Kolmogorovská zložitosť

08.01.2014

párovacia funkcia

$$\langle x, y\rangle = y + (x + y + 1)(x + y)/2$$
$$\langle x, y, z\rangle = \langle x, \langle y, z\rangle\rangle$$

$\mathcal{B}^*$ $\{\epsilon, 0, 1, 00, 01, 10, 11, \ldots\}$

reťazce–prirodzené čísla, $\mathcal{B}^* \leftrightarrow \mathcal{N}$

$\{(\epsilon, 0), (0, 1,), (1, 2), (00, 3), (01, 4), \ldots\}$

//nerozlišujeme medzi reťazcom 01 a číslom 4

//niekedy medzi reťazcom 100 a číslom 4

dĺžka reťazca w, kardinalita množiny A: $l(w), d(A)$

$D : \{0, 1\}^* \to \mathcal{N}$ $D(y) = x$, y je kód, x je vzor

prefixný kód $E_i(x) = \begin{cases} 1^x 0, & \text{i=0} \\ E_{i-1}(l(x))x, & \text{i>0} \end{cases}$

$\overline{x} = E_1(x) = E_0(l(x))x = 1^{l(x)}0x \quad \rightsquigarrow \quad l(E_1(x)) = 2l(x)+1$

$E_1(100000001011) = 1^{12}0100000001011,\ l(x) = 25$

$E_1(x) = 11111000010 \quad \rightsquigarrow x = 00010$

$E_2(x) = \overline{l(x)}x = 1^{l(l(x))}0l(x)x \quad \rightsquigarrow l(E_2(x)) = l(x) + 2l(l(x)) + 1$

$E_2(100000001011) = E_1(l(100000001011))100000001011 =$

$11110110010000001011,\ l(x) = 21$

$E_2(x) = 11101111000001 \quad \rightsquigarrow$ x=1000001

$\overline{x}y = 1110\mathbf{11}011 \quad \Rightarrow \quad x = 110,\ y = 11$

$\overline{x}\overline{y}z = 1110\mathbf{11}0111\mathbf{0}0\mathbf{1}011 \quad \Rightarrow \quad x = 110,\ y = 010,\ z = 11$

!!!viem, aké kódovanie používam!!!

popis objektu // má zmysel, ak z neho objekt vieme zrekonštruovať

- enumerácia objektu x z množiny S prirodzeným číslom $n(x)$
- $p \in \mathbb{N}$, potom $l(p)$ označuje dĺžku binárneho zápisu p

$$\mathbf{C_f(x) = \min_p\{l(p) \mid f(p) = n(x)\}}$$

zložitosť objektu x vzhľadom k metóde f

$\mathbf{p}$ $\sim$ program $\mathbf{f}$ $\sim$ počítač

- **f (aditívne) minorizuje g** ak $\exists c\ \forall x : \mathbf{C_f(x) \leq C_g(x) + c}$

  ak uvažujeme r metód $f_1, \ldots, f_r$, na identifikáciu metódy stačí povedať $\log r$ bitov

- metódy sú **ekvivalentné** ak sa navzájom aditívne minorizujú

- C je trieda čiastočných funkcií na Z^+. Funkcia **f je univerzálna (aditívne optimálna) pre C** ak

$$\mathbf{f \in C} \quad \& \quad \left[\forall \mathbf{g \in C}\ \exists \mathbf{c_{f,g}} :\ \mathbf{C_f(x) \leq C_g(x) + c_{f,g}}\right]$$

- univerzálne funkcie sú "ekvivalentné:"

$$|C_f(x) - C_g(x)| \leq \max\{c_{f,g}\ c_{g,f}\}$$

Example

Trieda $\mathcal{F}$ všetkých čiastočných funkcií/popisných metód nemá minimálny prvok

Sporom - nech f je univerzálna pre $\mathcal{F}$

- uvažujme postupnosť reťazcov $\mathcal{S} = x_1, x_2, \ldots$: $C_f(x_i) \geq i$
- definujme funkciu $\mathbf{g}$: $\mathbf{g(i) = x_i}$
- Kolmogorovská zložitosť reťazcov z $\mathcal{S}$

$\mathbf{C_g(x_i)} \leq 2 \log i + O(1)$

$2 \log i + O(1) \lll i \leq C_f(x_i)$

$\rightsquigarrow$ f nie je univerzálna $\diamond$

V triede čiastočne rekurzívnych/computable funkcií minimálny prvok existuje $\hookrightarrow$

metódy — $\{\Phi \mid \Phi$ je čiastočne rekurzívna funkcia$\}$

Theorem (invariance)

Existuje univerzálna čiastočne rekurzívna funkcia.

