# Задачи к курсу «Среднее время работы программ и колмогоровская сложность»

## Часть 1: асимптотические обозначения

*С практической точки зрения разница между полиномом и экспонентой зачастую важнее различия между конечным и бесконечным*

Джек Эдмондс, 1965

Пусть $f\colon \mathbb{N} \to \mathbb{N}$ и $g\colon \mathbb{N} \to \mathbb{N}$ суть две функции. Для сравнения скорости их роста используются следующие асимптотические обозначения:

- $f(n) = O(g(n))$, если для некоторой константы $C > 0$ найдётся $N$, такое что при всех $n > N$ выполнено $f(n) < Cg(n)$;
- $f(n) = \Omega(g(n))$, если для некоторой константы $c > 0$ найдётся $N$, такое что при всех $n > N$ выполнено $f(n) > cg(n)$;
- $f(n) = \Theta(g(n))$, если для некоторых констант $C > c > 0$ найдётся $N$, такое что при всех $n > N$ выполнено $Cg(n) > f(n) > cg(n)$;
- $f(n) = o(g(n))$, если для любой константы $c > 0$ найдётся $N$, такое что при всех $n > N$ выполнено $f(n) < cg(n)$;
- $f(n) = \omega(g(n))$, если для любой константы $C > 0$ найдётся $N$, такое что при всех $n > N$ выполнено $f(n) > Cg(n)$;
- $f(n) = \mathrm{poly}(g(n))$, если для некоторых констант $C > 0$ и $k > 0$ найдётся $N$, такое что при всех $n > N$ выполнено $f(n) < C(g(n))^k$.

**Задача 1.** В каких определениях можно убрать условие на положительность констант? В каких можно вычеркнуть «найдётся $N$, такое что»?

**Задача 2.** Какие из асимптотик совместимы друг с другом? Например, может ли одновременно быть $f(n) = o(g(n))$ и $f(n) = O(g(n))$?

**Задача 3.** Докажите, что:

а) $n^a = o(n^b)$ при $a < b$;

б) Если $f(n)$ и $g(n)$ — многочлены одинаковой степени, то $f(n) = \Theta(g(n))$;

в) $n^a = o(2^n)$;

г) $\log^a n = o(n^b)$.

**Задача 4.** Бывают ли такие $f$ и $g$, для которых ни одно из соотношений не выполнено?

**Задача 5.** Пусть $f_1(n) = o(g(n))$, а $f_2(n) = O(g(n))$. Что можно сказать про $f_1(n) + f_2(n)$? А про $f_1(n) \cdot f_2(n)$? Составьте полные «таблицы сложения и умножения».

**Задача 6.** Пусть $f(n) = o(g(n))$, а $g(n) = O(h(n))$. Что можно сказать про асимптотику $f$ в терминах $h$? Составьте полную «таблицу композиции».

**Задача 7.** Пусть $f(n) = 2^{n+\varepsilon(n)}$, где $\varepsilon(n) = o(1)$. Верно ли, что $f(n) = O(2^n)$?

## Часть 2: математическое ожидание

*Когда Билл Гейтс заходит в бар, все его посетители становятся в среднем миллиардерами*

Интернет-фольклор

Пусть дано множество $A$ из $N$ элементов. Для удобства мы будем считать, что $N = 2^n$. *Распределением вероятностей* на $A$ называется набор $(p_1, \ldots, p_N)$, такой что все $p_i \geqslant 0$ и $\sum p_i = 1$. Если все $p_i$ равны по $\frac{1}{N}$, то распределение называется *равномерным*. Вероятностью события (подмножества) $B \subset A$ называется величина $\sum_{a_i \in B} p_i$. Множество $A$ вместе с распределением вероятностей называется *вероятностным пространством*. Любая функция из $A$ в $\mathbb{N}$ называется (целочисленной неотрицательной) *случайной величиной*. Случайные величины можно складывать и умножать на число или друг на друга, как любые другие функции. *Математическим ожиданием* случайной величины $\xi$ называется величина $\sum p_i \xi(a_i)$. Обозначение: $\mathbb{E}\xi$.

