Лекция 3. Односторонняя перестановка и генератор псевдослучайных чисел

1 XOR-лемма Яо

Теорема 1. *Если существует односторонняя перестановка* $p : D_n \to D_n, D_n \subset \{0,1\}^{k(n)}$*, то существует генератор псевдослучайных чисел.*

Доказательство. Напоминание: схема доказательства теоремы:

- Односторонняя перестановка $f \mapsto$ односторонняя перестановка с трудным битом (декодирование списком кода Адамара и дерандомизации с помощью попарной независимости).
- Генератор $n \to n+1$: $G(x) = g(x)b(x)$ (XOR-лемма Яо).
- Генератор $n \to p(n)$: $g(g(x))b(g(x))b(x), g(g(g(x)))b(g(g(x)))b(g(x))b(x), \ldots$ (hybrid argument).

Сначала сделаем второй шаг.

Определение 1. $b(x)$ — трудный бит для $f(x)$, если он полиномиально вычислим и $\forall p(\cdot)\forall\{P_n\}_{n=1}^{\infty}\exists N : \forall n \geqslant N \to \left|P(P_n(f(x)) = b(x)) - \frac{1}{2}\right| < \frac{1}{p(n)}$.

Лемма. $b(x)$ — *трудный бит для* $f(x) \Rightarrow G(x) = f(x)b(x)$ — *генератор псевдослучайных чисел.*

Доказательство. Если существует отличитель для $G(x)$, то $\exists s(\cdot)\exists\{D_n\}_{n=1}^{\infty}\forall N\exists n > N : |P_x(D_n(f(x)b(x)) = 1) - P_y(D_n(y) = 1)| \geqslant \frac{1}{s(n)}$. Можно считать, что выражение под модулем положительно, так как для тех n, для которых это не так, можно инвертировать вывод D_n.

Рассмотрим варианты для $D(f(x)0) = \alpha, D(f(x)1) = \beta$.

- $\alpha = \beta \Rightarrow$ значение предсказателя случайно.
- $\alpha = 0, \beta = 1 \Rightarrow$ предсказатель возвращает 1.
- $\alpha = 1, \beta = 0 \Rightarrow$ предсказатель возвращает 0.

Обозначим A, B, C, D — события для $00, 01, 10, 11$ соответственно. $A_0, A_1 \subset AB_0, B_1 \subset B \ldots$ разбиения по значениям трудного бита, $a_0, a_1, \ldots$ — их вероятности.

$P(D_n(f(x)b(x)) = 1) = b_1 + c_0 + d_0 + d_1$,
$P(D_n(y) = 1) = \frac{b_0+b_1}{2} + \frac{c_0+c_1}{2} + d_0 + d_1$.
Тогда разность $\Delta = \frac{b_0+b_1}{2} + \frac{c_0+c_1}{2} \geqslant \frac{1}{s(n)}$.
Успех предсказателя: $\frac{a_0+a_1}{2} + b_1 + c_0 + \frac{d_0+d_1}{2} = \frac{1}{2} + \Delta \geqslant \frac{1}{2} + \frac{1}{s(n)}$. □

Почему XOR-лемма? Потому что $P(f(x)) = D(f(x)r) \oplus r \oplus 1$.

2 Построение генератора любой длины

Теперь сделаем генератор $n \to q(n)$. Для начала рассмотрим $G(x) = f(f(x))b(f(x))b(x)$, что должно быть вычислительно неотличимо от xr_1r_2.

$xr_1r_2 \sim f(x)r_1r_2$, так как x и $f(x)$ одинаково распределены (так как f — перестановка). $f(x)r_1r_2 \sim f(x)b(x)r_2$ по определению G. Далее, $xr_2 \sim f(x)b(x) \Rightarrow xr_1r_2 \sim f(f(x))b(f(x))b(x)$.

