$KP(x)$ преф. корр. декомпр.

$KP'(x)$ беспрефиксные

$m(x)$ наибольшая

$$\boxed{\begin{array}{l} \sum_x m(x) \le 1 \\ m(x,i) \xrightarrow[i\to\infty]{} m(x) \\ m(x,i) \in \mathbb{Q} \quad \nearrow \end{array}} \longleftrightarrow \quad m(x) = \Pr[M \text{ выдаст } x]$$

$\forall \mu$ – пер. сл. полумеры $\exists C$: $c \cdot m(x) \geq \mu(x)$ $\forall x$.

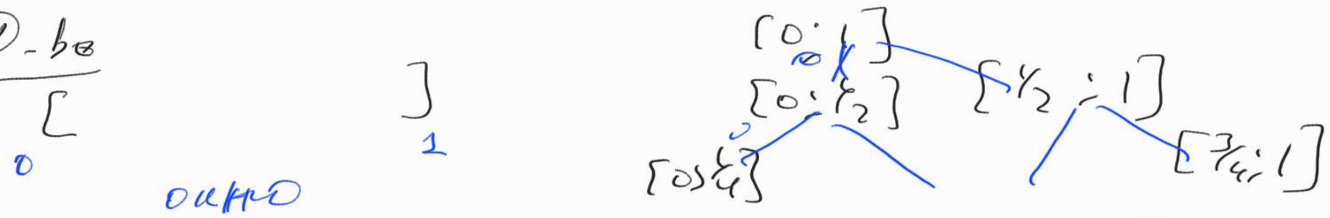
<u>Теорема</u> $-\log m(x) \overset{*}{\le} KP(x) \overset{*}{\le} KP'(x) \overset{(3)}{\le} -\log m(x)$

$O(1)$

<u>Часть (3)</u> $KP'(x) \le -\log m(x) + O(1)$

<u>Лемма</u> (о шторах) $l_1, l_2, \ldots$ — вычисл. посл-ть натур. чисел : $\sum 2^{-l_i} \le 1$.
Тогда $\exists$ вычислимая посл-ть неравных строк $x_1, x_2, \ldots$: $l(x_i) = l_i$

<u>Д-во</u>



<u>Инвариант</u>: своб. пр-во — это объединение непересек. виртуальных штор разной длины.

<u>В самом начале</u>: 1 виртуальная штора длина 1

<u>Шаг.</u> Приносят штору длины w.

а) Есть вирт. шторка длины w. Тогда мы вешаем штору на место этой виртуальной.

б) нет вирт. шторы длины w. Есть шторка длинее $> w$.

$\omega/2 + \omega/4 + \ldots < \omega$

1 самую маленькую метку, кот. $> \omega$. Пусть ее длина ω'.

$\omega' \quad \cancel{\omega} + \omega + 2\omega + 4\omega + \ldots + \omega'/2$

использует новые верт. метки

<u>Сл-ие</u>: l_i — вычислимая посл-ть натур. чисел $\sum 2^{-l_i} \le 1$. Тогда $KP'(i) \le l_i + O(1)$.

$|x_i| = l_i$ $\qquad x_i \longmapsto i$ — вычисл. беспрефиксн. декомпр.

$KP'(x) \le -\log m(x) + O(1)$

$\sum_x m(x) \le 1 \qquad m(x) = \lim_{i\to\infty} m(x,i)$

$m(x,i) \nearrow$

$m'(x,i)$ — округл $m(x,i)$ до ближайшей степени двойки вверх.

$2^{-k+1} < m(x,i) \le 2^{-k}$

$m'(x,i) = 2^{-k}$

$m'(x,i) \le 2\, m(x,i)$

<u>Опр.</u> (x,i) граничной, если $\underline{m'(x,i)} > m'(x,i+1)$ или $i=0$ и $m'(x,0) > 0$.

<u>Утв.</u> $\sum_{(x,i)\text{-гранич.}} m'(x,i) \le 4$

<u>Д-во</u> x — фикс.

