Сложность вычислений
Лекция 8: Дерандомизация ограниченных по памяти вычислений

Дмитрий Ицыксон

ПОМИ РАН

30 марта 2011

План

1. Экстракторы (напоминание)
2. Лемма о переиспользовании случайных битов
3. Дерандомизация вычислений с логарифмической памятью

Слабые источники и экстракторы

- Пусть D распределение на $\{0,1\}^n$.
- Минимальная энтропия
  $H_\infty(D) = \max\{k \in \mathbb{R} \mid \forall z\ D(z) \le 2^{-k}\}$.
- $0 \le H_\infty(D) \le n$; если $H_\infty(D) = n$, то X — равномерное.
- Распределение D на $\{0,1\}^n$ — это (n,k)-источник, если,
  $H_\infty(D) \ge k \iff \forall z\ D(z) \le 2^{-k}$.
- Определение. Функция $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ называется (k,ϵ)-экстрактором, если для любого (n,k)-источника D, распределение $\mathrm{Ext}(D, U_d) \approx_\epsilon U_m$.
- Теорема. $\forall k, n, \epsilon$ существует (k,ϵ)-экстрактор $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^k$, где $d = \log n + 2\log(1/\epsilon) + O(1)$.
- Лемма. $\forall \epsilon > 0, n \ge k$ можно построить (k,ϵ)-экстрактор $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^n$, где $t = O(n - k + \log \frac{1}{\epsilon})$.

Переиспользование случайных битов

- Пусть мы использовали n случайных битов для вычисления функции $f : \{0,1\}^n \to \{0,1\}^s$.
- Если $s << n$, то казалось бы, что можно переиспользовать эти случайные биты еще раз.
- Поскольку мы знаем $f(x)$, то примерно s случайных битов мы потеряли.
- Лемма. (Лемма о переиспользовании) $f : \{0,1\}^n \to \{0,1\}^s$, $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^n$ — это (k,ϵ)-экстрактор, где $k = n - s - \log \frac{1}{\epsilon}$. Тогда для $X \leftarrow U_n, W \leftarrow U_n, z \leftarrow U_t$ выполняется $f(X) \circ W \approx_\epsilon f(X) \circ \mathrm{Ext}(X,z)$.

Лемма о переиспользовании

Лемма. (Лемма о переиспользовании) $f : \{0,1\}^n \to \{0,1\}^s$, $\mathrm{Ext} : \{0,1\}^n \times \{0,1\}^t \to \{0,1\}^n$ — это $(k, \epsilon/2)$-экстрактор, где $k = n - (s+1) - \log \frac{1}{\epsilon}$. Тогда для $X \leftarrow U_n, W \leftarrow U_n, z \leftarrow U_t$ выполняется $f(X) \circ W \approx_\epsilon f(X) \circ \mathrm{Ext}(X, z)$.

Доказательство.

- X_v — случайная величина, равномерно распределенная на $f^{-1}(v)$, где $v \in \{0,1\}^s$.

$$\delta(f(X) \circ W, f(X) \circ \mathrm{Ext}(X,z)) = \sum_{v,w} |\Pr_{X,W}[f(X) = v \wedge W = w]$$

$$-\Pr_{X,z}[f(X) = v \wedge \mathrm{Ext}(X,z) = w]| = \sum_v \Pr[f(X) = v] \cdot \delta(W, \mathrm{Ext}(X_v, z))$$

- Пусть $V = \{v \mid \Pr[f(X) = v] \geq \epsilon/2^{s+1}\}$. Для $v \in V$, X_v — это (n,k) — источник для $k = n - (s+1) - \log \frac{1}{\epsilon}$.
- $\mathrm{Ext}(X_v, z) \approx_{\epsilon/2} W$ для $v \in V$.
- $\sum_{v \notin V} \Pr[f(X) = v] \leq 2^s \times \frac{\epsilon}{2^{s+1}} = \epsilon/2$

Вычисления с логарифмической памятью

- **L** — класс языков, принимаемых Машинами Тьюринга, использующих $O(\log n)$ памяти.
- $\mathbf{BP_HL}$ — это множество языков A, для которых существует вероятностная машина Тьюринга M:
  1. $\forall x \ \Pr[M(x) = A(x)] \geq \frac{3}{4}$
  2. M использует $O(\log n)$ памяти
  3. $\forall x \ M(x)$ останавливается для любой последовательности случайных битов.
- Легко показать, что $\mathbf{L} \subseteq \mathbf{P}$.
- Покажем, что $\mathbf{BP_HL} \subseteq \mathbf{P}$.

