ЛЕКЦИЯ 12

ВЕРОЯТНОСТНЫЕ ВЫЧИСЛЕНИЯ. КЛАСС BPP

1. Примеры вероятностных алгоритмов

1.1. Тест Соловея – Штриссена на простоту

Напомним некоторые определения:

Определение 46: Символ Лежандра *(здесь p — простое):*

$$\left(\frac{a}{p}\right) = \begin{cases} 1, & a \text{ — квадратичный вычет;} \\ -1, & a \text{ — квадратичный невычет;} \\ 0, & (a,p) \neq 1 \text{ (} a \text{ не является взаимно простым с } p\text{).} \end{cases}$$

Определение 47: Символ Якоби *(n — нечетное):*

$$\left(\frac{a}{n}\right), \qquad n = p_1 \dots p_k \quad \Rightarrow \quad \left(\frac{a}{p_1}\right) \cdot \dots \cdot \left(\frac{a}{p_k}\right).$$

В теории чисел часто используется следующее утверждение:

если n — простое, то $\left(\frac{a}{n}\right) = a^{\frac{n-1}{2}}$,

если n — нечетное составное, то $\Pr_a \left\{ \left(\frac{a}{n}\right) = a^{\frac{n-1}{2}} \right\} \leqslant \frac{1}{2}$.

Конструкция алгоритма:
Повторять k раз:

выбрать случайное $a \in \{1, \dots, n-1\}$;
если $\left(\frac{a}{n}\right) \neq a^{\frac{n-1}{2}}$ или НОД $(a,n) > 1$:

вернуть «составное»;
иначе:
вернуть «простое».

Если n — простое, то $\Pr\{\text{алгоритм выдал «простое»}\} = 1$.
Если n — нечетное составное, то $\Pr\{\text{алгоритм выдал «составное»}\} \geqslant 1 - 2^{-k}$.
Одно из преимуществ этого алгоритма — его полиномиальность.

1.2. Идентичность многочленов

В некоторых случаях сложные многочлены с различными коэффициентами можно представить в компактной записи (например, $(x+y)^{1000}$). По двум арифметическим схемам нужно понять, задают ли два выражения один и тот же многочлен.

Идея алгоритма: посчитать значения в случайных точках, проверить совпадения.

Возникают две проблемы:

1. Выбор множества, из которого следует брать значения.
2. Степень x^{2^n} задается арифметической схемой размера $Q(n)$. Если вместо x подставить число (например, 2), то для его записи потребуетя 2^n битов (т. е. число битов будет расти как экспонента от входа).

Обе эти проблемы решаются путем рассмотрения входа из конечного поля F_p, где $p = 2^{\text{poly}(n)}$ — размер поля. Такой размер поля позволяет на полиномиальной памяти проводить все операции за полиномиальное время.

Доказательство корректности

Док-во: Пусть d — степень многочлена, Q — ненулевой многочлен от k переменных. Тогда

$$\Pr_{x \in (F_p)^k}\{Q(x) \neq 0\} \geqslant \left(1 - \frac{d}{p}\right)^k \approx 1 - \frac{dk}{p}.$$

Например, подойдет значение $p = 2\,dk$.

Это можно доказать с помощью индукции по k.

База индукции: возьмем $k = 1$. Тогда из основной теоремы алгебры следует, что ненулевой многочлен степени d над полем имеет не более d корней.

$$\Pr\{Q(x) \neq 0\} \geqslant \frac{p-d}{p} = 1 - \frac{d}{p}.$$

Переход: многочлен от y, $x_1, \ldots, x_k$ можно рассматривать как многочлен от y с коэффициентами-многочленами от $x_1, \ldots, x_k$. Хотя бы один из этих коэффициентов будет ненулевым. Следовательно, с вероятностью $p' \geqslant \left(1 - \frac{d}{p}\right)^k$ он останется ненулевым после случайного выбора $x_1, \ldots, x_k$. Значит, с вероятностью $p' \geqslant \left(1 - \frac{d}{p}\right)^k$ после слу-

