{\ldots}01$», причем $|\widetilde{a}| > m + 1$. Таким образом, мы получим, что в $\widetilde{a}$ не менее m подряд идущих нулей. Построим строку $\widetilde{b}$ таким образом: $\widetilde{b} = \widetilde{a}\widetilde{x}_1\widetilde{a}\widetilde{x}_2 \ldots \widetilde{a}\widetilde{x}_n\widetilde{a}$. Отсюда сразу очевидно, что все $x \in X$ действительно будут удовлетворять условию, но почему же не будет других x? Пусть есть какая-то строчка s, которая соответствует какому-то элементу не из множества X. Тогда нужно доказать, что подстроки $\widetilde{a}s\widetilde{a}$ для любого такого s в строке $\widetilde{b}$ не окажется. Рассмотрим случаи, где такая строка могла бы начаться.

i. Пусть она началась с первого символа некоторого $\widetilde{a}$. Тогда первые $|\widetilde{a}|$ символов совпали. Далее нужно сравнить некоторое $\widetilde{x}$ и s. Нам известно, что они не равны друг другу. Сравним их. Если произошел конфликт по некоторому символу, то этот случай очевидно не выполняется. Если конфликта не было, то либо s префикс $\widetilde{x}$, либо наоборот. Случаи симметричны, поэтому пусть s – префикс $\widetilde{x}$. Тогда посмотрим на первый символ $\widetilde{a}$(он равен 1), которая стоит сразу за $\widetilde{x}$ и увидим, что поскольку $|\widetilde{a}| > |\widetilde{x}| + 1$, то на соответствующей позиции в s будет нолик.

ii. Если мы начали где-то из середины $\widetilde{a}$, то очевидно $0 \neq 1$.

iii. Если мы начали из последнего бита $\widetilde{a}$, то происходит поломка аналогично первому пункту.

iv. Если из любой позиции $\widetilde{x}$, то снова поломается первый бит следующего $\widetilde{a}$

□

8. $x = 6^n$. Решение:

$$\exists a, b[S_{\langle n,x\rangle,a,b} \wedge \forall m \forall y(S_{\langle m,y\rangle,a,b} \to ((m = 0 \wedge y = 1) \vee \exists z[(y = 6z) \wedge S_{\langle m-1,z\rangle,a,b}]))]$$

Великая и ужасная строчка, казалось бы. Разбираемся. На самом деле, мы такой формулой хотим утверждать, что если $x = 6^n$, то существуют строчки a и b, которые закодируют всё по схеме которую мы узнали чуть ранее так, что...

(a) Строчка b содержит в себе какие-то степени 6-ки.

(b) Элементы хранятся парами «степень, значение» (Пару n, x можно кодировать как $(n + x)^2 + x$)

(c) Пара n, x входит в такую строчку.

Мы немного считерили, когда сказали, что хотим кодировать пары, но в среднем должно быть понятно, как с этим быть. Если же теперь поговорить немного про формулу по частям, то $S_{\langle n,x\rangle,a,b}$ говорит нам о том, что n, x входит в нужное множество. Но тогда каким должно быть множество? Таким, что из того, что какой-либо элемент в нём лежит: $\forall m \forall y(S_{\langle m,y\rangle,a,b}$ следует, что он – степень 6-ки : $\to ((m = 0 \wedge y = 1) \vee \exists z[(y = 6z) \wedge S_{\langle m-1,z\rangle,a,b}]))$.

9. x – какая-то степень 6-ки. На самом деле, это проще. А именно:

$$\exists a, b[S_{x,a,b} \wedge \forall y(S_{y,a,b} \to (y = 1 \vee \exists z[(y = 6z) \wedge S_{z,a,b}]))]$$

1.3.3. Невыразимость предикатов: метод автоморфизмов

***Определение* 1.3.13.**
I – интерпретация сигнатуры F, P с носителем M.
$\alpha : M \to M$ называется автоморфизмом, если:

1. α – биекция
2. $\forall p^{(k)} \in P \forall x_1,x_2,\ldots,x_k \in M\ p^{(k)}(\alpha(x_1), \alpha(x_2), \ldots, \alpha(x_k)) = p^{(k)}(x_1,x_2,\ldots,x_k)$. Иначе говоря, все предикаты устойчивы относительно автоморфизма.
3. $\forall f \in F \forall x_1,x_2,\ldots,x_n \in M\ f(\alpha(x_1), \alpha(x_2), \ldots, \alpha(x_n)) = \alpha(f(x_1,x_2,\ldots,x_n))$ Иначе говоря, функция устойчива относительно автоморфизма.

Теорема 1.3.1 (метод автоморфизмов)**.**
α – автоморфизм интерпретации I. Тогда если k-местный предикат $p : M^k \to \{0, 1\}$ выразим в I, то p устойчив относительно α.

Соответственно, если такой предикат неустойчив относительно α, то p невыразим

Доказательство.
Сначала разберемся, что мы хотим доказать. У нас есть некоторый автоморфизм α и новый предикат $p(x_1,x_2,\ldots,x_n)$. Нужно доказать, что $p(x_1,x_2,\ldots,x_n) = p(\alpha(x_1), \alpha(x_2), \ldots, \alpha(x_n))$ (собственно, по определению устойчивости).

Раз уж p выразим, то он записывается с помощью предикатов из сигнатуры. Для них нам уже известно, что они устойчивы относительно автоморфизма (по определению автоморфизма). Поймём и для p по кусочкам, что он устойчив:

1. Любой терм устойчив по определению
2. Любой предикат устойчив по определению
3. p – комбинация каких-то предикатов из P. Заметим, что автоморфизм никак не влияет на предикаты, поэтому любая комбинация предикатов $\wedge, \vee, \to$ и т.д. не будет влиять на наш предикат
4. Осталось разобраться со случаями $\forall x\varphi$ и $\exists x\varphi$. Скажем, что φ соответствует некоторый предикат $g(x_1, x_2, \ldots, x_n)$. Тогда $\forall x\varphi$ соответствует $\forall y g(y, x_2, \ldots, x_n)$. Значит, мы хотим проверить $\forall y g(\alpha(y), \alpha(x_2), \ldots, \alpha(x_n))$. Но заметим, что квантор $\forall$ пробежится по всем переменным, а так как α – биекция, то любой квантор в итоге пройдет по любым переменным, как и было. Значит, и такие действия устойчивы относительно автоморфизма.

□

2. Множества и их порядки

2.1. Конечные множества

***Определение* 2.1.1.**

X — конечное множество, если $\exists$ биекция $\alpha : X \to [n] = \{1, 2, ..., n\}$, $|X| = n$

Число различных подмножеств n-элементного множества равно 2^n (то же, что число отображений $X \to \{0, 1\}$).

Если X, Y — два конечных множества, то количество отображений $X \to Y$ равно $|Y|^{|X|}$

$2^X = \{y : y \subseteq X\}$

***Определение* 2.1.2.**

Размещением из n элементов по k называется инъективное отображение $[k] \to [n]$.

A_n^k — количество размещений из n элементов по k.

Свойства.

1. $A_n^1 = n$
2. $A_n^{k+1} = nA_{n-1}^k$
3. $A_n^k = n \cdot (n-1) \cdot ... \cdot (n-k+1) = \frac{n!}{(n-k)!}$

***Определение* 2.1.3.**

Сочетание из n элементов по k называется k-элементное подмножество множества $[n]$.

$\binom{n}{k} = C_n^k$ — число сочетаний из n элементов по k.

Свойства.

1. $\binom{n}{k} = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!}$
2. $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + ... + \binom{n}{n} = 2^n$
3. $\binom{n}{k} = \binom{n}{n-k}$
4. $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$

Теорема 2.1.1. (**Бином Ньютона**)

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Доказательство.

Докажем по индукции.

База:

$n = 1$, $(a+b)^1 = \binom{0}{1}a + \binom{1}{1}b$.

Переход:

$$(a+b)^{n+1} = (a+b)(a+b)^n = (a+b)\sum_{k=0}^{n}\binom{n}{k}a^k\,b^{n-k} = \sum_{k=0}^{n}\binom{n}{k}a^{k+1}\,b^{n-k} + \sum_{k=0}^{n}\binom{n}{k}a^k\,b^{n-k+1} = \sum_{k=1}^{n+1}\binom{n}{k-1}a^k\,b^{n-k+1} + \sum_{k=0}^{n}\binom{n}{k}a^k\,b^{n-k+1} = \sum_{k=1}^{n}\left(\binom{n}{k} + \binom{n}{k-1}\right)a^k\,b^{n-k+1} + a^{n+1} + b^{n+1} = \sum_{k=0}^{n+1}\binom{n+1}{k}a^k\,b^{n-k} \quad \square$$

2.2. Характеристические функции множества

***Определение* 2.2.1.**

$A \subseteq X$

$\chi_A : X \to \{0, 1\}$ — характеристическая функция множества.

$$\chi_A(x) = \begin{cases} 1 & , \ x \in A \\ 0 & , \ x \notin A \end{cases}$$

Свойства.

1. $\chi_{A \cap B} = \chi_A \cdot \chi_B$
2. $\chi_{X \setminus A} = 1 - \chi_A$
3. $\chi_{A \cup B} = \chi_{\overline{\overline{A} \cap \overline{B}}} = 1 - \chi_{\overline{A} \cap \overline{B}} = 1 - \chi_{\overline{A}} \cdot \chi_{\overline{B}} = 1 - (1 - \chi_A)(1 - \chi_B)$
   Отсюда $\chi_{A \cup B} = \chi_A + \chi_B - \chi_A \cdot \chi_B$
4. $\chi_{A_1 \cup A_2 \cup \ldots \cup A_n} = \chi_{\overline{\overline{A_1} \cap \overline{A_2} \cap \ldots \cap \overline{A_n}}} = 1 - (1 - \chi_{A_1})(1 - \chi_{A_2}) \ldots (1 - \chi_{A_n})$
   Отсюда $\chi_{A_1 \cup A_2 \cup \ldots \cup A_n} = \sum\limits_i \chi_{A_i} - \sum\limits_{i<j} \chi_{A_i} \cdot \chi_{A_j} + \sum\limits_{i<j<k} \chi_{A_i} \cdot \chi_{A_j} \cdot \chi_{A_k} - \ldots$
5. $|A| = \sum\limits_{x \in X} \chi_A(x)$
6. $|A_1 \cup A_2 \cup \ldots \cup A_n| = \sum\limits_{x \in X} \chi_{A_1 \cup A_2 \cup \ldots \cup A_n}(x) = \sum\limits_i |A_i| - \sum\limits_{i<j} |A_i \cap A_j| + \sum\limits_{i<j<k} |A_i \cap A_j \cap A_k| - \ldots$

Утверждение 2.2.1.

Количество способов представить число n в виде суммы k неотрицательных слагаемых (порядок важен) равно $\binom{n+k-1}{k-1} = \binom{n+k-1}{n}$

Пример. (**Счастливые билетики**)

Количество счастливых билетиков — это количество таких шестизначных чисел $\overline{a_1 a_2 a_3 a_4 a_5 a_6}$, что $a_1 + a_2 + a_3 = a_4 + a_5 + a_6$.

Сделав замену, получаем, что это то же, что $(9 - b_1) + (9 - b_2) + (9 - b_3) = b_4 + b_5 + b_6 \implies b_1 + b_2 + b_3 + b_4 + b_5 + b_6 = 27$. Таким образом, свели задачу к поиску количества наборов b_i с заданной суммой.

Будем действовать от противного. Пусть B_i — множество неправильных "билетиков", в которых $b_i \geqslant 10$. Посчитаем тогда количество **всех** неправильных билетиков, то есть размер множества $X = \bigcup_{i=1}^{6} B_i$.

Заметим, что количество позиций i таких, что $b_i \geqslant 10$, может быть не больше двух (поскольку сумма должна быть равна 27).

Тогда посчитаем размер X по уже известной формуле: $|X| = \sum\limits_{i=1}^{6} |B_i| - \sum\limits_{i<j} |B_i \cap B_j|$

Заметим, что B_i равно количеству разбиений 17 элементов на 6 групп (или же количеству способов представить 17 в виде суммы 6 неотрицательных слагаемых). Формально это можно представить таким образом:

Положим, что у нас есть 6 корзин, по которым мы должны разместить 27 мячей, а также в конкретной корзине должно быть по крайней мере 10 мячей. Тогда мы можем сразу поместить 10 мячей в эту корзину, а оставшиеся 17 разместить произвольным образом.