$\mathbf{BP_HL} \subseteq \mathbf{P}$

- Конфигурация МТ (содержимое лент, положение головок, состояние): $O(\log n)$ бит.
- Число конфигураций: $\leq n^d$.
- Матрица перехода $n^d \times n^d$: бистохастическая матрица из $0, \frac{1}{2}, 1$.
- Из конечного состояния никуда уйти нельзя.
- Состояния в пути повторяться не могут.
- Достаточно возвести матрицу перехода в степень n^d и умножить на вектор начального распределения.
- (Упражнение) Докажите, что **BPL** $\subseteq$ **P** (в **BPL** машины останавливаются с вероятностью 1).

Псевдослучайный генератор Нисана

Теорема. Для любого d существует c и полиномиально вычислимая функция $g : \{0,1\}^{c \log^2 n} \to \{0,1\}^{n^d}$, что любая логарифмическая по памяти машина M, граф конфигураций которой имеет $\leq n^d$ вершин, обладает свойством:

$$\left| \Pr_{r \leftarrow U_{n^d}} [M(x,r) = 1] - \Pr_{z \leftarrow U_{c \log^2 n}} [M(x, g(z)) = 1] \right| < \frac{1}{10}$$

- Пусть M имеет $L \leq n^d$ конфигураций. Расстояние от начальной конфигурации до конечной $\leq L$.
- Считаем, что в каждой конфигурации используется 1 случайный бит.
- Разобьем случайные биты на две части: первые $L/2$ и вторые $L/2$.
- После использования первых $L/2$ случайных битов, единственная информация о них — это конфигурация ($O(\log n)$ битов). Можно применить лемму о переиспользовании.

Конструкция генератора

Генератор Нисана

- $r > 0$, $\mathrm{Ext}_k : \{0,1\}^{(k+1)r} \times \{0,1\}^r \to \{0,1\}^{(k+1)r}$ — экстрактор.
- $G_k : \{0,1\}^{(k+1)r} \to \{0,1\}^{2^k}$
- $G_0(z) = z_1$ (первый бит строки z).
- $G_k(a \circ z) = G_{k-1}(a) \circ G_{k-1}(\mathrm{Ext}_{k-1}(a, z))$, $|a| = kr, |z| = r$.

- Пусть s — это строка длины 2^k. Обозначим через $f_{u,2^k}(s)$ вершину, в которую придет машина за 2^k шагов, начав в конфигурации u, используя s в качестве случайных битов.
- $f_{u,2^k}(D)$ — распределение вероятностей на уровне 2^k.
- Основная цель: показать $f_{u,2^k}(U_{2^k}) \approx f_{u,2^k}(G_k(U_{(k+1)r}))$.

Генератор Нисана

Теорема. Пусть $r = O(\log n)$ такое, что для всех $k \leq d \log n$, $\mathrm{Ext}_k : \{0,1\}^{(k+1)r} \times \{0,1\}^r \to \{0,1\}^{(k+1)r}$ — это $((k+1)r - 2d \log n, \epsilon/2)$-экстрактор.

- $G_k : \{0,1\}^{(k+1)r} \to \{0,1\}^{2^k}$
- $G_0(z) = z_1$ (первый бит строки z).
- $G_k(a \circ z) = G_{k-1}(a) \circ G_{k-1}(\mathrm{Ext}_{k-1}(a,z))$, $|a| = kr, |z| = r$.

Тогда $\delta(f_{u,2^k}(U_{2^k}), f_{u,2^k}(G_k(U_{(k+1)r}))) \leq 3^k L\epsilon$.
Как применить эту теорему?

- u — начальная конфигурация
- $2^k = L$
- $3^k L\epsilon < \frac{1}{10} \implies \log 1/\epsilon = O(\log L) = O(\log n)$
- Экстрактор из случайного блуждания по экспандеру.