- usporiadanie TS $\rightsquigarrow$ usporiadanie PRF
- Φ_0, funkcia, kt. počíta UTS;
  vstup: $\mathbf{1^{l(n)}\ 0\ n\ p}$ $\quad$ $\mathbf{1^{l(n)}0n}$ — kód TS T_n $\quad$ $\mathbf{p}$ — program
- $\Phi_0(\langle n, p\rangle) = T_n(p)$
  ak $UTS = T_0$, potom $U(0p) = U(p)$

$$C_{\Phi_0}(x) \leq C_{\Phi_n}(x) + c_{\Phi_n}, \quad c_{\Phi_n} = 2l(n) + 1$$

□

$x, y, p \in \mathbb{N}$: PRF Φ spolu s p, y je popis $x \Leftrightarrow \Phi{<}\, y, p \,{>}= x$

podmienená Kolmogorovská zložitosť

$$C_\Phi(x|y) = \begin{cases} \min\{l(p) : \ \Phi(\langle y, p\rangle) = x\} & \text{ak } p \text{ existuje} \\ \infty, & \text{ak } p \text{ neexistuje} \end{cases}$$

Theorem (conditional invariance)

*Existuje univerzálna PRF Φ_0 pre triedu PFR, ktoré počítajú **x** z **y**, pričom*

$$C_{\Phi_0}(x|y) \leq C_\Phi(x|y) + c_\Phi$$

- $\Phi_0(\langle y, \langle n, p\rangle\rangle) = \Phi_n(\langle y, p\rangle)$
- $C_{\Phi_0}(x|y) \leq C_{\Phi_n}(x|y) + c_{\Phi_n}, \quad c_{\Phi_n} = 2l(n) + 1$ □

na voľbe univerzálnej funkcie nezáleží

- $\forall$ aditívne optimálne funkcie Ψ, Ψ' $\quad \exists c_{\Psi,\Psi'}$

$$\forall x, y \quad |C_\Psi(x|y) - C_{\Psi'}(x|y)| \leq c_{\Psi,\Psi'}$$

$$\boxed{C(x|y) = C_{\Phi_0}(x|y)} \qquad \boxed{C(x) = C_{\Phi_0}(x|\epsilon)}$$

hľadanie správneho modelu

- $\mathbf{T}$ (samoodeľujúco) popisuje stroj; model, resp. regularita
  $\mathbf{p}$ popisuje program; neregularita

$$\boxed{\mathbf{C(x) = \min\{l(T) + l(p) : T(p) = x\} + O(1)}}$$

- $\min_{\mathbf{T}}\{\mathbf{l(T) + C(x|T)} \mid \mathbf{T} \in \{\mathbf{T_0}, \mathbf{T_1}, \ldots\}\}$

Theorem

$\exists c, \forall x, y \quad C(x) \leq l(x) + c, \qquad C(x|y) \leq C(x) + c$

- $C_T(x|y) = C(x)$ keď T so vstupom $\langle z, y \rangle$ vypočíta x práve vtedy, keď UTS vypočíta x pri vstupe $\langle z, \epsilon \rangle$
- $C(x|y) \leq C_T(x|y) + c = C(x) + c$ □

Example

- $C(xx) = C(x) + O(1)$
- ak Φ je totálna injektívna rekurzívna, potom $\exists c: \quad |C(\Phi(x)) - C(x)| \leq c$
- $C(x + C(x)) \leq ?$

Example

- $C(x,y) \leq C(x) + C(y) + 2\log(\min\{C(x), C(y)\}) + 1$
- $C(x,y|C(x)) \leq C(x) + C(y) + O(1)$

x^* je "prvý najkratší" program, kt.generuje x. Vieme vypočítať x^* , keď poznáme $x, C(x)$?

- ÁNO — systematicky simulujeme programy dĺžky $C(x)$; x^* je "prvý", kt. zastane s x
- výpočet $C(x)$ pri znalosti x sme zredukovali na výpočet x^* pri znalosti $x, C(x)$

$$C(x^*|x) = C(C(x)|x) \leq C(C(x)) \leq \log l(x) \qquad \boxed{\mathbf{C(x) = l(x^*)}}$$

Theorem

Nech $A \subseteq \mathbb{N} \times \mathbb{N}$ *je RE,* $y \in \mathbb{N}$. *Nech* $Y = \{x : \ (x, y) \in A\}$ *konečná. Potom* $\exists c_A \quad \forall x \in Y \ C(x|y) \leq l(d(Y)) + c_A$

- postupne vymenovávame $(x_1, y_1), \ldots$ a počítame tie, keď $y_j = y$ $(x_{i_1}, y), (x_{i_2}, y), \ldots$
- $+$ poradové číslo i_j stačí na určenie $x = x_{i_j}$ □

$\hookrightarrow$ Ak $A \subseteq \mathbb{N}, d(A^{\leq n}) \leq p(n)$, je RE, potom

- $\forall x, l(x) \leq n \quad C(x|n) \leq l(p(n)) + O(1)$
- $C(x) \leq C(x|n) + 2l(n) + O(1) \leq l(p(n)) + 2l(n) + O(1)$
- $C(x) = O(\log n)$

Definition

Nech $\Phi_1, \Phi_2, \ldots$ a $\Psi_1, \Psi_2, \ldots$ sú enumerácie PRF, pričom $\Psi_i = \Phi_{f(i)}, \Phi_i = \Psi_{g(i)}$. Ak f, g sú čiastočne rekurzívne, potom Φ, Ψ sú **rekurzívne izomorfné**.