чайного выбора $x_1, \dots, x_k$ останется ненулевым многочлен от y.

$$\left(1-\frac{d}{p}\right)^k \cdot \left(1-\frac{d}{p}\right) = \left(1-\frac{d}{p}\right)^{k+1}.$$

2. Класс BPP

Теорема 24 (Гача – Сипсера) $\mathrm{BPP} \subset \sum_2^p \bigcap \prod_2^p$. *

Док-во: Достаточно доказать, что $\mathrm{BPP} \subset \sum_2^p$ (поскольку $L \in \mathrm{BPP} \Leftrightarrow \overline{L} \in \mathrm{BPP}$, то $\overline{L} \in \sum_2^p \Leftrightarrow L \in \sum_2^p$).

Пусть $L \in \mathrm{BPP}$. Рассмотрим M — машину с двумя аргументами, такую что

$$x \in L \quad \Rightarrow \quad \Pr_r\{M(x,r)=1\} > 1-2^{-n},$$

$$x \notin L \quad \Rightarrow \quad \Pr_r\{M(x,r)=1\} < 2^{-n}.$$

Обозначим $|r| = m = \mathrm{poly}(n)$.

$$S_x = \{r \subset \{0,1\}^n | M(x,r)=1\}.$$

Либо

$$|S_x| > 2^m\left(1-\frac{1}{2^n}\right),$$

либо

$$|S_x| < 2^{m-n}.$$

В первом случае S_x покрывает почти всю область.

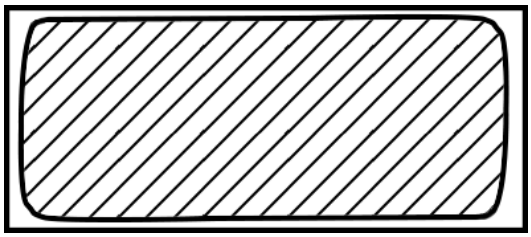

Рис. 12.1

Во втором случае S_x покрывает лишь малую часть области.

Определим

$$S_x + u = \{y+u : y \in S_x\};$$

здесь сложение происходит побитово.

Итоговая формула будет иметь следующий вид:

$$\exists\, u_1, \dots, u_k : \forall\, r \bigvee_{i=1}^{k} (r \in S_x + u_i).$$

Рис. 12.2

Это означает, что любая точка r покрыта каким-то сдвигом.

Поскольку k — полином, и

$$r \in S_x + u_i \quad \Leftrightarrow \quad r + u_i \in S_x \quad \Leftrightarrow \quad M(x, r + u_i) = 1,$$

значит, эта формула лежит в $\sum_2^p$.

Если $x \notin L$, то итоговая формула заведомо неверна при $k < 2^n$ (принцип Дирихле).

Допустим, $x \in L$. Применим вариационный метод. Проведем оценку:

$$\Pr_{u_1,\dots,u_k}\{\bigcup_{i=1}^{k}(r \in S_x + u_i) \neq \{0,1\}^m\} = \Pr_{u_1,\dots,u_k}\{\exists\, r \in \{0,1\}^m : r \notin \bigcup_{i=1}^{k}(S_x + u_i)\} \leqslant$$

$$\leqslant 2^m \Pr_{u_1,\dots,u_k}\{r \notin \bigcup_{i=1}^{k}(S_x + u_i)\} \leqslant 2^m \Pr_{u_1,\dots,u_k}\{r \in \bigcap_{i=1}^{k}\overline{(S_x + u_i)}\} =$$

$$= 2^m \prod_{i=1}^{k} \Pr\{r \in \overline{(S_x + u_i)}\} < 2^m \cdot \left(\frac{1}{2^n}\right)^k = 2^{m-nk}.$$

Для удовлетворения неравенства

$$2^{m-nk} < 1$$

достаточно выбрать $k > \dfrac{m}{n}$.