Получаем, что $|B_i| = \binom{17+6-1}{6-1} = \binom{22}{5}$.

Аналогичным образом, $\forall\, i < j : |B_i \cap B_j| = \binom{7+6-1}{6-1} = \binom{12}{5}$.

Таким образом, $|X| = 6\binom{22}{5} - \frac{6\cdot 5}{2}\binom{12}{5} = 6\binom{22}{5} - 15\binom{12}{5}$.

Посчитаем теперь количество корректных билетиков. По сути, это число всевозможных разбиений 27 на 6 слагаемых за вычетом числа неправильных разбиений.

Отсюда ответ: $\binom{27+6-1}{6-1} - |X| = \binom{32}{5} - 6\binom{22}{5} + 15\binom{12}{5}$

2.3. Равномощные множества

***Определение* 2.3.1.**

Множества A и B называются равномощными, если между ними существует биекция.

Пример.

1. $[0, 1]$ и $[0, 2]$
2. $(0, 1) \sim (1, +\infty) \sim (0, +\infty)$
3. $A = 2^{\mathbb{N}}$ и $B =$ *множество последовательностей из 0 и 1*
4. 2^X равномощно множеству отображений $X \to \{0,\, 1\}$

Счётное множество равномощно $\mathbb{N}$.

Теорема 2.3.1. (**Свойства счётных множеств**)

1. Любое подмножество счётного множества счётно или конечно.
2. В любом бесконечном множестве есть счётное подмножество.
3. Объединение конечного или счётного числа конечных или счётных множеств конечно или счётно.

Доказательство.

1. Если подмножество счётного множества не конечно, то оно бесконечно. Однако заметим, что, поскольку оно является подмножеством счётного, то все его элементы можно занумеровать. Следовательно, можно получить биекцию между элементами подмножества и множество $\mathbb{N} \implies$ оно счётно.
2. Рассмотрим бесконечное множество A. Будем последовательно строить наше счётное подмножество.

   Выберем в качестве первого элемента a_1 произвольный. Поскольку A бесконечно, то в нём есть ещё элементы (помимо a_1). Таким образом, можем выбрать элемент $a_2 \neq a_1$.

   В общем случае, если у нас уже выбрано подмножество элементов $\{a_1,\, a_2,\, \ldots,\, a_{n-1}\}$, то, поскольку оно конечно, в A существуют элементы, отличные от уже выбранных. Следовательно, можем выбрать элемент a_n.

   В итоге получаем бесконечную последовательность элементов из A.

3. Покажем, что объединение счётного количества счётных множеств счётно (самое сильное из всех вышеописанных утверждений).

   Рассмотрим счётное количество счётных множеств A_1, A_2, A_3, ... Будем полагать, что все элементы в них различны (если это не так, то в конце сможем выкинуть их). Расположим элементы наших множеств в качестве таблицы:

$$\begin{array}{lccccc} A_1: & a_{11} & a_{12} & a_{13} & a_{14} & \dots \\ A_2: & a_{21} & a_{22} & a_{23} & a_{24} & \dots \\ A_3: & a_{31} & a_{32} & a_{33} & a_{34} & \dots \\ A_4: & a_{41} & a_{42} & a_{43} & a_{44} & \dots \\ & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

   Будем теперь обходить таблицу "змейкой" по диагоналям следующим образом:

$$a_{11},\ a_{21},\ a_{12},\ a_{31},\ a_{22},\ a_{13},\ a_{41},\ a_{32},\ a_{23},\ a_{14},\ \dots$$

   Таким образом, любой элемент таблицы будет стоять на некотором вполне конкретном месте в последовательности, а само множество выписанных элементов будет счётно (поскольку каждому элементу соответствует некоторый номер — его позиция в последовательности).

□

Примеры счётных множеств.

1. $\mathbb{Q}$ — счётное объединение счётных множеств.
2. $\mathbb{N}^2 = \bigcup\limits_{i\in\mathbb{N}} \mathbb{N} \times \{i\}$ (счётное объединение счётных).
3. $\mathbb{N}^k = \bigcup\limits_{i\in\mathbb{N}} \mathbb{N}^{k-1} \times \{i\}$
4. Множество конечных последовательностей натуральных чисел:
   $\mathbb{N}^0 \cup \mathbb{N}^1 \cup \mathbb{N}^2 \cup \mathbb{N}^3 \cup \dots$
5. Σ — конечный алфавит.

   Σ^* — множество конечных строк, поскольку существует соответствие между символами алфавита и $\mathbb{N}$.

   Тогда Σ^* — подмножество $\mathbb{N}$.
6. Множество программ на языке C++.
7. Множества алгебраических чисел.

   Алгебраическое число — корень многочлена с рациональными коэффициентами.

Теорема 2.3.2.

A — бесконечное множество, B — не больше чем счётное. Тогда $A \cup B$ равномощно A.

Доказательство.

Пусть $C \subseteq A$, C — счётно.

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

Поскольку C счётно, то $C \cup B$ тоже счётно $\implies$ существует биекция между C и $C \cup B$.

Таким образом, $A \cup B = (A \setminus C) \cup (C \cup B)$. □

Теорема 2.3.3.

[0, 1] равномощно множеству бесконечных последовательностей из 0 и 1.

Доказательство.

Представим любое число из [0, 1] в двоичной системе счисления. Если же запись некоторого $x \in [0,\,1]$ в двоичной системе счисления представляется конечным числом нулей и единиц, то сделаем из дроби периодическую. К примеру, $0,00101101 \to 0,00101100(1)$.

Тогда $[0,\,1] \sim A \subseteq 2^{\mathbb{N}}$. Однако заметим, что A — это $2^{\mathbb{N}}$ с выкинутым счётным подмножеством (без конечных последовательностей из 0 и 1), а потому $[0,\,1] \sim 2^{\mathbb{N}}$. □

Теорема 2.3.4.

$[0,\,1] \sim [0,\,1] \times [0,\,1]$

Доказательство.

Как и в доказательстве прошлой теоремы, представим все конечные десятичные дроби в виде бесконечных.

Построим теперь биекцию $f : [0,\,1] \times [0,\,1] \to [0,\,1]$.

Рассмотрим два произвольных числа a и b, что $a = \overline{0, \alpha_1\alpha_2\alpha_3\ldots},\ b = \overline{0, \beta_1\beta_2\beta_3\ldots}$

Тогда биекция будет выглядеть следующим образом: $f(a,b) = \overline{0, \alpha_1\beta_1\alpha_2\beta_2\alpha_3\beta_3\ldots}$

А поскольку существует биекция между множествами, то они равномощны. □

Теорема 2.3.5. (**теорема Кантора-Берштейна**)

Если A равномощно подмножеству B, а B равномощно подмножеству A, то A и B равномощны.

Другая формулировка: Пусть $A_0 \supseteq A_1 \supseteq A_2$ и $|A_0| = |A_2|$. Тогда $|A_0| = |A_1|$.

Доказательство.

Пусть $f : A_0 \to A_2$ — биекция. Тогда положим $f(A_1) = A_3 \subseteq A_2$ (поскольку $A_1 \subseteq A_0$), $f(A_2) = A_4,\ f(A_3) = A_5$ и так далее.

Получили систему вложенных множеств: $A_0 \supseteq A_1 \supseteq A_2 \supseteq A_3 \supseteq A_4 \supseteq \ldots$

Представим теперь множество A_0 в виде объединения непересекающихся слоёв:

$C_k = A_k \setminus A_{k+1},\ C = \bigcap_{k=0}^{\infty} A_k,\ A_0 = (\bigcup_{k=0}^{\infty} C_k) \cup C$

Мы знаем, что $\forall k$ существует биекция между множествами A_{2k} и A_{2k+2}, а также A_{2k+1} и A_{2k+3}. Поскольку $A_{2k+1} \subseteq A_{2k}$, а $A_{2k+3} \subseteq A_{2k+2}$, то существует биекция между $A_{2k} \setminus A_{2k+1} = C_{2k}$ и $A_{2k+2} \setminus A_{2k+3} = C_{2k+2}$.

Следовательно, можем установить соответствие между A_0 и A_1 следующим образом:

$$\begin{aligned} A_0 &= C_0 \sqcup C_1 \sqcup C_2 \sqcup C_3 \sqcup C_4 \sqcup \ldots \sqcup C \\ &\quad\ \downarrow \qquad\ \downarrow \qquad\ \downarrow \\ A_1 &= \qquad C_1 \sqcup C_2 \sqcup C_3 \sqcup C_4 \sqcup \ldots \sqcup C. \end{aligned}$$

Таким образом, получили биекцию между A_0 и $A_1 \implies |A_0| = |A_1|$. □

Теорема 2.3.6. (**теорема Кантора**)

Множество последовательностей из 0 и 1 несчётно.

Доказательство.

Случай обобщённой теоремы Кантора при $X = \mathbb{N}$. □

Теорема 2.3.7. (**обобщённая теорема Кантора**)

X неравномощно 2^X.

Доказательство.

От противного. Пусть X равномощно 2^X. Тогда существует биекция $f : X \to 2^X$.

Рассмотрим множество $A = \{x \in X : x \notin f(x)\}$.

Поскольку f — биекция, то $\exists y \in X : f(y) = A$. Тогда рассмотрим два случая:

1. $y \in A$

   Поскольку $f(y) = A$, то $y \in f(y) \implies y \notin A \implies$ противоречие

2. $y \notin A$

   Тогда аналогично получаем, что $y \notin f(y) \implies y \in A \implies$ противоречие.

Следовательно, наше предположение было неверно и X неравномощно 2^X.

А поскольку $X \subset 2^X$, то $|X| < \left|2^X\right|$. □

Примеры.

1. [0, 1] несчётно.
2. $\mathbb{R} = (-\infty, 0) \cup \{0\} \cup (0, +\infty)$ несчётно.

   $(-\infty, 0) \sim (-1, 0)$; $(0, +\infty) \sim (0, 1)$
3. Множество трансцендентных (не алгебраических) чисел несчётно (поскольку $\mathbb{R} = A \cup T$).
4. $\exists$ подмножество $\mathbb{N}$, для которого нет программы на C++, вычисляющей её характеристическую функцию.

   Пояснение:

   Существует счётное число программ на языке C++. А различных подмножеств $\mathbb{N}$ — несчётное количество. Следовательно, не существует биекции между этими множествами.

3. Теория графов

3.1. Введение в графы

***Определение* 3.1.1.**

V — множество вершин.

$E \subseteq V \times V$ — множество рёбер.

Запись $(u, v) \in E$ обозначает существование ребра между вершинами u и v.

Пара (V, E) образует ориентированный граф.

Определение* 3.1.2.** (Путь между вершинами***)

Между вершинами u и v существует путь, если существует такой набор вершин w_1, w_2, ..., w_n, что (u, w_1), (w_1, w_2), ..., $(w_n, v) \in E$.

$u \rightsquigarrow v$ — u и v связаны путём.

$u \leftrightsquigarrow v \iff u \rightsquigarrow v$ и $v \rightsquigarrow u$

***Свойства* $\rightsquigarrow$.**

1) $u \rightsquigarrow u$

2) $u \rightsquigarrow v$ и $v \rightsquigarrow w \implies u \rightsquigarrow w$

***Свойства* $\leftrightsquigarrow$.**

1) $u \leftrightsquigarrow u$

2) $u \leftrightsquigarrow v$ и $v \leftrightsquigarrow w \implies u \leftrightsquigarrow w$

3) $u \leftrightsquigarrow v \implies v \leftrightsquigarrow u$

Таким образом, отношение $\leftrightsquigarrow$ является отношением эквивалентности.

***Определение* 3.1.3.**

Классы эквивалентности по отношению $\leftrightsquigarrow$ — компоненты сильной связности.

Граф, построенный на компонентах сильной связности как на вершинах, является ациклическим.

***Определение* 3.1.4.**

Путь между двумя вершинами простой, если все вершины в нём разные.

Утверждение 3.1.1.

Две вершины соединены путём $\iff$ они соединены простым путём.