$\sum_{i:\,(x,i)\text{-гр}} m'(x,i) \le 2m'(x,i_0) \le 4m(x,i_0) \le 4m(x)$ для некот. i_0

$\sum_{(x,i)\text{-гр}} m'(x,i) \le 4\sum_x m(x) \le 4$, Утв. Мн-во граничных пар (x, i) разрешимо.

Сл-ие: это мн-во перечислимо.

Д-во: перечисление (алгоритмом)

(x_0, i_0), (x_1, i_1), — —. $(\_ .)$

— все граничные пары. $l_n = -\log m'(x_n, i_n) + 2$

l_n : $2^{-l_n} = m'(x_n, i_n)/4$

l_n — вычислимая последовательность

$$\sum_n 2^{-l_n} \leq 1$$

По сл-ию из леммы о шторах

$KP'(n) \leq l_n + O(1)$

$KP'(x_n) = KP'(n) + O(1) \leq l_n + O(1) \leq$

$\leq -\log m'(x_n, i_n) + O(1)$

Для каждого x $\exists n!$ $x = x_n$

Пусть (x, i_n) — это такая гр. пара,

при которой i_n — max.

$x = x_n$ (x_n, i_n)

$KP'(x_n) \leq -\log m'(x_n, i_n) + O(1) \leq$

$\leq -\log m(x) + O(1)$

$KP'(x)$ ($= KP'(x_n)$)

$m'(x_n, i_n) \geq m(x, i)$ $\forall i$

$\Rightarrow m'(x_n, i_n) \geq m(x)$

$KP(x)$ $KP'(x)$ $-\log m(x)$

$KS(x) \le KP(x) + O(1)$

$KP(x) \le 2KS(x) + O(1)$

D — оптим. декомпр.

$D(p) = x$ — оптим. описа. x

$\widehat{|p|}\ p$

$KP(x) \le KS(x) + 2\log KS(x) + O(1)$

$KP(x) \le KS(x) + O(\log n)$ $n = |x|$

$$\sum_x 2^{-KP(x)} \le 1$$

К. Сл-ть и энтропия

$A = \{a_1, \dots, a_k\}$

$x \in A^N$ $p_1, p_2, \dots, p_k$

p_i — частота буквы a_i в x.

$$H(p_1, p_2, \dots, p_k) = -\sum_{i=1}^{k} p_i \log \frac{1}{p_i}$$

<u>Теорема</u> $\frac{KS(x)}{N} \le H(p_1, p_2, \dots, p_k) + \frac{O(\log N)}{N}$

Для большинства слов x выполняется рав-во.

Д-во

$$KS(x \mid N, p_1, \ldots, p_k) \le \log C(N, p_1, \ldots, p_k) + O(1)$$

$C(N, p_1, \ldots, p_k)$ — число слов длины N в алф. A с частотами $p_1, \ldots, p_k$

$$C(N, p_1, \ldots, p_k) = \frac{N!}{(p_1 N)!\,(p_2 N)! \ldots (p_k N)!}$$

$$= \mathrm{poly}(N) \cdot 2^{N H(p_1, \ldots, p_k)}$$

$$\log C(N, p_1, \ldots, p_k) = N H(p_1, \ldots, p_k) + O(\log N)$$

$$KS(x) \le KS(x \mid N, p_1, \ldots, p_k) + \underbrace{KS(N, p_1, \ldots, p_k)}_{\le\, O(k \cdot \log N)} + O(\log n)$$

$|A| = k$ $\quad p_1, p_2, \ldots, p_k \ge 0$ распред.

$\sum p_i = 1$

ξ прин. знач. a_i с вер p_i

ξ^N N независимых копий ξ

Теорема $E\left[KP(\xi^N \mid N)\right] = N \cdot H(\xi) + O(1)$

Д-во $E\left[KP(\xi^N \mid N)\right] \geq N \cdot H(\xi)$

$\exists p_i > 0$ ξ_n принимает знач. A^N.