Доказательство

- Нужно доказать: $\delta(f_{u,2^k}(U_{2^k}), f_{u,2^k}(G_k(U_{(k+1)r}))) \leq 3^k L\epsilon$.
- Пусть ϵ_k — это максимум левой части по всем f и u.
- $\epsilon_0 = 0$.
- Индукцией докажем, что $\epsilon_k \leq 2\epsilon_{k-1} + \epsilon L$.
- Чтобы оценить расстояние между $f_{u,2^k}(D_1)$ и $f_{u,2^k}(D_4)$, мы рассмотрим два дополнительных распределения D_2 и D_3 и воспользуемся неравенством треугольника:
  $\delta(f_{u,2^k}(D_1), f_{u,2^k}(D_4)) \leq \delta(f_{u,2^k}(D_1), f_{u,2^k}(D_2)) +$
  $\delta(f_{u,2^k}(D_2), f_{u,2^k}(D_3)) + \delta(f_{u,2^k}(D_3), f_{u,2^k}(D_4))$.
- $D_1 = U_{2^k}$
- $D_4 = G_k(U_{(k+1)r})$
- $D_2 = U_{2^{k-1}} \circ G_{k-1}(U_{kr})$
- $D_3 = G_{k-1}(U_{kr}) \circ G_{k-1}(U_{kr})$
- $D_1 \overset{\epsilon_{k-1}}{\to} D_2 \overset{\epsilon_{k-1}}{\to} D_3 \overset{L\epsilon}{\to} D_4$

$$\delta(f_{u,2^k}(D_1), f_{u,2^k}(D_2)) \leq \epsilon_{k-1}$$

- $D_1 = U_{2^k}$
- $D_2 = U_{2^{k-1}} \circ G_{k-1}(U_{kr})$
- $p_{uw} = \Pr[f_{u,2^{k-1}}(U_{2^{k-1}}) = w]$
- $q_{uw} = \Pr[f_{u,2^{k-1}}(G_{k-1}(U_{kr})) = w]$
- $\sum_w |p_{uw} - q_{uw}| = \delta(f_{u,2^{k-1}}(U_{k-1}), f_{u,2^{k-1}}(G_{k-1}(U_{kr}))) \leq \epsilon_{k-1}$
- $D_1 = U_{2^k} = U_{2^{k-1}} \circ U_{2^{k-1}}$

$$\begin{aligned}\delta(f_{u,2^k}(D_1), f_{u,2^k}(D_2)) &= \sum_v \left| \sum_w p_{uw} p_{wv} - \sum_w p_{uw} q_{wv} \right| \\ &= \sum_w p_{uw} \sum |p_{wv} - q_{wv}| \leq \epsilon_{k-1}\end{aligned}$$

$$\delta(f_{u,2^k}(D_2), f_{u,2^k}(D_3)) \leq \epsilon_{k-1}$$

- $D_2 = U_{2^{k-1}} \circ G_{k-1}(U_{kr})$
- $D_3 = G_{k-1}(U_{kr}) \circ G_{k-1}(U_{kr})$

АНАЛОГИЧНО!

$$\delta(f_{u,2^k}(D_3), f_{u,2^k}(D_4)) \leq \epsilon L$$

- $D_3 = G_{k-1}(U_{kr}) \circ G_{k-1}(U_{kr})$
- $D_4 = G_k(U_{(k+1)r}) = G_{k-1}(U_{kr}) \circ G_{k-1}(\mathrm{Ext}_{k-1}(U_{kr}, U_r))$
- Пусть $g_u : \{0,1\}^{kr} \to [1, L]$, $g_u(a) = f_{u,2^{k-1}}(G_{k-1}(a))$
- Пусть $X, Y \leftarrow U_{kr}$, $z \leftarrow U_r$.
- Лемма о переиспользовании:
  $g_u(X) \circ Y \approx_\epsilon g_u(X) \circ \mathrm{Ext}_{k-1}(X, z)$.
- $g_u(X) \circ g_w(Y) \approx_\epsilon g_u(X) \circ g_w(\mathrm{Ext}_{k-1}(X, z))$ для всех w на расстоянии 2^{k-1} от u.

$$\delta(f_{u,2^k}(D_3), f_{u,2^k}(D_4)) = \sum_v |\sum_w Pr[g_u(X) = w \wedge g_w(Y) = v] -$$
$$\sum_w Pr[g_u(X) = w \wedge g_w(Ext_{k-1}(X, z)) = v]| < \epsilon L$$

Заключительные замечания

- Нетрудно проверить, что очередной бит G_k можно вычислить, за полиномиальное время, используя $O(\log^2 n)$ памяти.
- Из этого следует, что $\mathbf{BP_HL} \in \mathbf{DSpace}[\log^2 n]$
- Мы сократили число случайных битов до $O(\log^2 n)$, проиграли мультипликативный $\log n$ по памяти, алгоритм все еще полиномиальный по времени.
- (Задача) Докажите, что любой язык из $\mathbf{BP_HL}$ можно решить одновременно за полиномиальное детерминированное время и $O(\log^2)$ память.