Доказательство.

Рассмотрим самый короткий путь.

Если в нём есть две повторяющиеся вершины, то отрезок пути между ними можно выкинуть, то есть он не самый короткий.

Следовательно, самый коротки путь — простой. □

***Определение* 3.1.5.**

Цикл — это путь w_1, w_2, ..., w_n, что $w_1 = w_n$.

Длина цикла — количество рёбер в нём.

Цикл простой, если w_1, w_2, ..., w_n различны и $n > 1$.

***Определение* 3.1.6.**

Граф $G = (V, E)$ называется неориентированным, если $\forall\, u, v \in V : (u, v) \in E \iff (v, u) \in E$.

$v \sim u$, если существует путь между вершинами v и u.

$\sim$ — отношение эквивалентности.

Классы эквивалентности — компоненты связности.

Граф связный, если в нём всего одна компонента связности.

3.2. Деревья

***Определение* 3.2.1.**

Дерево — связный неориентированный граф без циклов.

***Определение* 3.2.2.**

$\deg v = |\{u : (u, v) \in E\}|$ — степень вершины (количество исходящих рёбер).

Лемма.

В дереве всегда есть хотя бы 2 вершины степени 1.

Доказательство.

Рассмотрим самый длинный простой путь.

Его концы — искомые вершины. □

Теорема 3.2.1.

Следующие утверждения эквивалентны:

1) G — дерево.

2) G — связный граф с n вершинами и $n - 1$ ребром.

3) G — граф без циклов с n вершинами и $n - 1$ ребром.

4) G — связный граф, что при удалении $\forall$ ребра связность теряется.

5) G — граф без циклов, что при добавлении $\forall$ ребра появляется цикл.

Доказательство.

1) $\iff$ 4) :

Если в графе есть цикл, то можно удалить ребро не нарушив связность.

1) $\implies$ 5) :

Между любыми двумя вершинами в графе есть путь $\implies$ при добавлении любого ребра появится цикл.

5) $\implies$ 1) :

Если граф не связен, то можно соединить две вершины, не создав цикла $\implies$ противоречие.

1) $\implies$ 2) и 1) $\implies$ 3) :

Покажем, что в связном графе без циклов n вершин, то в нём $n - 1$ ребро.

Индукция по количеству вершин.

База: 1 вершина. **Переход:** $n \to n + 1$.

Рассмотрим связный граф без циклов из $n + 1$ вершины.

Выкинем из него любую вершину степени один. Такая обязательно найдётся.

Формально, можно объяснить это от противного:

Пусть это не так, то есть степень каждой вершины по крайней мере 2. Тогда можно выбрать произвольную вершину и начать гулять от неё по графу (не посещая никакую вершину дважды). Тогда мы либо найдём цикл (чего быть не может), либо попадём в вершину степени один.

Таким образом, в оставшемся графе будет n вершин и $n-1$ ребро.

Затем вернём вершину. Поскольку её степень была равна единице, то получим граф из $n+1$ вершины и n рёбер.

2) $\Longrightarrow$ 1) :

Пусть в графе есть цикл. Удалим из этого цикла любое ребро.

Тогда мы получим связный граф на n вершинах и $n-2$ рёбрах. Однако это невозможно.

3) $\Longrightarrow$ 1)

Пусть граф не связен. Положим, что в нём две компоненты связности.

Тогда в него можно добавить ребро, объединив две компоненты и не получив цикл.

Таким образом, мы получим граф из n вершин и n рёбер. Но в таком графе обязательно найдётся цикл $\Longrightarrow$ противоречие.

□

Теорема 3.2.2.
Из произвольного связного графа можно удалить несколько рёбер так, чтобы получить дерево.

Доказательство.
Пока в графе есть циклы, удаляем из них рёбра. □

Следствие.
В связном графе $|E| \geqslant |V| - 1$.

3.3. Эйлеров путь и цикл

***Определение* 3.3.1.**
Эйлеров цикл — цикл, проходящий по каждому ребру графа ровно 1 раз.

Теорема 3.3.1.
В связном графе есть Эйлеров цикл $\Longleftrightarrow$ степени всех вершин чётны.

Доказательство.
“$\Longrightarrow$”

Очевидно.

“$\Longleftarrow$”

Рассмотрим самый большой по количеству рёбер цикл C. Если в него входят все рёбра графа, то мы нашли Эйлеров цикл.

Пусть G' — это граф G, из которого выкинули все рёбра из C. В G' степени всех вершин чётны.

Поскольку G связен, то в G' найдётся ребро, один из концов которого лежит в C. Назовём этот конец v.

Тогда заметим, что в G' обязательно найдётся цикл, содержащий v. Это так, поскольку степень каждой вершины чётная, и если мы входим в какую-нибудь вершину, то мы всегда сможем из неё выйти по ещё непосещённому ребру.

Таким образом, рано или поздно мы попадём в v, найдя цикл C'.

В конце мы можем объединить циклы C и C' и рассматривать новый граф без них. Продолжать будем до тех пор, пока граф не опустеет. □

***Определение* 3.3.2.**
Эйлеров путь — путь, проходящий по всем рёбрам графа ровно 1 раз.

Теорема 3.3.2.
Эйлеров путь существует тогда и только тогда, когда степени ровно двух вершин нечётны, либо когда существует Эйлеров цикл.

Доказательство.
"$\Longrightarrow$"

Очевидно.

"$\Longleftarrow$"

Если есть Эйлеров цикл, то есть и Эйлеров путь.

Если же есть 2 вершины нечётной степени, то соединим их ребром и найдём Эйлеров цикл, а затем выкинем это ребро. □

3.4. Раскраски

***Определение* 3.4.1.**
$G = (V, E)$ — неориентированный граф.

$h : V \to [k]$ называется правильной раскраской в k цветов, если никакое ребро не соединяет две вершины одинакового цвета.

***Определение* 3.4.2.**
Граф k-раскрашиваем, если существует правильная раскраска его в k цветов.

***Определение* 3.4.3.**
Двудольный граф — 2-раскрашиваемый граф.

Теорема 3.4.1. (**Критерий двудольности**)

Граф двудольный $\iff$ в нём нет простых циклов нечётной длины.

Доказательство.
"$\Longrightarrow$"

Очевидно (поскольку цвета на циклах должны чередоваться).

"$\Longleftarrow$"

Докажем для связных графов.

Выберем какую-нибудь вершину и начнём из неё обходить граф (dfs или bfs). Таким образом, мы можем все вершины распределить по "уровням" (т.е. каждой вершине сопоставить глубину в дереве обхода), а затем по уровням же покрасить их.

Покажем, что если в графе все циклы чётной длины, то не может быть ребра между вершинами одного цвета.

От противного. Пусть $h(v) = h(u)$ и $(v, u) \in E$. Обозначим вершину, из которой мы начали обход, за s.

Рассмотрим пути $s \rightsquigarrow v$ и $s \rightsquigarrow u$. Поскольку $h(v) = h(u)$, то $|s \rightsquigarrow v| + |s \rightsquigarrow u| \vdots 2$. Следовательно, вместе с ребром (v, u) мы получим цикл нечётной длины $\Longrightarrow$ противоречие. $\square$

4. Дискретная теория вероятностей

4.1. Введение в дискретную теорию вероятностей

***Определение* 4.1.1.** Ω — конечное множество, пространство элементарных событий.

Пример. $\{1, 2, 3, 4, 5, 6\}$ — значения, которые могут выпасть на кубике.

***Определение* 4.1.2.** Событие — подмножество Ω.

***Определение* 4.1.3.** Вероятностная мера — $\Pr : 2^\Omega \to [0, 1]$.
Пусть $A \subseteq \Omega$ — некоторое событие. Тогда вероятность события A — $\Pr(A)$.

Свойства.

1) $\Pr(\Omega) = 1$

2) $A \cap B = \varnothing \implies \Pr(A \cup B) = \Pr(A) + \Pr(B)$

***Определение* 4.1.4.** $(\Omega, \Pr)$ — вероятностное пространство.

Свойства вероятностного пространства.

1) $\Pr(\varnothing) = 0$

2) $A \subseteq B \implies \Pr(A) \leqslant \Pr(B)$

3) $\Pr(A_1 \cup A_2 \cup ... \cup A_n) \leqslant \Pr(A_1) + \Pr(A_2) + ... + \Pr(A_n)$

4) Формула включения-исключения:

$\Pr(\bigcup_i A_i) = \sum_i \Pr(A_i) - \sum_{i<j} \Pr(A_i \cap A_j) + \sum_{i<j<k} \Pr(A_i \cap A_j \cap A_k) - ...$

5) $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$

$p_i = \Pr(\{\omega_i\})$ — вероятность события ω_i

$\Pr(A) = \sum_{\omega_i \in A} p_i$

Доказательство.

1) $\Pr(\varnothing) + \Pr(\Omega) = 1$

2) $B = A \cup (B \setminus A),\ A \cap (B \setminus A) = \varnothing \implies \Pr(B) = \Pr(A) + \Pr(B \setminus A) \geqslant \Pr(A)$

3) По индукции:

$\Pr(A \cup B) \leqslant \Pr(A) + \Pr(B)$, поскольку $\Pr(A \cup B) = \Pr(A) + \Pr(B \setminus A) \leqslant \Pr(A) + \Pr(B)$ □

Пример.

В школе учится n детей, все они ходят на кружки. В каждом кружке учится ровно d человек, всего в школе кружков $k \leqslant 2^{d-1}$, $d \geqslant 2$. Можно ли выдать некоторым детям по айфону так, чтобы в любом кружке были дети как с айфонами, так и без?

Решение:

Опишем наше вероятностное пространство.

Ω — всевозможные способы раздать айфоны детям. Каждое событие — бинарная строка из n нулей и единиц (в i-ой позиции стоит 1, если мы выдаём айфон i-ому ребёнку, и 0 иначе), $|\Omega| = 2^n$. Вероятность каждого события равна $\frac{1}{2^n}$.

Положим событие A_i — i-ый кружок нарушает описанное правило.

Тогда $\Pr(A_i) = \frac{1}{2^n} \cdot (2^{n-d} + 2^{n-d}) = \frac{2^{n-d+1}}{2^n} = 2^{1-d}$ (то есть считаем, что в кружке или у всех детей есть айфоны, или у всех детей они отсутствуют).

$\Pr(\exists$ *кружок, нарушающий правило*$) = Pr(A_1 \cup A_2 \cup ... \cup A_k)$

$\Pr(A_1 \cup A_2 \cup ... \cup A_k) \leqslant \Pr(A_1) + \Pr(A_2) + ... + \Pr(A_k) \leqslant 2^{1-d} \cdot 2^{d-1} = 1.$

Наихудший случай — когда кружков ровно 2^{d-1}. Однако заметим, что и в таком случае не будет достигаться равенство, поскольку для любых двух кружок существует ненулевая вероятность того, что оба будут нарушать правило.

Таким образом, $\Pr($*все кружки хорошие*$) > 0$. Следовательно, можно так распределить айфоны, что все кружки будут удовлетворять правилу.

4.2. Теорема Эрдеша-Ко-Радо

Теорема 4.2.1.

$S = \{0, 1, ..., n-1\}$

$\mathcal{F} \subseteq 2^S$

Пусть выполняются следующие условия:

1) $\forall\, A \in \mathcal{F} \quad |A| = k,\ k \leqslant \frac{n}{2}$

2) $\forall\, A, B \in \mathcal{F} \quad A \cap B \neq \varnothing$

Тогда $|\mathcal{F}| \leqslant \binom{n-1}{k-1}$.

Доказательство.

1) Докажем небольшую лемму.

Рассмотрим все множества $A_s = \{s, s+1, ..., s+k-1\}$ (считаем, что суммирование производится по модулю n). Тогда утверждается, что в $\mathcal{F}$ войдёт не более k множеств такого вида.

Положим, что для некоторого $s \in S : A_s \in \mathcal{F}$.

Помимо A_s, в $\mathcal{F}$ могут входить только некоторые множества из A_{s-k+1}, ..., A_{s-1}, A_{s+1}, ..., A_{s+k-1} (таким образом мы требуем, чтобы эти множества пересекались с A_s).

