Лекция 10. Экстракторы II

Про каждый из описанных объектов есть вероятностное доказательство существововования, которое не приводится

1 Комбинаторная интерпретация

Seeded-экстрактор можно представить как двудольный граф с долями $\{0,1\}^n$ и $\{0,1\}^m$ и рёбрами проведенными естественным образом. Тогда условие на экстактор запишется как

$$\forall S : |S| = 2^k \to \left| \frac{|E(S,T)|}{|S|D} - \frac{|T|}{2^m} \right|$$

.

Multisource-экстрактор удобно рассматривать как таблицу, раскрашенную в один из $\{0,1\}^m$ цветов. Тогда условие на экстрактор будет выглядеть так: для любых достаточно больших наборов столбцов и строк S, T число клеток x, покрашенных в цвета из множества $Q \subset \{0,1\}^m$ удовлетворяет следующему неравенству:

$$\left| \frac{x}{|S||T|} - \frac{|Q|}{2^m} \right| < \varepsilon.$$

2 Некоторые усиления и родственные объекты

Если, например, взять multisource-экстрактор, и испортить в нём распределение битов в первой строке, то общее распределение пострадает не сильно. Поэтому можно рассматривать экстратkторы в сильном смысле.

Определение 1. Multisource-экстрактор называется *экстрактором в сильном смысле*, если условные распределения $MEXt(x,y) \mid y$ и $Mext(x,y) \mid x$ тоже ε-близки к равномерному.

Определение 2. Seeded-экстрактор называется *экстрактором в сильном смысле*, если $(y, Ext(x,y))$ — ε-близко к равномерному.

Определение 3. Двудольный граф, в котором можно пошагово на запрос вершины в левой доле говорить соседа в правой доле (так, чтобы набор рёбер оставался парасочетанием), называется *графом, допускающим online-парасочетание.*

Определение 4. *Дисперсер* это функция $Disp(x,y)$ такая, что для $\forall \xi, H_\infty(\xi) \geqslant k, \eta \sim U_{2^d}, \eta \perp \xi \to Disp\left(\{0,1\}^n \times \{0,1\}^d\right)$ занимает $\geqslant 1-\varepsilon$ от $\{0,1\}^m$.

3 Конструкции экстракторов

Пусть $H : \{0,1\}^n \to \{0,1\}^m$ — семейство хеш-функций, тогда организуем экстрактор следующим образом: $Ext(x,h) = h(x)$.

Лемма (Leftover hash lemma)**.** *Если* $m = k - 2\log\frac{1}{\varepsilon}$, *то полученный объект — сильный* $\left(k, \frac{\varepsilon}{2}\right)$*-экстрактор.*

Доказательство. Нужно доказать, что $(h, h(x)) \sim U_d \times U_m$. Обозначим $D = 2^d, M = 2^m, N = 2^n$, тогда $M = K\varepsilon^2$.

Оценим вероятность коллизии: $P_{x,h,x',h'}\{Ext(x,h) = Ext(x',h') \wedge h = h'\} = P_{x,h,x',h'}\{h = h' \wedge (x = x' \vee (x \neq x' \wedge h(x) = h(x')))\} \leqslant \frac{1}{DK} + \frac{1}{DM} = \frac{1}{DK}(1+\varepsilon^2)$.

Теперь оценим L_2 расстояние от нашего распределения до равномерного: $\|(h, h(\xi)) - U_d \times U_m\|^2 = \sum\limits_{z,t}(P(h - z, h(\xi) = t) - \frac{1}{DM})^2 = P(\text{коллизии}) - \frac{2}{DM}\sum\limits_{z,t} P(h = z, h(\xi) = t) - \frac{1}{DM} \leqslant \frac{1}{DK} + \frac{1}{DM} - \frac{1}{DM} = \frac{\varepsilon^2}{DM}$.

Тогда $|(h, h(\xi)) - U_d \times U_m|_1 \leqslant \varepsilon$, а значит статистическое расстояние не больше $\frac{\varepsilon}{2}$. □

Это довольно плохой экстрактор, однако, лучшие построенные ограничиваются $O(\log^2 n)$ дополнительными чисто-случайными битами.