**Задача 8.** Докажите, что математическое ожидание линейно: $\mathbb{E}(\xi+\eta) = \mathbb{E}\xi+\mathbb{E}\eta$ и $\mathbb{E}\alpha\xi = \alpha\mathbb{E}\xi$ для любого числа $\alpha$.

**Задача 9.** Придумайте такую $\xi$, что $\mathbb{E}\xi = O(n)$, а $\mathbb{E}(\xi^2) = \Omega(2^n)$.

**Задача 10.** Покажите, что при подходящем подборе распределения $\mathbb{E}\xi$ может принять любое значение между $\min\xi$ и $\max\xi$.

**Задача 11. (Неравенства Маркова)** Напомним, что мы рассматриваем только неотрицательные случайные величины. Докажите, что вероятность того, что $\xi > k\mathbb{E}\xi$, меньше $\frac{1}{k}$.

## Часть 3: колмогоровская сложность

.

*Чему равно минимальное натуральное число, которое нельзя описать фразой на русском языке, состоящей не более чем из семнадцати слов?*

Парадокс Берри

Пусть $V$ — некоторый алгоритм, получающий на вход два слова из нулей и единиц и возвращающий третье слово. Колмогоровской сложностью $KS_V(x|y)$ слова $x$ при условии $y$ и способе описания $V$ называется длина кратчайшего слова $p$, такого что $V(p,y) = x$.

**Задача 12. (Теорема Колмогорова–Соломонова)** Пусть алгоритм $U$ работает так: свой первый аргумент $q$ он понимает как пару $(v,p)$, где $v$ — код программы $V$. Затем он возвращает $V(p,y)$, где $y$ — второй вход. Докажите, что для любого $V$ найдётся константа $c$, такая что при всех $x$ и $y$ выполнено

$$KS_U(x|y) \leqslant KS_V(x|y) + c.$$

Такой алгоритм $U$ мы будем называть *универсальным способом описания*, в дальнейшем под $KS$ будем понимать $KS_U$. Под безусловной колмогоровской сложностью $KS(x)$ понимается $KS(x|\varepsilon)$, где $\varepsilon$ — пустое слово. Все равенства будут выполнены с точностью до аддитивной константы.

**Задача 13.** Докажите, что при всех $x$ и $y$ выполнено $KS(x|y) < KS(x) + O(1)$ и $KS(x) < |x| + O(1)$, где $|x|$ — длина слова $x$.

**Задача 14.** Докажите, что для любого $n$ существует не больше $2^n - 1$ слов сложности меньше $n$. Докажите, что доля слов длины $n$ сложности $\geqslant n-k$ не меньше $1 - \frac{1}{2^k}$.

**Задача 15.** Докажите, что математическое ожидание сложности случайного слова длины $n$ превышает $\gamma n$ (для любого $\gamma \in (0,1)$ при всех достаточно больших $n$).

**Задача 16.** Докажите, что колмогоровская сложность невычислима. (Указание: формализуйте парадокс Берри).

**Задача 17.** Слово $x$ длины $n$ называется *случайным в колмогорвском смысле*, если $KS(x|n) \geqslant n$. Докажите, что достаточно длинное случайное слово не может иметь 40% нулей и 60% единиц.

**Задача 18.** Докажите, что $KS(xy) \leqslant KS(x)+KS(y)+2\log|x|+O(1)$. (Указание: дополнительное место тратится на указание раздела между описаниями $x$ и $y$).

**Задача 19.** Докажите, что при всех $x$ и $y$ выполнено $KS(x) \leqslant KS(y) + KS(x|y) + O(1)$.

**Задача 20.** Докажите, что сумма $\sum_{x\in\{0,1\}^n} 2^{-2KS(x|n)}$ ограничена некоторой константой, не зависящей от $n$.