Разобьём описанные множества по парам следующим образом:

(A_{s-k+1}, A_{s+1}), (A_{s-k+2}, A_{s+2}), ..., (A_{s-1}, A_{s+k-1})

Заметим, что внутри каждой пары множества не пересекаются. Следовательно, из каждой пары мы можем выбрать не более одного представителя. Поскольку всего пар $k-1$, то мы можем выбрать не более k таких множеств.

2) Заметим, что описанную лемму можно обобщить для случая перестановок:

Пусть $\sigma : S \to S$ — биекция.

$A_s^\sigma = \{\sigma(s), \sigma(s+1), ..., \sigma(s+k-1)\}$

Тогда $\forall\, \sigma$ $\mathcal{F}$ содержит $\leqslant k$ множеств A_s^σ.

3) Пусть σ — случайная биекция, i — случайный элемент S.

Положим вероятностное пространство $\Omega = \{(\sigma, i) : \sigma \in S_n,\ i \in S\},\ \ |\Omega| = n! \cdot n$.

Рассмотрим множество $\mathcal{U} = \{(\sigma, i) : A_i^\sigma \in \mathcal{F}\}$.

Заметим, что, с одной стороны, $\Pr(\mathcal{U}) \leqslant \frac{1}{n! \cdot n} \cdot (k \cdot n!) = \frac{k}{n}$ (суммируем по всем перестановкам). С другой стороны, $\Pr(\mathcal{U}) = \frac{|\mathcal{F}|}{\binom{n}{k}}$. Следовательно, $\frac{|\mathcal{F}|}{\binom{n}{k}} \leqslant \frac{k}{n} \implies |\mathcal{F}| \leqslant \frac{k}{n} \cdot \binom{n}{k} = \binom{n-1}{k-1}$.

□

4.3. Случайная величина и её математическое ожидание

***Определение* 4.3.1.**

$(\Omega, \Pr)$ — конечное вероятностное пространство.

Случайной величиной называется функция $X : \Omega \to \mathbb{R}$, сопоставляющая каждому событию некоторое числовое значение.

$[X \in A] = \{\omega \in \Omega : X(\omega) \in A\}$, где $A \subseteq \mathbb{R}$.

Пример.

$\Omega = \{1, 2, 3, 4, 5, 6\}$ — значения, которые могут выпасть на игральном кубике.

$X(1) = 1^2,\ X(2) = 2^2,\ ...,\ X(6) = 6^2$

$[X \in \{1, 2, ..., 10\}] = \{1, 2, 3\}$

***Определение* 4.3.2.**

$(\Omega, \Pr)$ — конечное вероятностное пространство.

$X : \Omega \to \mathbb{R}$ — случайная величина.

$\mathbb{E}[X]$ — математическое ожидание.

$\mathbb{E}[X] = \sum\limits_{\omega \in \Omega} \Pr(\omega) \cdot X(\omega)$

Свойства.

1) Линейность.

$X, Y : \Omega \to \mathbb{R},\ \alpha, \beta \in \mathbb{R}$

Тогда $\mathbb{E}[\alpha X + \beta Y] = \alpha\mathbb{E}[X] + \beta\mathbb{E}[Y]$

2) Принцип усреднения.

$X : \Omega \to \mathbb{R}$

Тогда $\Pr\big[X \geqslant \mathbb{E}[X]\big] > 0$ и $\Pr\big[X \leqslant \mathbb{E}[X]\big] > 0$

Доказательство.

1) Распишем по определению:

$\mathbb{E}[\alpha X + \beta Y] = \sum\limits_{\omega \in \Omega} \Pr(\omega) \cdot \big(\alpha X(\omega) + \beta Y(\omega)\big) = \alpha \sum\limits_{\omega \in \Omega} \Pr(\omega) X(\omega) + \beta \sum\limits_{\omega \in \Omega} \Pr(\omega) Y(\omega) = \alpha\mathbb{E}[X] + \beta\mathbb{E}[Y]$

2) От противного.

Пусть $\forall\, \omega \in \Omega : \Pr(\omega) > 0 \quad X(\omega) < \mathbb{E}[X]$

$\mathbb{E}[X] = \sum\limits_{\omega \in \Omega} \Pr(\omega) X(\omega) < \sum\limits_{\omega \in \Omega} \Pr(w)\mathbb{E}[X] = \mathbb{E}[X] \sum\limits_{\omega \in \Omega} \Pr(\omega) = \mathbb{E}[X] \implies$ противоречие $\square$

***Определение* 4.3.3.**

Турнир — орграф, между любыми двумя вершинами которого ровно 1 ребро.

Теорема 4.3.1.

$\forall\, n\ \exists$ турнир на n вершинах, в котором хотя бы $\frac{n!}{2^{n-1}}$ гамильтоновых путей.

Доказательство.

Ω — множество всех турниров на n вершинах, $|\Omega| = 2^{\frac{n(n-1)}{2}}$.

$X(\omega)$ — число гамильтоновых путей в ω.

Пусть σ — перестановка вершин.

Скажем, что $X_\sigma(\omega) = \begin{cases} 1 & \text{, если } \sigma \text{ задаёт гамильтонов путь в } \omega \\ 0 & \text{, иначе} \end{cases}$

Таким образом, $X(\omega) = \sum\limits_{\sigma \in S_n} X_\sigma(\omega)$.

$\mathbb{E}[X] = \sum\limits_{\sigma \in S_n} \mathbb{E}[X_\sigma] = \sum\limits_{\sigma \in S_n} (\Pr[X_\sigma = 1] \cdot 1 + \Pr[X_\sigma = 0] \cdot 0) = \sum\limits_{\sigma \in S_n} \Pr[X_\sigma = 1]$

Заметим, что σ задаёт гамильтонов путь, если существуют рёбра $\sigma(1) \to \sigma(2) \to ... \to \sigma(n)$. Направления остальных рёбер нас не интересуют.

Таким образом, $\Pr[X_\sigma = 1] = \frac{2^{\frac{n(n-1)}{2} - (n-1)}}{2^{\frac{n(n-1)}{2}}} = \frac{1}{2^{n-1}}$.

Отсюда получаем, что $\mathbb{E}[X] = \sum\limits_{\sigma \in S_n} \Pr[X_\sigma = 1] = \frac{n!}{2^{n-1}}$.

А по принципу усреднения существует такой турнир, что количество гамильтоновых путей в нём по крайней мере $\frac{n!}{2^{n-1}}$. □

4.4. 3-КНФ и неравенство Маркова

Теорема 4.4.1.

φ — формула в 3-КНФ, m — число дизъюнктов в φ.

Тогда $\exists$ набор переменных, который выполняет $\geqslant \frac{7}{8}m$ дизъюнктов.

Доказательство.

Пусть n — число переменных.

$\Omega = \{0,1\}^n$ — множество всевозможных значений переменных, Pr — равномерная мера.

Положим Y — число невыполнимых дизъюнктов.

$Y_i = \begin{cases} 1 & \text{, если не выполняется } i\text{-ый дизъюнкт} \\ 0 & \text{, иначе} \end{cases}$

Тогда $\mathbb{E}[Y] = \mathbb{E}[Y_1 + Y_2 + ... + Y_m] = \sum\limits_{i=1}^{m} \mathbb{E}[Y_i]$.

$\mathbb{E}[Y_i] = \Pr[Y_i = 1] \cdot 1 + \Pr[Y_i = 0] \cdot 0 = \Pr[Y_i = 1] = \frac{2^{n-3}}{2^n} = \frac{1}{8}$ (значения в дизъюнкте мы выбираем однозначно, а остальные — не принципиально, каким образом)

Следовательно, $\mathbb{E}[Y] = \frac{1}{8}m \implies \exists$ набор, что $Y(\omega) \leqslant \frac{1}{8}m$. Т.е. выполняет $\geqslant \frac{7}{8}m$ дизъюнктов. □

Теорема 4.4.2. (**Неравенство Маркова**)

$X : \Omega \to \mathbb{R}_{\geqslant 0}$, $\mathbb{E}[X] > 0$

Тогда $\Pr\big[X \geqslant C \cdot \mathbb{E}[X]\big] \leqslant \frac{1}{C} \quad \forall\, C > 0$

Доказательство.

$\Pr\big[X \geqslant C \cdot \mathbb{E}[X]\big] = \Pr\big(\{\omega : X(\omega) \geqslant C \cdot \mathbb{E}[X]\}\big) = \sum\limits_{X(\omega) \geqslant C \cdot \mathbb{E}[X]} \Pr(\omega)$

От противного. Пусть $\Pr\big[X \geqslant C \cdot \mathbb{E}[X]\big] > \frac{1}{C}$.

$\mathbb{E}[X] = \sum\limits_{\omega \in \Omega} \Pr(\omega) X(\omega) = \left(\sum\limits_{X(\omega) \geqslant C \cdot \mathbb{E}[X]} \Pr(\omega) X(\omega) \right) + \left(\sum\limits_{X(\omega) < C \cdot \mathbb{E}[X]} \Pr(\omega) X(\omega) \right)$

Отсюда $\mathbb{E}[X] \geqslant \sum\limits_{X(\omega) \geqslant C \cdot \mathbb{E}[X]} \Pr(\omega) X(\omega) \geqslant \sum\limits_{X(\omega) \geqslant C \cdot \mathbb{E}[X]} \Pr(\omega)(C \cdot \mathbb{E}[X]) > C \cdot \mathbb{E}[X] \cdot \frac{1}{C} = \mathbb{E}[X]$

Получили противоречие. □

Пример.

Приведём алгоритм поиска набора переменных в 3-КНФ, выполняющего $\geqslant \frac{7}{8}m$ дизъюнктов:

1) Выберем случайный набор.

2) Закончим, если набор выполняет $\geqslant \frac{7}{8}m$ дизъюнктов.

Воспользуемся неравенством Маркова и посчитаем вероятность успеха:

По-прежнему будем считать, что Y — число невыполненных дизъюнктов, $\mathbb{E}[Y] = \frac{m}{8}$.

$\Pr[Y > \frac{1}{8}m] = \Pr[8Y > m] = \Pr[8Y \geqslant m+1] = \Pr[Y \geqslant \frac{m+1}{8}] = \Pr[Y \geqslant \frac{m}{8} \cdot \frac{m+1}{m}] \leqslant \frac{m}{m+1}$

Следовательно, $\Pr[\text{алгоритм нашёл такой набор}] \geqslant \frac{1}{m+1}$.

Посчитаем вероятность того, что нам не повезёт спустя k запусков алгоритма. Это $\left(1 - \frac{1}{m+1}\right)^k$.

Таким образом, если мы положим $k = (m+1)t$, то нам не повезёт с вероятностью $< \frac{1}{e^t}$.

***Определение* 4.4.1.**

$(\Omega, \Pr)$ — вероятностное пространство.

A, B — независимые события, если $\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$.

В частности, A_1, A_2, ..., $A_k \subseteq \Omega$ — независимые, если $\Pr(\bigcap A_i) = \prod \Pr(A_i)$.

***Определение* 4.4.2.**

X_1, X_2, ..., $X_n : \Omega \to \mathbb{R}$ — независимые случайные величины, если $\forall\, A_1$, A_2, ..., A_n :

$\Pr[X_1{=}A_1, X_2{=}A_2, ..., X_n{=}A_n] = \Pr[X_1{=}A_1] \cdot \Pr[X_2{=}A_2] \cdot ... \cdot \Pr[X_n{=}A_n]$

Теорема 4.4.3.

X_1, X_2, ..., $X_n : \Omega \to \mathbb{R}$ — независимые случайные величины.

Тогда $\mathbb{E}[X_1 X_2 ... X_n] = \mathbb{E}[X_1]\mathbb{E}[X_2]...\mathbb{E}[X_n]$

Доказательство.

