Определение (это определение чего?) $\mu_G(x) \stackrel{def}{=} P[G(Un\ (...)) = x] = \frac{\#\{s:\ G(s)=x\}}{2^k}$

Определение (статистическое расстояние) $d(\mu,\nu) \stackrel{def}{=} \frac{\sum_x |\mu(x)-\nu(x)|}{2} = \max_{A\subset\{0,1\}^{WTF}} |\mu(A)-\nu(A)|$ (WTF)

Определение Семейство функций $\{G_n \mid n \in \mathbb{N}\}$ называется PRG, если

- оно полиномиально вычислимо;
- $\forall$ полиномиального по n алгоритма B существует пренебрежимо малая последовательность $\{\alpha_n\}$ такая, что

  $$\forall x, y\ |P[B(x,r)=y] - P[B(x,G_n(s))]| \le \alpha_n$$

  где $r = Un_{\ l(n)}, s = Un_{\ k(n)}$. (ДОПИСАТЬ МНОГОЧЛЕНЫ И ОПРЕДЕЛЕНИЯ ФУНКЦИЙ)

Гипотеза $\exists$ PRG.

$\exists$ PRG $\Rightarrow \exists$ OWF (one-way function) $\Rightarrow$ P $\neq$ NP

Будет описана конструкция Блюма-Микеля(1), HILL(3), ?V?(2) - конструкции PRG, исходя из односторонних функций. Помимо этого существует конструкция Нисан-Вандерсон, F - PRG, исходя из трудноразрешимых языков.

Определение (схема из функциональных элементов) понятно что такое. размер схемы - количество элементов в ней.

Определение Последовательность $\{\alpha_n\}$ называется пренебрежимо малой, если $\forall poly(n)\ \exists N:\ \forall n > N\ |\alpha_n| < \frac{1}{poly(n)}$.

Определение $\{f_n\},\ f_n : \{0,1\}^{l(n)} \to \{0,1\}^{m(n)}$ - семейство односторонних функций, если:

- $\{f_n\}$ - полиномиально вычислимо относительно n;
- $\forall\{C_n\}$ - последовательности схем полиномиального размера

  $$P[C_n(f_n(x)) \in f_n^{-1}(f_n(x))] \sim 0$$

- $Un\ (Dom\ f_n)$ - доступно.

Определение $d(\alpha_n,\beta_n) = \sum\limits_{x\in Dom} \frac{|\alpha_n(x)-\beta_n(x)|}{2}$

Определение Распределение μ_n называется полиномиально моделируемым, если существует (вероятностный) алгоритм A, получающий на вход $Un\ p(n)$ и $\forall x \in Dom\ \mu_n\ \ P[A=x] = \mu_n(x)$.

Определение Распределение μ_n называется доступным, если существует полиномиально моделируемое распределение η_n такое, что $d(\mu_n, \eta_n) \sim 0$.

Свойства

1) $\alpha_n \sim \beta_n, \beta_n \sim \gamma_n \Rightarrow \alpha_n \sim \gamma_n$

2) $\alpha_n \sim \beta_n$, и γ_n независима от α_n и β_n. Тогда $\alpha_n\gamma_n \sim \beta_n\gamma_n$ (конкатенация).

   $\triangle$ Пусть это не так. Тогда существует $\{C_n\}$ полиномиального размера, такие, что $|P[C_n(\alpha_n\gamma_n) = 1] - P[C_n(\beta_n\gamma_n) = 1]|$ - не пренебрежимо малая последовательность. Заметим, что $|P[C_n(\alpha_n\gamma_n) = 1] - P[C_n(\beta_n\gamma_n) = 1]| \leq |E_{\gamma_n}(P[C_n(\alpha_n\gamma_n) = 1] - P[C_n(\beta_n\gamma_n) = 1])| \leq E_{\gamma_n}(|P[C_n(\alpha_n\gamma_n) = 1] - P[C_n(\beta_n\gamma_n) = 1]|) \leq |P[C_n(\alpha_n\gamma_{max}) = 1] - P[C_n(\beta_n\gamma_{max}) = 1])|$, то есть $\alpha_n \ncong \beta_n$. Противоречие.$\square$

3) Пусть $\{T_n\}$ - последовательность схем полиномиального размера, и $\alpha_n \sim \beta_n$. Тогда $T_n(\alpha_n) \sim T_n(\beta_n)$

Определение $h_n(x)$ называется трудным битом для односторонней $f_n(x)$, если $h_n(x)$ полиномиально вычислима, и $\forall\{C_n\}$ - схем полиномиального размера $P[C_n(f_n(x)) = h_n(x)] \sim \frac{1}{2}$.