беспрефиксное

c_i — это описание строки $\alpha_i \mid N$

c_i — обр. префиксный код

$E[c_i] \geq H(\xi^N) = N H(\xi)$

$\exists$ префиксный код со средней длиной символа $\leq H(\xi^N) + 1$.

$E\left[KP(\xi^N / N)\right] \leq NH(\xi) + O(1)$

§ Случайные по Мартин-Лёфу посл-ти

0 1 01 01 01 01

0010100010000000010000101010

Опр. $x_1, x_2, \ldots$ $\in \{0, 1\}$ посл-то случайная, если $\exists c$ $\forall n$ $KS(x_1, x_2, \ldots, x_n) \geq$

$\geq n - c$.

Утв. Это определение плохое, т.к.

$\forall x_1 \ldots x_n \ldots$ $\exists n$!

$K(x_1, \ldots, x_n) \leq n - \log n + O(1)$.

Д-во $\underbrace{x_1 \ldots x_k}_{m} \underbrace{x_{k+1} \ldots x_{k+m}}_{y}$

$n = m + k$

$KS(x_1 x_2 \ldots x_k x_{k+1} \ldots x_{k+m}) \leq$

$\leq KS(y) + c \leq m + O(1) \leq n - k + O(1) \leq$

$\leq n - \log n + O(1)$

$k \geq \log n$

Опр. Посл-ть $x_1, x_2, \ldots$ — наз-ся случайной по Мартин-Лёфу, если $\exists c : \forall n \ KP(x_1, \ldots, x_n) \geq n - c$.

Теорема Мера (Бернулли) не случ. последовательностей равна 0.

Д-во

$$A_c = \bigcup_{\substack{n \\ x_1 \ldots x_n \\ KP(x_1 \ldots x_n) \leq n - c}} \Omega_{x_1 \ldots x_n}$$

$\Omega_{x_1 \ldots x_n}$ — мн-во посл-тей, начин. с $x_1 \ldots x_n$

$\forall c \ A_c$ покрывает не случ. посл-ть

$$\mu(A_c) \leq \sum_{\substack{n \\ x_1 \ldots x_n \\ KP(x_1, \ldots, x_n) \leq n - c}} 2^{-n} \leq \sum_{\substack{n \\ x_1 \ldots x_n \\ KP(\ldots) \leq n - c}} 2^{-KP(\ldots) - c} \leq$$

$-n \leq -KP(\ldots) - c$

$$\leq \sum_{\substack{n \\ x_1 \ldots x_n}} 2^{-KP(x_1 \ldots x_n) - c} \leq 2^{-c} \underbrace{\sum_{x} 2^{-KP(x)}}_{\leq 1} \leq 2^{-c}$$

Теорема Закон больших чисел в форме Харди-Литлвуда.

Для почти всех посл-тей $x_1 \ldots$ ен – —

выполняется

$$\left| \frac{x_1 + \ldots + x_n}{n} - \frac{1}{2} \right| = O\left(\sqrt{\frac{\ln n}{n}}\right)$$

Д-во $x_1 \ldots x_n$ $\quad p_n \cdot n \quad 1$

$(1-p_n) \cdot n \quad 0$

$$KP(x_1 \ldots x_n) \le KS(x_1, \ldots, x_n) + O(\log n) \le$$

$$\le H(p_n) \cdot n + O(\log n)$$

$$H(1/2 + x) = 1 - c \cdot x^2 + o(x^2)$$

$$\le 1 - c' x^2$$

$$p_n = \frac{1}{2} + \delta_n$$

(Достаточно д-ть, что нер-во выполн. для случайных посл-тей

$\exists c \quad \forall n$

$$\underline{n - c} \le \underline{KP(x_1, \ldots x_n)} \le \underline{n - c' \delta_n^2 \cdot n + O(\log n)}$$

$$\underline{\delta_n^2 \le O\left(\frac{\log n}{n}\right)} \qquad \delta_n \le O\sqrt{\frac{\log}{n}}$$