$A_1 = X_1(\Omega)$, $A_2 = X_2(\Omega)$, ..., $A_n = X_n(\Omega)$

$\mathbb{E}[X_1 \cdot X_2 \cdot ... \cdot X_n] = \sum\limits_{x \in \mathbb{R}} x \cdot \Pr[X_1 \cdot X_2 \cdot ... \cdot X_n = x] = \sum\limits_{a_1 \in A_1} \sum\limits_{a_2 \in A_2} ... \sum\limits_{a_n \in A_n} a_1 a_2 ... a_n \cdot \Pr[X_1{=}a_1, ..., X_n{=}a_n] =$

$= \sum\limits_{a_1 \in A_1} \sum\limits_{a_2 \in A_2} ... \sum\limits_{a_n \in A_n} a_1 a_2 ... a_n \cdot \Pr[X_1{=}a_1] \Pr[X_2{=}a_2] ... \Pr[X_n{=}a_n] =$

$= \left(\sum\limits_{a_1 \in A_1} a_1 \cdot \Pr[X_1{=}a_1]\right) \left(\sum\limits_{a_2 \in A_2} a_2 \cdot \Pr[X_2{=}a_2]\right) ... \left(\sum\limits_{a_n \in A_n} a_n \cdot \Pr[X_n{=}a_n]\right) = \mathbb{E}[X_1]\mathbb{E}[X_2]...\mathbb{E}[X_n]$

□

4.5. Энтропия случайной величины

Определение* 4.5.1.** (Энтропия***)

(Ω, Pr) – конечное вероятностное пространство

$X : \Omega \to \mathbb{R}^k$ – случайная величина

$X(\Omega) = \{a_1, a_2, \ldots, a_n\}$

$p_i = Pr[x = a_i]$

Тогда энтропия случайной величины выражается следующим образом:

$$H[X] = \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i}$$

Анонс-пояснение: энтропия – это количество информации, которую несёт в себе случайная величина.

Пример 1.
Пусть $X \in \{1,0\}^n$ – случайная величина, $\forall S \in \{0,1\}^n \ Pr[x = S] = 2^{-n}$. Тогда:

$$H[X] = \sum_{S \in \{0,1\}^n} \frac{1}{2^n} \log_2 2^n = n$$

Теорема 4.5.1. (**Неравенство Йенсена**)
Пусть $f : (a,b) \to \mathbb{R}$ – выпуклая. То есть:

$$\forall x, y \in (a,b), \alpha \in [0,1]$$

$$f(\alpha x + (1-\alpha)y) \leqslant \alpha f(x) + (1-\alpha) f(y)$$

Тогда:

$$\forall x_1, x_2, \ldots, x_n \in (a,b), \ \forall \lambda_1, \lambda_2, \ldots, \lambda_n \in [0,1], \sum_{i=1}^{n} \lambda_i = 1$$

$$f(\lambda_1 x_1 + \lambda_2 x_2 + \ldots + \lambda_n x_n) \leqslant \lambda_1 f(x_1) + \lambda_2 f(x_2) + \ldots + \lambda_n f(x_n)$$

Замечание.
$\log_2 x$ – вогнутая функция (иначе говоря, $-log_2 x$ – выпуклая)

Лемма. (**о чужих логарифмах**)

$$x_1, x_2, \ldots, x_n > 0, \ \sum_{i=1}^{n} x_i = 1, \ \ y_1, y_2, \ldots, y_n > 0, \ \sum_{i=1}^{n} y_i \leqslant 1$$

Тогда выполняется следующее:

$$\sum_{i=1}^{n} x_i \log_2 \frac{1}{x_i} \leqslant \sum_{i=1}^{n} x_i \log_2 \frac{1}{y_i}$$

Доказательство.

$$\sum_{i=1}^{n} x_i \log_2 \frac{1}{x_i} - \sum_{i=1}^{n} x_i \log_2 \frac{1}{y_i} = \sum_{i=1}^{n} x_i \log_2 \frac{y_i}{x_i} \underset{\text{Неравенство Йенсена}}{\leqslant} \log_2 \sum_{i=1}^{n} y_i \leqslant 0$$

□

Замечание.
$(X,Y) : \Omega \to (X(\Omega), Y(\Omega))$
$X : \Omega \to \mathbb{R}^k$
$Y : \Omega \to \mathbb{R}^l$
$(X,Y) : \Omega \to \mathbb{R}^{k+l}$

Теорема 4.5.2.

$H[X,Y] \leqslant H[X] + H[Y]$, причём равенство достигается тогда и только тогда, когда X и Y независимы

Доказательство.

Пусть X принимает значения $\{a_1, a_2, \ldots, a_n\}$, Y — значения $\{b_1, b_2, \ldots, b_m\}$, а также $p_{a_i} = \Pr[x = a_i],\ q_{b_i} = \Pr[y = b_i],\ \rho_{a_i,b_j} = \Pr[x = a_i, y = b_j]$. Тогда:

$$H[X,Y] = \sum_{i=1}^{n}\sum_{j=1}^{m} \rho_{a_i,b_j} \log_2 \frac{1}{\rho_{a_i,b_j}}$$

$$H[X] + H[Y] = \sum_{i=1}^{n} p_{a_i} \log_2 \frac{1}{p_{a_i}} + \sum_{j=1}^{n} q_{b_j} \log_2 \frac{1}{q_{b_j}}$$

Заметим, что $p_{a_i} = \sum_{j=1}^{m} \rho_{a_i,b_j}$ и $q_{b_j} = \sum_{i=1}^{n} \rho_{a_i,b_j}$. Действительно, вероятность того, что произойдет какое-то конкретное событие из X – сумма вероятностей того, что произойдет такое фиксированное событие и при этом любое событие из Y. И наоборот. Тогда:

$$H[X] + H[Y] = \sum_{i=1}^{n}\sum_{j=1}^{m} \rho_{a_i,b_j} \log_2 \frac{1}{p_{a_i}} + \sum_{i=1}^{n}\sum_{j=1}^{m} \rho_{a_i,b_j} \log_2 \frac{1}{q_{b_j}} = \sum_{i=1}^{n}\sum_{j=1}^{m} \rho_{a_i,b_j} \log_2 \frac{1}{p_{a_i} q_{b_j}}$$

Осталось проверить, что $\sum_{i=1}^{n}\sum_{j=1}^{m} p_{a_i} q_{b_j} = 1$.

Поменяем порядок суммирования: $\sum_{i=1}^{n} p_{a_i} \sum_{j=1}^{m} q_{b_j} = \sum_{i=1}^{n} p_{a_i} = 1$. Значит, можно использовать лемму о чужих логарифмах, по которой:

$$H[X,Y] = \sum_{i=1}^{n}\sum_{j=1}^{m} \rho_{a_i,b_j} \log_2 \frac{1}{\rho_{a_i,b_j}} \leqslant \sum_{i=1}^{n}\sum_{j=1}^{m} \rho_{a_i,b_j} \log_2 \frac{1}{p_{a_i} q_{b_j}} = H[X] + H[Y]$$

□

Замечание.

Если X и Y при этом были независимы, то $\Pr[x{=}a_i, y{=}b_j] = \Pr[x{=}a_i]\Pr[y{=}b_j] \implies \rho_{a_i,b_j} = p_{a_i} q_{b_j}$ и неравенство обращается в равенство.

Следствие. (***Полуаддитивность энтропии***)

Пусть $X_1, X_2, \ldots, X_n$ – случайные величины. Тогда $H[X_1, X_2, \ldots, X_n] \leqslant H[X_1] + \ldots + H[X_n]$

Доказательство.

Упражнение.

Подсказка: доказывается по индукции. □

***Определение* 4.5.2.**

$h(p) = p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p}$ – бинарная энтропия

$h(p) \uparrow$ при $p \in [0, \frac{1}{2}]$ и $h(p) \downarrow$ при $p \in [\frac{1}{2}, 1]$

Утверждение 4.5.3.

Пусть имеем некоторую вероятность p такую, что $0 \leqslant p \leqslant \frac{1}{2}$. Тогда:

$$\binom{n}{0} + \binom{n}{1} + \ldots + \binom{n}{\lfloor pn \rfloor} = \sum_{0 \leqslant i \leqslant \lfloor pn \rfloor} \binom{n}{i} \leqslant 2^{n \cdot h(p)}$$

Доказательство.

Докажем, что если есть некоторое множество $\mathcal{F} \subseteq \{0,1\}^n : \forall s \in \mathcal{F}$ в $s \leqslant pn$ единиц, то $|\mathcal{F}| = \sum_{0 \leqslant i \leqslant pn} \binom{n}{i}$.

Пусть X – случайная величина, равномерная на $|\mathcal{F}|$. Тогда $H[X] = \sum_{s \in \mathcal{F}} \frac{1}{|\mathcal{F}|} \log_2 |\mathcal{F}| = \log_2 |\mathcal{F}|$.

Что же это была за величина? $X = x_1 + x_2 + \ldots + x_n$, где $x_i = 1$, если i-й бит s равен 1, и $x_i = 0$ иначе. Тогда получаем, что:

$H[X] = H[x_1 + x_2 + \ldots + x_n] \leqslant H[x_1] + H[x_2] + \ldots + H[x_n] = nH[x_1] = n \cdot h(q)$, где q — вероятность того, что символ строки из $\mathcal{F}$ будет равен 1.

Заметим теперь, что поскольку количество единиц в таких строках не превышает np, то $q \leqslant p \Rightarrow h(q) \leqslant h(p)$ (т.к. бинарная энтропия на отрезке $[0, \frac{1}{2}]$ строго возрастает).

Таким образом, мы имеем следующую оценку:

$H[X] = \log_2 |\mathcal{F}|,\ H[X] \leqslant n \cdot h(p) \Rightarrow |\mathcal{F}| \leqslant 2^{n \cdot h(p)}$ □

4.6. Однозначно декодируемые коды. Неравенство Крафта

***Определение* 4.6.1.**

Пусть имеется некоторый алфавит $\Sigma = \{a_1, a_2, \ldots, a_n\}$, где a_i — i-й сивол алфавита.

Тогда $C : \Sigma \to \{0,1\}^*$ – однозначно декодируемый код, если $\forall$ 2-х слов их код не совпадает.

$$C : \Sigma^* \to \{0,1\}^*,\ \ C(x_1, x_2, \ldots, x_k) = C(x_1)C(x_2)\ldots C(x_k)$$

$$\forall\, x, y \in \Sigma^*,\ x \neq y \implies C(x) \neq C(y)$$

***Определение* 4.6.2.**

Код называется префиксным, если $\forall\, x \neq y \in \Sigma:\ C(x)$ не является префиксом $C(y)$

Пример.

$\Sigma = \{a, b, c\},\ C(a) = 0,\ C(b) = 10,\ C(c) = 11$. Здесь C – префиксный однозначно декодируемый код.

Лемма. (**Неравенство Крафта**)

1. Пусть $C : \Sigma \to \{0,1\}^*$ – однозначно декодируемый код. Тогда $\sum_{a \in \Sigma} 2^{-|C(a)|} \leqslant 1$
2. Пусть $l_1, l_2, \ldots, l_n \in \mathbb{N},\ \sum_{i=1}^n 2^{-l_i} \leqslant 1$. Тогда $\exists$ префиксный код $C : \Sigma \to \{0,1\}^*,\ \ l_i = |C(a_i)|$

Доказательство.

1. Положим, что $C : \Sigma \to \{u, v\}^*$ — однозначно декодируемый код, где $\Sigma = \{a_1, a_2, \ldots, a_k\}$, а u и v — некоторые произвольные значения.

   Тогда рассмотрим следующую формальную сумму: $(C(a_1) + C(a_2) + \ldots + C(a_k))^N$, где a_i — символы исходного алфавита Σ. В этой сумме не будет подобных слагаемых, поскольку код однозначно декодируемый. В частности, слагаемые вида uv и vu будут считаться различными.

   От противного. Пусть $\sum\limits_{i=1}^{k} 2^{-|C(a_i)|} > 1$. Положим $u = v = \frac{1}{2}$.

   Тогда, с одной стороны, $(C(a_1) + C(a_2) + \ldots + C(a_k))^N = \left(\sum\limits_{i=1}^{k} 2^{-|C(a_i)|}\right)^N$. Несложно заметить, что значение этого выражения растёт экспоненциально в зависимости от N.

С другой стороны, исходную формальную сумму можно раскрыть по биному Ньютона. Заметим, что сумма по словам одинаковой длины тогда будет не больше 1, т.к. мы рассматриваем однозначно декодируемый код, а потому все слагаемые будут различны. В то же время, длина максимального слова будет не больше $N \cdot \max\limits_{i=1}^{k}\{|C(a_i)|\}$, а эта величина растёт линейно.

Таким образом, мы получили противоречие.