Определение Две последовательности α_n, β_n называются вычислительно и неотличимыми, если $\forall\{C_n\}$ - схем полиномиального размера $P[C_n(\alpha_n) = 1] \sim P[C_n(\beta_n) = 1]$.

Лемма Пусть $\{f_n\}$ - семейство односторонних функций, являющихся перестановками, а $\{h_n\}$ - ее трудный бит. Тогда

$$h_n(x)f_n(x) \sim r_n f_n(x) \sim r_n x$$

где r_n - чистый случайный бит.

$\triangle$ Докажем правую эквивалентность. Поскольку $P[C_n(x) = 1] = P[C_n(f_n(x)) = 1]$, и используем свойство III.

Докажем левую эквивалентность. От противного. Пусть $\exists\{C_n\}$ - схем полиномиального размера таких, что $\exists s(n) = poly(n)$, и $\forall N\ \exists n > N$:

$$|P[C_n(h_n(x)f_n(x)) = 1] - P[C_n(r_n f_n(x)) = 1]| > \frac{1}{s(n)}$$

Построим $\{R_n\}$:

I $R_n(r_n, f_n(x)) = r_n$, если $C_n(0f_n(x)) = C_n(1f_n(x))$;

II $R_n(r_n, f_n(x)) = 0$, если $C_n(0f_n(x)) = 1, C_n(1f_n(x)) = 0$;

III $R_n(r_n, f_n(x)) = 1$, если $C_n(1f_n(x)) = 1, C_n(0f_n(x)) = 0$;

Тогда легко проверить, что

$$|R_n(r_n, x)| = |P[C_n(h_n(x)f_n(x)) = 1] - P[C_n(r_n f_n(x)) = 1]| > \frac{1}{s(n)}$$

Это значит, что $\{R_n\}$ обращает функцию f_n. Противоречие. $\square$

Определение (код Адамара) Пусть $x \in \{0,1\}^n$. Тогда кодом Адамара $H = H_x : \{0,1\}^n \mapsto \{0,1\}$ называется функция: $\forall y \ H_x(y) = x \odot y$. Эта функция является линейной по y.

Лемма Пусть H_1, H_2 - коды Адамара, и $H_1 \neq H_2$. Тогда H_1 отличается от H_2 ровно в половине точек.

$\triangle$ Пусть $H = H_1 + H_2$. Тогда H - линейная непостоянная функция. Пусть $e_1, \ldots, e_n$ - базис в $\{0,1\}^n$. $\exists j \ H(e_j) = 1$. Установим биективное соответствие между нулями и единицами H по следующему правилу: каждому вектору сопоставим вектор, коэффицент при e_j у которого инвертирован относительно исходного вектора, а остальные коэффициенты такие же. Значит, нулей и единиц у H - одинаковое число. Следовательно, H_1 и H_2 различаются ровно в половине значений. $\square$

Лемма (восстановление по испорченному коду Адамара) Пусть имеется испорченный в не более чем ε доле точек код Адамара $\widetilde{H}$ (известно, что есть такое x, что $\widetilde{H}$ отличается от H не более чем в ε доле точек, где $H = H_x$ - правильный код Адамара для x; при этом само x не известно), $0 < \varepsilon < \frac{1}{4}$. Тогда по $\widetilde{H}$ можно однозначно восстановить x.

$\triangle$ Восстановим правильное значение в точке y. Зафиксируем y. $H(y) = H(y + r) - H(r) \overset{2\varepsilon}{=} \widetilde{H}(y + r) - \widetilde{H}(r)$ (то есть последнее равенство верно с точностью до ошибки в 2ε доле всех r). Первое равенство - по линейности, последнее равенство следует из того, что максимальная ошибка при сложнеии или вычитании суммируется. Следовательно, $H(y) = \widetilde{H}(y+r) - \widetilde{H}(r)$ выполняется в $1 - 2\varepsilon > \frac{1}{2}$ случаев, значит, $H(y)$ определяется точно, как наиболее часто встречающееся по всем возможным r значение $\widetilde{H}(y + r) - \widetilde{H}(r)$. Определив значения H в базисе $\{0,1\}^n$, получим координаты x (т.к. скалярное произведение с базисным вектором - это соответствующая координата x). $\square$

Лемма Пусть события $A_1, \ldots, A_N$ попарно независимы, и каждое происходит с вероятностью $\leq \frac{1}{2} - \varepsilon$. Тогда $P[\text{произошло больше половины событий}] \leq \frac{1}{\varepsilon^2 N}$ <TODO>

Теорема (восстановление списка по испорченному коду Адамара) Существует вероятностный алгоритм, который по n, ε имея в качестве внешней процедуры $\widetilde{H}$, такой, что расстояние Хемминга (???) $(\widetilde{H}_x, H_x) \leq \frac{1}{2} - \varepsilon$, за время $poly(n, \frac{1}{\varepsilon})$ с вероятностью $> \frac{1}{2}$ находит список длины $poly(n, \frac{1}{\varepsilon})$, содержащий x. (список экспонентациальной длины, втф???)