Для любого константного увеличения можно сделать точно также. Для $n \to q(n)$ делаем так:

$$h_0(x) = xr_1r_2 \dots r_{q(n)}$$

$$\vdots$$

$$h_{q(n)}(x) = f^{q(n)}(x)b(f^{q(n)-1}(x)) \dots b(x)$$

Хотим доказать, что $h_{q(n)} \sim h_0(x)$. Если $P(D_n(h_{q(n)}(x)) = 1) - P(D_n(h_0(x)) = 1) \geqslant \frac{1}{s(n)}$, то $\exists m : P(D_n(h_m(x)) = 1) - P(D_n(h_{m-1}(x)) = 1) \geqslant \frac{1}{s(n)q(n)}$, что невозможно аналогично пункту $n \to n + 2$.

Теорема 2 (Левин-Голдрайх)**.** *Пусть f — односторонняя перестановка, то $g(xy) = f(x)y$ тоже одностороняя перестановка, а $b(xy) = x \odot y = \bigoplus_{i=1}^{n} x_i y_i$ — трудный бит для g.*

Доказательство. Первая часть очевидна, если f — односторонняя пересатновка, то и g тоже перестановка, легко вычисляется и если g можно обратить, то обратить можно и f. Для доказательства второй части воспользуемся кодом Адамара.

Код Адамара: $x \mapsto (x \odot z)_{z \in \{0,1\}^n}$ слово длины n превращает в слово длины 2^n. Его можно воспринимать как значение всех линейных функций на входе x или как значение на всех входах линейной функции, заданной x.

Пусть $\hat{f}(z)$ совпадает с $f(z)$ на доле входов z равной $\frac{3}{4} + \varepsilon$. Тогда можно восстановить $f(z) = \hat{f}(z + r) + \hat{f}(r)$ и с вероятностью $> \frac{1}{2}$ мы восстановим $f(z)$. Повторив много раз, можем узнать $f(e_i) = x_i$.

Для доли повреждения $\frac{1}{2}$ декодировать уже не получится, но можно декодировать списком: имея доступ к $\hat{f}(z)$ как к оракулу, напечатать полиномиальный список в котором с вероятностью $\geqslant \frac{1}{2}$ находится вектор x, определяющий f.

Задача. В шаре с центром в любой точке и радиусом (в смысле расстояния Хемминга) $\frac{1}{2} - \varepsilon$ находится $poly\left(\frac{1}{\varepsilon}\right)$ кодовых слов.

Запишем равенство: $f(z) = \hat{f}(z + r) + f(r)$, которое должо быть выполнено в $\geqslant \frac{1}{2}$ случаев. Непонятно только, откуда взять $f(r)$.

Идея попарной независимости: проведем процедуру выше для некоторого числа попарно независимых случайных r. Возьмем случайные неависимые в совокупности вектора $u_1, \dots, u_l$ и вектора $r_1, \dots, r_{2^l-1}$ построим как $r_a = a_1u_1 + \dots + a_lu_l$. Тогда $r_1, \dots, r_{2^l-1}$ попарно независимы. Алгоритм будет следующий:

```
u_1, ..., u_l := random()
for (f(u_1), ... f(u_l) in {{0,1}^n}^l) {
    for (int a = 1; a < 2^l - 1; ++a) {
        f(r_a) = a_1 f(u_1) + ... + a_l f(u_l)  // linearity
        f(e_i) = f_hat(e_i + r_a) + f(r_a)
    }
    choose f(e_i) as majority for all a
    add f(e_1), ..., f(e_l) in list
}
```

Утверждается, что по неравенству Чебышёва при большом числе повторений с вероятностью больше, чем $\frac{1}{2}$ декодирование произведено верно.

Теперь, если g — это односторонняя функция, $h(xy) = g(x)y, b(xy) = x \odot y = f(y)$ и есть предсказатель b, то можно с его помощью построить $\hat{f}(y)$, совпадающую на доле $\frac{1}{2} + \varepsilon$, что можно декадировать списком $x_1, \ldots, x_m$ и каждый x проверить непосредственно.

Несмотря на то, что $\hat{f}$ экспоненциально длинная, но нам нужно только занчение в полиноме точек, которые мы и запомним (или можно относиться к $\hat{f}$ как к оракулу).

□

□