Следовательно, $\sum\limits_{i=1}^{k} 2^{-|C(a_i)|} \leqslant 1$.

2. Назовём правильно отформатированным отрезком длины 2^{-k} полуинтервал следующего вида: $\left[2^{-k} \cdot n;\ 2^{-k} \cdot (n+1)\right),\ 0 \leqslant n \leqslant 2^k - 1$.

   Отсортируем все отрезки следующим образом: $l_1 \leqslant l_2 \leqslant \ldots \leqslant l_n$. Будем располагать отрезки на прямой последовательно друг за другом в таком порядке.

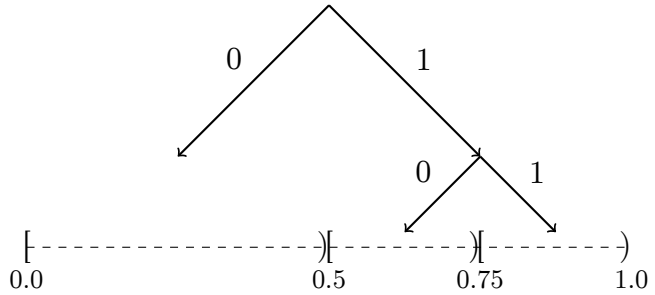
   Заметим, что каждый отрезок при этом будет правильно отформатирован. Это можно показать по индукции:

   Пусть $l_1 \leqslant l_2 \leqslant \ldots \leqslant l_n \leqslant l_{n+1},\ \sum\limits_{i=1}^{n+1} 2^{-l_i} \leqslant 1$, и первые n отрезков длины $l_1, l_2, \ldots, l_n$ соответственно правильно отформатированы. Тогда $n+1$-ый отрезок также будет правильно отформатирован, поскольку среди прочих его длина минимальна (l_{n+1} максимально), а потому уложенный префикс отрезков может быть разбит на целое количество блоков длины $2^{-l_{n+1}}$.

   Тогда заметим, что такому набору правильно отформатированных отрезков можно сопоставить двоичное дерево, пути до листьев в котором будут соответствовать префиксным кодам.

   Пример: $l_1 = 1,\ l_2 = 2,\ l_3 = 2$.

   Префиксные коды, соответствующие заданным длинам: 0, 10 и 11 соответственно.



□

Теорема 4.6.1.

1. Если $C : \Sigma \to \{0,1\}^*$ — однозначно декодируемый код, то $\forall\, p_1,p_2,\ldots,p_n \in [0,1] : \sum_{i=1}^{n} p_i = 1$, будет выполняться неравенство $\sum_{i=1}^{n} p_i\,|C(a_i)| \geqslant H[p_1,p_2,\ldots,p_n]$
2. $\forall\, p_1,p_2,\ldots,p_n \in [0,1] : \sum_{i=1}^{n} p_i = 1$, где p_i — частота встречаемости символа a_i в языке, существует префиксный код $C : \Sigma \to \{0,1\}^* : \sum_{i=1}^{n} p_i\,|C(a_i)| < H[p_1,p_2,\ldots,p_n] + 1$.

Доказательство.

1. Из неравенства Крафта имеем $\sum\limits_{i=1}^{n} 2^{-|C(a_i)|} \leqslant 1$.

   Тогда $\sum\limits_{i=1}^{n} p_i \log_2 \frac{1}{2^{-|C(a_i)|}} \geqslant \sum\limits_{i=1}^{n} p_i \log_2 \frac{1}{p_i} = H[p_1,p_2,\ldots,p_n]$

2. Скажем, что $l_i = \lceil \log_2 \frac{1}{p_i} \rceil \ \ \forall\, i \in [n] \implies \sum\limits_{i=1}^{n} 2^{-l_i} \leqslant \sum\limits_{i=1}^{n} p_i = 1$

   Тогда по второй части леммы Крафта существует код с такими длинами. Осталось доказать, что требуемое неравенство выполнится:

   $l_i = \lceil \log_2 \frac{1}{p_i} \rceil \implies l_i - \frac{1}{p_i} < 1$

   $\sum\limits_{i=1}^{n} p_i l_i - \sum\limits_{i=1}^{n} p_i \log_2 \frac{1}{p_i} = \sum\limits_{i=1}^{n} p_i (l_i - \log_2 \frac{1}{p_i}) < \sum\limits_{i=1}^{n} p_i = 1 \implies \sum\limits_{i=1}^{n} p_i l_i < H[p_1,p_2,\ldots,p_n] + 1$

□

4.7. Закон больших чисел для распределения Бернулли

***Определение* 4.7.1.**

$X : \Omega \to \mathbb{R}$

$\Pr[X = 1] = p, \ \ \Pr[X = 0] = 1 - p$

$\mathbb{E}[X] = \Pr[X = 1] = p$

Теорема 4.7.1.

$X_1, X_2, \ldots, X_N : \Omega \to \{0, 1\}$ — независимые случайные величины.

$\mathbb{E}[X_1] = \mathbb{E}[X_2] = \ldots = \mathbb{E}[X_N] = \mu$

Тогда $\Pr\left[\left|\frac{X_1+X_2+\ldots+X_N}{N} - \mu\right| \geqslant \varepsilon\right] \leqslant 2C^N$, где $0 < C < 1$ и C зависит от μ и ε.

Доказательство.

Пусть $\exists\, 0 < C_1, C_2 < 1$, что выполняется следующее:

$\Pr\left[\frac{X_1+X_2+\ldots+X_N}{N} \geqslant \mu + \varepsilon\right] \leqslant 2C_1^N$

$\Pr\left[\frac{X_1+X_2+\ldots+X_N}{N} \leqslant \mu - \varepsilon\right] \leqslant 2C_2^N$

Тогда $C = \max(C_1, C_2)$

Рассмотрим первый случай.

$A = X_1 X_2 \ldots X_N$

$\Pr\left[\frac{X_1+X_2+\ldots+X_N}{N} \geqslant \mu + \varepsilon\right] \leqslant 2C_1^N = \sum\limits_{\omega(S) \geqslant (\mu+\varepsilon)N}^{S \in \{0,1\}^N} \Pr[A = S] \leqslant$

Пояснение: $\omega(S)$ — число единиц в строке S

Пусть $Y_1, Y_2, \ldots, Y_N$ — независимые случайные величины, $\Pr[Y_i = 1] = \mu + \varepsilon$.

$B = Y_1 Y_2 \ldots Y_N$

Будем рассматривать такие строчки S, что $\omega(S) \geqslant (\mu + \varepsilon)N$. Тогда:

$\Pr[A = S] = \mu^{\omega(S)} \cdot (1 - \mu)^{N - \omega(S)}$

$\Pr[B = S] = (\mu + \varepsilon)^{\omega(S)} \cdot (1 - \mu - \varepsilon)^{N - \omega(S)}$

$\Pr[A = S] = \Pr[B = S] \cdot \left(\frac{\mu}{\mu+\varepsilon}\right)^{\omega(S)} \left(\frac{1-\mu}{1-\mu-\varepsilon}\right)^{N-\omega(S)} \leqslant \Pr[B = S] \cdot \left(\frac{\mu}{\mu+\varepsilon}\right)^{(\mu+\varepsilon)N} \left(\frac{1-\mu}{1-\mu-\varepsilon}\right)^{(1-\mu-\varepsilon)N}$

Отсюда получаем $\Pr[A = S] \leqslant \Pr[B = S] \cdot \left(\left(\frac{\mu}{\mu+\varepsilon}\right)^{\mu+\varepsilon} \left(\frac{1-\mu}{1-\mu-\varepsilon}\right)^{1-\mu-\varepsilon} \right)^N = \Pr[B = S] \cdot C_1^N$.

$\leqslant \sum\limits_{\omega(S)\geqslant(\mu+\varepsilon)N}^{S\in\{0,1\}^N} (\Pr[B = S] \cdot C_1^N) = C_1^N \cdot \left(\sum\limits_{\omega(S)\geqslant(\mu+\varepsilon)N}^{S\in\{0,1\}^N} \Pr[B = S] \right) \leqslant C_1^N$

Осталось показать, что $C_1 = \left(\frac{\mu}{\mu+\varepsilon}\right)^{\mu+\varepsilon} \left(\frac{1-\mu}{1-\mu-\varepsilon}\right)^{1-\mu-\varepsilon} < 1$.

Заметим, что это то же, что и $(\mu+\varepsilon)\log\frac{\mu}{\mu+\varepsilon} + (1-\mu-\varepsilon)\log\frac{1-\mu}{1-\mu-\varepsilon} < 0$

Следовательно, $(\mu+\varepsilon)\log\frac{1}{\mu+\varepsilon} + (1-\mu-\varepsilon)\log\frac{1}{1-\mu-\varepsilon} < (\mu+\varepsilon)\log\frac{1}{\mu} + (1-\mu-\varepsilon)\log\frac{1}{1-\mu}$, что верно по лемме о чужих логарифмах (а также в силу того, что log — строго вогнутая функция, а $\varepsilon > 0$).

Случай второго неравенства рассматривается аналогично. □

4.8. Условные вероятности

***Определение* 4.8.1.**

$(\Omega, \Pr)$ — конечное вероятностное пространство.

$A, B \subseteq \Omega,\ \Pr(B) \neq 0$

Условной вероятностью $\Pr(A \mid B) = \frac{\Pr(A\cap B)}{\Pr(B)}$ называется вероятность наступления события A при условии уже произошедшего события B.

Определение* 4.8.2.** (Формула полной вероятности***)

B_1, B_2, ..., $B_n \subseteq \Omega$

$\Pr(B_i) > 0,\ B_i \cap B_j \neq \varnothing$

Тогда $\Pr(A) = \sum\limits_{i=1}^{n} \Pr(A \cap B_i) = \sum\limits_{i=1}^{n} \Pr(A \mid B_i) \cdot \Pr(B_i)$

Теорема 4.8.1. (**Формула Байеса**)

$\Pr(A \mid B) = \frac{\Pr(A\cap B)}{\Pr(B)} = \frac{\Pr(B \mid A)\cdot\Pr(A)}{\Pr(B)}$

В частности, $\Pr(A \mid B)\Pr(B) = \Pr(B \mid A)\Pr(A)$

4.9. Дисперсия

***Определение* 4.9.1.**

$(\Omega, \Pr)$ — конечное вероятностное пространство.

X — случайная величина.

$D[X] = \mathbb{E}\big[(X - E[X])^2\big]$ — дисперсия случайной величины.

$D[X] = \mathbb{E}\big[(X - E[X])^2\big] = \mathbb{E}\big[X^2 + E[X]^2 - 2X\mathbb{E}[X]\big] = \mathbb{E}[X^2] + \mathbb{E}[X]^2 - 2\mathbb{E}[X]^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

Из определения матожидания очевидно, что $\mathbb{E}[X^2] - \mathbb{E}[X]^2 \geqslant 0 \implies \mathbb{E}[X^2] \geqslant \mathbb{E}[X]^2$.

Лемма.

X_1, X_2, ..., X_n — попарно независимые случайные величины.

Тогда $D[X_1 + X_2 + ... + X_n] = D[X_1] + D[X_2] + ... + D[X_n]$

Замечание. $D[\alpha X] = \alpha^2 D[X]$

Доказательство.

$D[X_1 + X_2 + ... + X_n] = \mathbb{E}[(X_1 + X_2 + ... + X_n)^2] - \mathbb{E}[X_1 + X_2 + ... + X_n]^2 =$

$= \left(\sum_{i=1}^{n} \mathbb{E}[X_i^2] + 2\sum_{i<j} \mathbb{E}[X_i \cdot X_j]\right) - \left(\sum_{i=1}^{n} (\mathbb{E}[X_i])^2 + 2\sum_{i<j} \mathbb{E}[X_i] \cdot \mathbb{E}[X_j]\right)$

Заметим, что $\forall\, i \neq j : \mathbb{E}[X_i \cdot X_j] = \mathbb{E}[X_i] \cdot \mathbb{E}[X_j]$ (поскольку случайные величины независимы). Следовательно, получаем, что:

$D[X_1 + X_2 + ... + X_n] = \sum_{i=1}^{n} \mathbb{E}[X_i^2] - \sum_{i=1}^{n} \mathbb{E}[X_i]^2 = \sum_{i=1}^{n} (\mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2) = \sum_{i=1}^{n} D[X_i]$. □

4.10. Неравенство Чебышёва

Теорема 4.10.1. (**Неравенство Чебышёва**)

$\forall C > 0 \quad \Pr[|X - \mathbb{E}[X]| \geqslant C] \leqslant \frac{D[X]}{C^2}$

Доказательство.