Теорема (Левина-Голдрайха) Пусть f - односторонняя функция. Тогда $x \odot y$ - трудный бит для функции $[x, y] \to [f(x), y]$.

Определение (универсальное семейство хэш-функций) Пусть H - семейство функций (не обязательно всех) вида $\{0,1\}^n \to \{0,1\}^s$, и $h = Un \ H$. H называется универсальным, если

I $\forall x \ h(x)$ равномерно распределено в $\{0,1\}^s$

II $\forall x_1 \neq x_2 \quad h(x_1)$ и $h(x_2)$ - независимы, или, иначе говоря, пары вида (x_1, x_2) равномерно распределены в $\{0,1\}^s$(??? иначе говоря а что с неравными иксами)

Как пример можно привести линейные функции.

Далее еще будет важно, чтобы семейство задавалось полиномиальным количеством параметров (?) от m, s.

Определение (энтропия) Пусть α - случайная велечина с n значениями, и вероятности исходов - $p_1, \ldots, p_n$. Энтропией называется:

$$\begin{array}{rrcl} \text{Шеннона:} & H_0 & = & \sum p_i \log_2 \frac{1}{p_i} \\ \text{Ренье:} & H_1 & = & \log_2 \frac{1}{\sum p_i^2} \\ \text{минимальная:} & H_\infty & = & \min_i log_2 \frac{1}{p_i} \end{array}$$

и верно соотношение $H_\infty \leq H_1 \leq H_0$. Вообще говоря, $2^{-H_r} = \sqrt[r]{\sum p_i^{r+1}}$.

(без доказательства) Максимум всех энтропий при числе исходов n достигается при равной вероятности всех исходов и равен $\log_2 n$.

$\triangle$ ($H_\infty \leq H_1$) <TODO>

Лемма (о сглаживании) Пусть H - универсальное семейство хэш-функций с параметрами (m, s), $h = Un\ (H)$, x - случайная велечина в $\{0,1\}^m$, $H_1(x) \geq k$, $r = Un\ (\{0,1\}^s)$ (!!!! почему и там и там s), $L_1(\alpha, \beta) = \sum\limits_y |P[\alpha = y] - P[\beta = y]|$.

Тогда

$$(h(x), h) \sim_{2^{\frac{s-k}{2}}} (r, h)$$

где $\sim$ понимается в смысле L_1 расстояния.

$\triangle$ Пусть $|H| = 2^l$. Одно из неравенств далее следует из того, что $E\xi^2 \geq (E\xi)^2$.

$$L_1 = \sum_{h,a} |2^{-l} P_x[h(x) = a] - 2^{-l-s}| \leq |E_{h,a}| P_x[h(x) = a] 2^s - 1| \leq$$

$$\leq \sqrt{E_{h,a}(P_x[h(x) = a] 2^s - 1)^2} \leq \sqrt{E_{h,a}(2^s \sum_x P(x) \mathbb{I}[h(x) = a] - 1)^2} \leq$$

$$\leq \sqrt{E_{h,a}(2^s \sum_{x_1} P(x_1) \mathbb{I}[h(x_1) = a] - 1)} \sqrt{E_{h,a}(2^s \sum_{x_2} P(x_2) \mathbb{I}[h(x_2) = a] - 1)} =$$

$$= \sqrt{E_{h,a}(\sum_{x_1, x_2} 2^{2s} P(x_1) P(x_2) \mathbb{I}[h(x_1) = h(x_2) = a]) + Q} = (*)$$

Q - остаток, и $Q = E_{h,a}(1 - 2^{s+1} \sum\limits_x P(x) \mathbb{I}[h(x) = a]) = 1 + (-2) = -1$.

$E_h(\mathbb{I}[h(x_1) = h(x_2) = a]) = 2^{-2s}$, если $x_1 \neq x_2$, и 2^{-s} в другом случае. Для того, чтобы посчитать сумму в (*), прибавим и вычтем этот член.