1) Пусть $\Pr\big[X = \mathbb{E}[X]\big] = 1$. Тогда неравенство очевидно.

2) Пусть X не всегда равен $\mathbb{E}[X]$.

Положим $Y = (X - \mathbb{E}[X])^2$, $\mathbb{E}[Y] > 0$ (следует из описанного выше условия).

В частности, $D[X] = \mathbb{E}[Y]$.

$\Pr\big[\,|X - \mathbb{E}[X]| \geqslant C\big] = \Pr\big[(X - \mathbb{E}[X])^2 \geqslant C^2\big] = \Pr\left[Y \geqslant \frac{C^2}{\mathbb{E}[Y]}\mathbb{E}[Y]\right] \leqslant \frac{\mathbb{E}[Y]}{C^2}$ (по неравенству Маркова). □

Теорема 4.10.2. (**Закон больших чисел для попарно независимых случайных величин**)

X_1, X_2, ..., X_n — попарно независимые случайные величины.

$\mathbb{E}[X_i] = \mu,\ D[X_i] = \sigma^2$

Тогда $\Pr\left[\left|\frac{X_1+X_2+...+X_n}{n} - \mu\right| \geqslant \varepsilon\right] \leqslant \frac{\sigma^2}{n\varepsilon^2}$

Доказательство.

$Y = \frac{X_1+X_2+...+X_n}{n},\ \mathbb{E}[Y] = \mu$

$\Pr\big[\,|Y - \mu| \geqslant \varepsilon\big] = \Pr\big[\,|Y - \mathbb{E}[Y]| \geqslant \varepsilon\big] \leqslant \frac{D[Y]}{\varepsilon^2}$ (по неравенству Чебышёва)

$D[Y] = \sum_{i=1}^{n} D\big[\frac{X_i}{n}\big] = \sum_{i=1}^{n} \frac{1}{n^2} D[X_i] = \sum_{i=1}^{n} \frac{\sigma^2}{n^2} = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$

Таким образом, $\Pr\big[\,|Y - \mu| \geqslant \varepsilon\big] \leqslant \frac{\sigma^2}{n\varepsilon^2} \implies \Pr\left[\left|\frac{X_1+X_2+...+X_n}{n} - \mu\right| \geqslant \varepsilon\right] \leqslant \frac{\sigma^2}{n\varepsilon^2}$ □

5. Коды, исправляющие ошибки

5.1. Игра с угадыванием числа

Боб загадал число из множества $[n]$, Алиса может задавать ему вопросы вида «принадлежит ли загаданное число множеству $\{7, 13, 42\}$?» Какое минимальное число вопросов понадобится Алисе, чтобы угадать число?

Представим число в двоичной системе счисления. Теперь Алисе достаточно задать $\lceil \log_2 n \rceil$ вопросов вида «является ли i-тый бит загаданного числа единицей?»

5.2. Игра с одной ошибкой

Теперь Бобу разрешено дать до одного неверного ответа.

Пусть Алиса задала $m = \lceil \log_2 n \rceil$ вопросов про каждый двоичный разряд числа. Теперь осталось $m+1$ чисел-кандидатов: само число K, найденное Алисой, и m чисел K_i с i-тым битом, отличающимся от сказанным Бобом. Теперь необходимо $\lceil \log_2(m+1) \rceil$ дополнительных вопросов, чтобы однозначно определить число.

Немного изменим формат игры. Алиса заранее сообщит Бобу, какие вопросы она будет задавать при всех возможных исходах. Теперь Бобу достаточно записать ответы в виде бинарной строки и отправить строку Алисе.

5.3. Код Хэмминга

***Определение* 5.3.1.** Код $C : \{0,1\}^k \to \{0,1\}^n$.

***Определение* 5.3.2.** $x, y \in \{0,1\}^n$.

$\delta(x, y) = |\{i | x_i \neq y_i\}|$ — расстояние Хэмминга.

***Определение* 5.3.3.** Код $C : \{0,1\}^k \to \{0,1\}^n$ имеет расстояние $\geqslant d$, если $\forall x, y \in \{0,1\}^k, x \neq y$ выполняется $\delta(x, y) \geqslant d$.

***Определение* 5.3.4.** Код исправляет t ошибок, если его расстояние $\geqslant 2t + 1$.

6. Матричные игры

6.1. Неразрешимость линейных сравнений

Лемма (Фаркаша)**.**

Если система из m линейных неравенств не имеет решений, то $\exists q_1,q_2,\ldots,q_m \geqslant 0$ Такие, что если сложить неравенства, домноженные на коэффициенты $q_1, q_2, \ldots, q_m$, то получится $0 \geqslant 1$

$$\forall j \sum_{i=1}^{m} q_i a_{i,j} = 0, \sum_{i=1}^{m} q_i b_i = 1$$

Доказательство.

Индукция по числу переменных x.

База: $n = 1$.

$$\begin{cases} x_1 \geqslant c_1 \\ \ldots \\ x_1 \geqslant c_k \\ x_1 \leqslant d_1 \\ \ldots \\ x_1 \leqslant d_l \end{cases}$$

Очевидно, что такая система будет невыполнима только если $\max(c) > \min(d)$. Иначе говоря, система не имеет рещений, если $\exists c_j, d_t$ такие, что $c_j > d_t$

$x_1 \geqslant cj > d_t \geqslant x_1$ – действительно, если так, то получаем неавенство $x_1 > x_1$, решений нет. Ещё можно показать это так: $x_1 \geqslant c_j,\ -x_1 \geqslant -d_t \Rightarrow 0 \geqslant c_j - d_t > 0$

Переход: $n \to n + 1$. Разделим все неравеснтва на 3 группы:

1. Неравенства, в которые x_1 не входит
2. m неравенств вида $x_1 \geqslant f_i(x_2, x_3, \ldots, x_n)$, $i \in [m]$
3. l неравенств вида $x_1 \leqslant g_j(x_2, x_3, \ldots, x_n)$, $j \in [l]$

Напишем эквивалентную систему неравенств. Первая группа неравенств останется без изменений. Вторую и третью совместим, записав в новую систему все неравенства вида $f_i \leqslant g_j, \forall i, j : i \in [m], j \in [l]$. Покажем теперь, что исходная система выполнима тогда и только тогда, когда выполнима полученная. В одну сторону очевидно. Пусть исходная выполнима, тогда подставим выполняющий набор в полученную систему и она выполнится. В другую сторону. Пусть полученная система имеет решение. Тогда подставим в неё выполняющие значения и найдем $a = \max(f_i)$ и $b = \min(g_j)$. Для них известно, что $a \leqslant b$ т.к соотношение «меньше» было введено на всех парах f_i, g_j. Выберем любое число из отрезка $[a, b]$, подставим на место x_1. Получили выполняющий набор для исходной системы.

Вернемся к поиску коэффициентов. Теперь мы знаем, что если исходная система невыполнима, то и полученная тоже. Тогда по предположению индукции (мы только что избавились от переменнной), значит, для полученной систмемы есть нужный набор коэффициентов. Значит, он есть и для исходной системы и его можно вывести из коэффициентов для полученной системы. □

Лемма.

Пусть A – вещественная матрица $m \times n$, тогда выполнится ровно одно из двух:

1. $\exists p_1,p_2,\ldots,p_m \geqslant 0 : \forall j \in [n] \sum_{i=1}^{m} A_{i,j}p_i \geqslant 0$
2. $\exists q_1,q_2,\ldots,q_n \geqslant 0 : \forall i \in [m] \sum_{j=1}^{n} A_{i,j}q_j < 0$

Замечание. Можно переносить строгость между утверждениями. (Проверяется переворотом матрицы и домножением на -1)

Доказательство.

Пусть 1. не выполнится ни для каких p, подходящих под условия выше. Тогда составим следующую систему уравнений, о которой будем знать, что она невыполнима:

$$\begin{cases} \sum_{i=1}^{m} A_{i,j}p_i \geqslant 0 & \forall j \in [n] \text{ сопоставим этим неравнествам коэффициенты} q_j \\ p_i \geqslant 0 & \forall i \in [m] \text{ этим } - \alpha_i \\ \sum_{i=1}^{m} p_i \geqslant 1 & \text{а этим } - \gamma \end{cases}$$

Раз такая система невыполнима, то найдутся такие коэффициенты из леммы Фаркаша, что сумма всех неравенств выше даст неравенство $0 \geqslant 1$. Найдем эти коэффициенты. Начнем с очевидного. А именно, $\gamma = 1$ (нам больше неоткуда получить единицу в правой части уравнения). Выпишем теперь отдельно сумму всех неравенств (раз мы в результате должны получить неравенство $0 \geqslant 1$, то сумма всех левых частей должна быть равна нулю):

$$\sum_{j=1}^{n} q_j \sum_{i=1}^{m} A_{i,j}p_i + \sum_{i=1}^{m} \alpha_i p_i + \sum_{i=1}^{m} p_i = 0$$

$$\sum_{i=1}^{m} p_i \sum_{j=1}^{n} A_{i,j}q_j + \sum_{i=1}^{m} \alpha_i p_i + \sum_{i=1}^{m} p_i = 0$$

$$\sum_{i=1}^{m} p_i (\sum_{j=1}^{n} A_{i,j}q_j + \alpha_i + 1) = 0$$

Вспомним теперь, что это равенство должно быть нулем при любых значениях переменных p. Значит, каждый коэффициент при p_i должен быть равен 0. Отсюда:

$$\sum_{j=1}^{n} A_{i,j}q_j + \alpha_i + 1 = 0, \ \forall i \in [m]$$

$$\sum_{j=1}^{n} A_{i,j}q_j = -\alpha_i - 1, \ \forall i \in [m]$$

Заметим, что $-\alpha_i - 1 < 0$, $\forall i \in [m]$ (по лемме Фаркаша, коэффициенты неотрицательны $\Rightarrow -\alpha_i \geqslant 0$). Отсюда получаем выполнение пункта 2 т.к. мы нашли неотрицательные коэффициенты q такие, что постолбцовые суммы оказались меньше нуля.

□

6.2. Основные понятия матричных игр

***Определение* 6.2.1.**

A – платежная матрица, на которой ведётся игра. Тогда если столбцовый игрок выбрал столбец выбрал столбец i, а строчный игрок выбрал строку j, то столбцовый игрок платит строчному $A_{i,j}$ (это число может быть в том числе и отрицательным, тогда строчный игрок платит столбцовому $-A_{i,j}$).

***Определение* 6.2.2** (Стратегия)**.**

1. Чистая. В таком случае строчный/столбцовый игрок всегда выбирает только один столбец/одну строку. Будем обозначать (i) чистую стратегию, которая выбирает i-й столбец/строчку.

2. Смешанная. Тогда она представляет собой распределение вероятностей выбора того или иного столбца или строчки.

   Иначе говоря, стратегия столбцового (первого) игрока — p_1,p_2,$\ldots$,p_m, $p_i \geqslant 0$, $\sum_{i=1}^{m} p_i = 1$

   Строчного (второго) — q_1,q_2,$\ldots$,q_n, $q_i \geqslant 0$, $\sum_{i=1}^{n} q_i = 1$

Замечание. Игроки играют независимо.

***Определение* 6.2.3.**

$A(p,q)$ – математическое ожидание выигрыша I игрока, если он играет по стратегии p, а игрок II – по стратегии q.

$$A(p,q) = \sum_{i=1}^{m}\sum_{j=1}^{n} p_i q_j a_{i,j} = \sum_{j=1}^{n} q_j (\sum_{i=1}^{m} p_i a_{i,j}) = \sum_{j=1}^{n} q_j A(p,(j))$$

Последний переход нетривиален, поясню. Давайте внимательно посмотрим на $\sum_{i=1}^{m} p_i a_{i,j}$ и поймем, что она эквивалентна тому, что стратегия q – чистая, выбирающая строчку j. При этом стратегия p – по-прежнему данная нам смешанная. (Если всё ещё нет, то представьте, что у вас есть некоторая p, а q при этом – все нули и единичка на позиции j. Посчитайте мат. ожидание. Оно должно получиться равным именно такой сумме).