Из условия $H_1(x) \geq k$ вытекает, что $\sum\limits_x P^2(x) \leq 2^{-k}$ (используется в последнем неравенстве).

$$(*) = \sqrt{1 - (\sum_{(x,x)} P^2(x) - \sum_{(x,x)} 2^s P^2(x)) - 1} = \sqrt{\sum_x (2^s - 1) P^2(x)} =$$

$$= \sqrt{(2^s - 1)\sum_x P^2(x)} \leq \sqrt{2^s 2^{-k}} = 2^{\frac{s-k}{2}}$$

□

Лемма Если α, β - независимые случайные велечины с конечным числом значений, то $H_1((\alpha, \beta)) = H_1(\alpha) + H_1(\beta)$ (вообще говоря, верно для любого количества велечин).

△ Пусть $p_1, \ldots, p_n$ - вероятности значений α, $q_1, ..., q_m$ - вероятности значений β. Тогда вероятности значений (α, β) - $p_i q_j; 1 \leq i \leq n, 1 \leq j \leq m$. Очевидно, что

$$2^{-H_1((\alpha,\beta))} = \sum_{i,j}(p_i q_j)^2 = \sum_{i,j} p_i^2 q_j^2 = (\sum_i p_i^2)(\sum_j q_j^2) = 2^{-H_1(\alpha)} 2^{-H_1(\beta)}$$

из чего следует требуемое. □

Лемма Пусть f - односторонняя функция, b - ее сложный бит, $r_1, \ldots r_n$ - чисто случайные биты. Тогда

$$f(x_1) \ldots f(x_n) b(x_1) \ldots b(x_n) \sim f(x_1) \ldots f(x_n) r_1 \ldots r_n$$

△ Будем заменять по одному сложному биту на случайные, и именно, докажем, что

$$f(x_1) \ldots f(x_n) b(x_1) \ldots b(x_{k-1}) b(x_k) r_{k+1} \ldots r_n \sim f(x_1) \ldots f(x_n) b(x_1) \ldots b(x_{k-1}) r_k r_{k+1} \ldots r_n$$

Имеем

$$f(x_k) b(x_k) \sim f(x_k) r_k$$

Видно, что, кроме этого, в больших выражениях все одинаковое, значит, из маленкой эквивалентности следует большая, так как можно дописать в нужные места одинаковые части (по свойству III получившиеся выражения будут эквивалентны). Далее доказательство по индукции.

□

Теорема Пусть существует односторонняя функция $f : \{0,1\}^n \to \{0,1\}^m$, $H_1(f(Un_n)) = n - c$, $0 < c < 1$. Тогда существует PRG.

△ Будем считать, что у f есть трудный бит b. Пусть также имеется случайная хэш - функция из универсального семейства $h : \{0,1\}^{nm} \to \{0,1\}^{n^2 - \sqrt{n} - cn}$. Будем обозначать $x = x_1 \ldots x_n$, $f^n(x) = f^n(x_1 \ldots x_n) = f(x_1) \ldots f(x_n)$, можно тут же заметить, что $H_1(f^n(x)) = (n - c)n$.

Определим генератор G следующим образом:

$$[h, x_1 \ldots x_n]$$

$$\downarrow$$

$$[h, h(f(x_1) \ldots f(x_n)), b(x_1) \ldots b(x_n)]$$

Необходимо доказать, что

$$[h, h(f^n(\bullet)), b(x_1) \ldots b(x_n)] \sim [h, h(f^n(\bullet)), Un_n] \sim [h, Un_{n^2 - \sqrt{n} - cn}, Un_n]$$

Докажем первую эквивалентность. По предыдущей лемме имеем ($r_1 \dots r_n$ - чисто случайные биты)

$$f^n(\bullet)b(x_1)\dots b(x_n) \sim f^n(\bullet)r_1\dots r_n$$

припишем слева h и применим его к $f_n(\bullet)$. Тогда по свойству III получится как раз требуемая эквивалентность.

Докажем вторую эквивалентность. По лемме о сглаживании имеем $[h, h(f^n(\bullet))] \sim [h, Un_{\ n^2-\sqrt{n}-cn}]$ (в качестве случайной велечины берем $f^n(\bullet)$). От дописывания в конец n случайных битов статистическое расстояние не меняется. Следовательно, имеем нужную эквивалентность с точностью до $2^{-\frac{\sqrt{n}}{2}}$, что является пренебрежимо малой функцией.

Длина входа генератора: $|h| + n^2$

Длина выхода генератора: $|h| + n^2 - \sqrt{n} - cn + n$

Видно, что длина входа больше, чем длина выхода.

(дописать про длину x !!!)

Доказано, что G - генератор, (что такое генератор??), у него (

□