Аналогично

$$A(p,q) = \sum_{i=1}^{m}\sum_{j=1}^{n} p_i q_j a_{i,j} = \sum_{i=1}^{m} p_i (\sum_{j=1}^{n} q_j a_{i,j}) = \sum_{j=1}^{m} p_i A((i),q)$$

***Определение* 6.2.4.**

Стратегия p гарантирует выигрыш хотя бы C, если $\forall q\ A(p,q) \geqslant C$.

Свойства стратегий.

1. Стратегия p гарантирует выигрыш хотя бы C тогда и только тогда, когда $\forall$ чистой стратегии q $A(p,q) \geqslant C$.

   Доказательство.

   В одну сторону очевидно. Если p гарантирует выигрыш хотя бы C для любой стратегии q, то для чистых тем более.

   В обратную: $A(p,q) = \sum_{j=1}^{n} q_j A(p,(j)) \geqslant \sum_{j=1}^{n} q_j C = C \sum_{j=1}^{n} q_j = C$ □

2. а) Либо у первого игрока есть стратегия, гарантирующая ему выиграть $\geqslant 0$, b) либо у второго игрока есть стратегия, гарантирующая ему проиграть < 0.

Доказательство.
Ну в самом деле, посмотрим, что мы хотим получить:

a) $\exists p : \forall j \ A(p,(j)) \geqslant 0$
$\sum_{i=1}^{m} p_i A_{i,j} \geqslant 0, \sum p_i = 1$
b) $\exists q : \forall i \ A((i),q) < 0$
$\sum_{j=1}^{n} q_j A_{i,j} < 0, \sum q_i = 1$

Заметим, что мы это уже видели в этой лемме. Действительно, по ней мы знаем, что нужный знак можно получить только одним из двух способов. Но заметим, что сумма наших вероятностей должна быть равна единичке. Мы можем получить такие из коэффициентов, полученных по лемме, домножением на константу. □

3. $\forall \mu \in \mathbb{R}$ либо у первого игрока есть стратегия, гарантирующая ему выиграть $\geqslant \mu$, либо у второго игрока есть стратегия, гарантирующая ему проиграть $< \mu$

Доказательство.
Из всех элементов вычтем μ и воспользуемся предыдущим свойством. □

***Определение* 6.2.5.**
$C \in \mathbb{R}$ называется ценой игры, если $\exists p^*, q^*$ такие, что p^* гарантирует первому игроку выиграть не менее C, а q^* гарантирует второму игроку проиграть не более C.

4. Если C_1, C_2 – цены игры, то $C_1 = C_2$.
Рассмотрим, не теряя общности $C_1 < C_2$. Тогда пусть стратегии p^*, q^* дают нам цену игры C_1, а p^{**}, q^{**} – C_2. Тогда $C_1 \geqslant A(p^{**}, q^*) \geqslant C_2$. Противоречие.

6.3. Теорема Фон-Неймана

Теорема 6.3.1 (Фон-Неймана)**.**
В любой матричной игре $\exists!$ цена игры.

Доказательство.
Единственность была доказана выше. Докажем существование.

$$f(q) = \max_{\text{p – стратегия}} A(p,q) = \max_{\text{p – стратегия}} \sum_{i=1}^{m} p_i A((i),q) = \max_{\text{p – чистая стратегия}} A(p,q)$$

$$g(p) = \min_{\text{q – стратегия}} A(p,q) = \min_{\text{q – стратегия}} \sum_{j=1}^{n} q_j A(p,(j)) = \min_{\text{q – чистая стратегия}} A(p,q)$$

$$\forall p,q \ f(q) \geqslant A(p,q) \geqslant g(p)$$
$$f(q) \geqslant \sup g(p)$$
$$\inf f(q) \geqslant \sup g(p)$$

Докажем, что нам на самом деле равенство. Пусть $\inf f(q) > \mu$ и $\mu > \sup g(p)$. Но по свойству 3 либо $\exists p : g(p) \geqslant \mu$, либо $\exists q : f(q) < \mu$. Противоречие. Следовательно, $\inf f(q) = \sup g(p) = C$. Будем пока просто утверждать, что найденное C – цена игры.

До сих пор было непонятно, почему же мы вообще имеем право писать max и min. Вдруг они просто не достигаются? По свойству 3 либо $\exists p : g(p) \geqslant C$, либо $\exists q : f(q) < C \Rightarrow \exists p^* : g(p^*) = C$. Аналогично по этому же свойству либо $\exists p : g(p) > C$, либо $\exists q : f(q) \leqslant C \Rightarrow \exists q^* : f(q^*) = C$. □

7. Числа Рамсея

7.1. Числа Рамсея

***Определение* 7.1.1.**

$R(m,n)$ — минимальное натуральное число N, что если в полном графе из N вершин все рёбра покрасить в два цвета, то либо найдётся m вершин, что все рёбра между ними покрашены в первый цвет, либо найдётся n вершин, что все рёбра между ними покрашены во второй цвет.

$R(m,n)$ — число Рамсея.

Пример.

$R(3,3) = 6$ (разбиралась на практике)

Замечание.

Неочевидно, что $R(m,n) < +\infty$.

Свойства.

1. $R(2,n) = n$
2. $R(m,n) = R(n,m)$
3. $R(m,n) \leqslant R(m,n-1) + R(n-1,m)$ при $m,n \geqslant 3$

   Доказательство.

   Рассмотрим полный граф на $R(m,n-1) + R(m-1,n)$ вершинах (предполагаем, что эта величина конечна).

   Рассмотрим одну вершину v. Она соединена с C_1 вершинами ребром первого цвета и с C_2 вершинами рёбрами второго цвета.

   Утверждение. Либо $|C_1| \geqslant R(m-1,n)$, либо $|C_2| \geqslant R(m,n-1)$.

   Доказательство. От противного. $R(m-1,n) + R(m,n-1) - 1 = |C_1| + |C_2| \leqslant R(m-1,n) + R(m,n-1) - 2$

   Рассмотрим случаи:

   1) $|C_1| \geqslant R(m-1,n) \implies$ в C_1 либо есть n вершин, соединённых рёбрами цвета 1, либо $m-1$ вершина, соединённая рёбрами цвета 2.

   2) $|C_2| \geqslant R(m,n-1) \implies$ либо в C_1 есть m вершин, соединённых рёбрами цвета 1, либо $n-1$ вершина, соединённая рёбрами цвета 2. □
4. $R(m,n) \leqslant \binom{m+n-2}{m-1} = \binom{m+n-2}{n-1}$

   Доказательство.

   Индукция по m и n.

   База: $R(2,n) = n \leqslant \binom{n}{1} = n$

   Переход: Пусть $m,n \geqslant 3$

   $R(m,n) \leqslant R(m-1,n) + R(m,n-1) \leqslant \binom{m+n-1}{m-2} + \binom{m+n-1}{m-1} = \binom{m+n-2}{m-1}$ □
5. $R(n,n) \leqslant \binom{2n-2}{n-1} \leqslant 4^{n-1}$

6. $R(n,n) \geqslant (\sqrt{2})^n$

Доказательство.
От противного. Рассмотрим граф из $N < (\sqrt{2})^n$ вершин.

Рассмотрим случайную раскраску в два цвета.

Посчитаем вероятность того, что найдётся n вершин, все рёбра между которыми покрашены в один цвет.

$\Pr[\exists] = \Pr[\bigcup_{A\subseteq[N]}^{|A|=n} \text{в } A \text{ все рёбра одного цвета}] \leqslant \sum_{A\subseteq[N]}^{|A|=n} \Pr[\text{в } A \text{ все рёбра одного цвета}] = \binom{N}{n}\frac{2}{2^{n(n-1)/2}} \leqslant \frac{N^n}{n!}\cdot\frac{2}{2^{n(n-1)/2}} \leqslant \frac{2^{n^2/2}\cdot 2}{n!\cdot 2^{n^2/2-n/2}} = \frac{2^{n/2+1}}{n!} < 1 \implies$ для $n \geqslant 3$ выполняется.

То есть при $n \geqslant 3$ с ненулевой вероятностью найдётся полный граф на N вершинах, который нельзя раскрасить в два цвета описанным образом.

□

***Определение* 7.1.2.**
$R(S,m,n)$ — минимальное натуральное число N, что если все S-элементные подмножества $[N]$ покрасить в 2 цвета, то либо найдётся m-элементное подмножество, все его S-элементные подмножества покрашены в цвет 1, либо найдётся n-элементное подмножество, все его S-элементные подмножества покрашены в цвет 2.

Свойства.

7. $R(1,m,n) = m+n-1$.

8. $R(s,m,n) \leqslant R(s-1, R(s,m-1,n), R(s,m,n-1)) + 1$

Доказательство.
Выделим элемент $v \in [N]$.

Раскрасим все $(s-1)$-элементные подмножества $[N]\setminus\{v\}$ в 2 цвета: в тот цвет, в который покрашено s-элементное подмножество, получается добавлением v

По индукционном предположению, найдётся либо 1) $R(s,m-1,n)$ элементов, в которых $(s-1)$-элементные подмножества покрашены цвет 1, либо 2) $R(s,m,n-1)$ элементов, все $(s-1)$-элементные подмножества покрашены в цвет 2.

1) $R(s,m-1,n) \to \begin{cases} m-1 \text{ элемент: все } s\text{-ки покрашены в цвет 1 добавлением } v \implies \text{OK} \\ n \text{ элементов: все } s\text{-ки покрашены в цвет 2} \end{cases}$

2)$R(s,m,n-1) \to \begin{cases} m \text{ элементов: все } s\text{-ки покрашены в цвет 1} \\ n-1 \text{ элемент: все } s\text{-ки покрашены в цвет 2 добавлением } v \end{cases}$ □

***Определение* 7.1.3.**
$R_k(s,\, m_1,\, m_2,\, ...,\, m_k)$ — обобщение чисел Рамсея на k цветов.
$R_k(1,\, m_1,\, m_2,\, ...,\, m_k) = m_1 + m_2 + ... + m_k - k + 1$
$R_k(s,\, m_1,\, m_2,\, ...,\, m_k) \leqslant R_k(s-1,\, R_k(s,\, m_1-1,\, m_2,\, ...,\, m_k),\, ...,\, R_k(s,\, m_1,\, m_2,\, ...,\, m_k-1)) + 1$

Пример.

1. (Теорема Шура).

   $\forall r \in \mathbb{N}\ \exists n \in \mathbb{N}$: если раскрасить числа $[n]$ в r цветов, то обязательно найдутся $x,y,z \in [n]$: они одного цвета и $x+y=z$

Доказательство.

$n = R_r(2, \overbrace{3, 3, ..., 3}^{r \text{ times}})$

$\varphi : [n] \to [r], \quad \varphi((x, y)) = \varphi(|x - y|)$

$\exists$ цвет и $\exists x, y, z : \varphi(|x - y|) = \varphi(|y - z|) = \varphi(|x - z|)$

Пусть $x < y < z$. Тогда $(y - x) + (z - y) = z - x$.

TODO Осознать это □

2. (Теорема Эрдеша-Секереша)

   $\forall m \in \mathbb{N}\ \exists N \in \mathbb{N}$: из любых N точек общего положения (никакие три точки не лежат на одной прямой) на плоскости можно выбрать выпуклый m угольник.

   Доказательство.

   $N = R(3, m, m)$

   Цвет тройки $\varphi(A, B, C) =$ числу точек внутри ABC (`mod` 2).

   Т.е. найдутся m точек, что все тройки покрашены в один цвет.

   Отсюда можем вывести противоречие.

   TODO Нужен рисунок и чей-нибудь конспект. □

Лемма. (**Кёнига**)

В любом бесконечном корневом дереве конечной степени найдётся бесконечная ветвь.

Доказательство.

Найдём сына бесконечного размера (такой обязательно есть) и перейдём в него. □

Теорема 7.1.1. (**Бесконечная теорема Рамсея**)

Рёбра полного бесконечного графа покрашены в r цветов. Тогда найдётся ∞ множество вершин, что все рёбра между ними покрашены в один цвет.

Доказательство.

TODO □