$\forall$ rekurzívne izomorfné $\Psi, \Phi \; \exists \; c_{\Phi,\Psi} : \forall x \; |C_\Phi(x) - C_\Psi(x)| < c_{\Phi,\Psi}$

ALE

Existujú enumerácie Ψ, Φ také, že rozdiel $|C_\Phi(x) - C_\Psi(x)|$ je neohraničený.

$\forall$ rekurzívne izomorfné Ψ, Φ $\exists$ $c_{\Phi,\Psi} : \forall x \; |C_\Phi(x) - C_\Psi(x)| < c_{\Phi,\Psi}$

- f, g čiastočne rekurzívne $\Rightarrow \exists n(f), m(g)$ tak, že $\Phi_{n(f)}, \Psi_{m(g)}$ počíta f, resp. g.
- $C_{\Phi_0}(x) \leq C_{\Phi(i)}(x) + c_i \leq C_{\Psi_{g(i)}}(x) + c_i + c_{m(g)}$
  $C_{\Psi_0}(x) \leq C_{\Psi(i)}(x) + c_i \leq C_{\Phi_{f(i)}}(x) + c_i + c_{n(f)}$

$$C_\Phi(x) = l(p), \Phi_0(p) = x \qquad c_f$$

$$C_\Psi(x) = l(q), \Psi_0(q) = x \qquad c_g$$

$$C_\Psi(x) \leq l(q) + 2l(c_g) + 1 \qquad C_\Phi(x) \leq l(p) + 2l(c_f) + 1$$

Pre $c = \max\{2l(c_g), 2l(c_f)\} + 1$ $\qquad |C_\Phi(x) - C_\Psi(x)| \leq c$

$\diamond$

Existujú enumerácie Ψ, Φ také, že rozdiel $|C_\Phi(x) - C_\Psi(x)|$ je neohraničený.

Definujeme Ψ k Φ:

$$\begin{aligned} \Psi_{2i}(1) &= y_i & C_\Phi(y_i) \geq i^2 \\ \Psi_{2i}(x) &= \Phi_i(x) & \text{pre } x > 1 \\ \Psi_{2i+1}(x) &= \Phi_i(x) \end{aligned}$$

$$C_\Psi(y_i) \leq C_{\psi_{2i}}(y_i) + c_{\Psi_{2i}} = 1 + c_{\Psi_{2i}}, \; c_{\Psi_{2i}} \leq 2\log 2i + O(1)$$

$$|C_\Phi(y_i) - C_\Psi(y_i)| \geq i^2 - c_{\Psi_{2i}} - O(1) \rightsquigarrow \infty$$

$\diamond$

Definition

Reťazec x je **c-nestlačiteľný** ak $\mathbf{C(x) \geq l(x) - c}$

- počet reťazcov dĺžky n je 2^n
- počet programov dĺžky menšej ako n je $\sum_{i=0}^{n-1} 2^i = 2^n - 1$

  $\hookrightarrow$ $\forall n \; \exists$ aspoň jeden nestlačiteľný reťazec
- # programov dĺžky menšej ako $n - c$ je $\sum_{i=0}^{n-c-1} 2^i = 2^{n-c} - 1$

  $\frac{2^n - (2^{n-c} - 1)}{2^n}$ aspoň (1-2^{-c})-tina reťazcov dĺžky n je c-nestlačiteľná

Theorem (incompressibility)

Nech $c \in \mathbb{N}$, y je fixované a A je konečná množina kardinality aspoň m. Potom aspoň $m(1-2^{-c})+1$ prvkov $x \in A$ má $C(x|y) \geq \log m - c$.

- programov dĺžky menšej ako $\log m - c$ je $\sum_{i=0}^{\log m - c - 1} 2^i = 2^{\log m - c} - 1 = \frac{m}{2^c} - 1$
- v A existuje aspoň $m - (\frac{m}{2^c} - 1) > m(1 - \frac{1}{2^c})$ prvkov s $C(x) \geq \log m - c$

□

Pritom $x \in A$ stačí identifikovať indexom a popisom A

$$C(x) \leq c_A + l(d(A)) \geq c_A + l(m)$$

Sú podreťazce nestlačiteľných reťazcov nestlačiteľné?

c-nestlačiteľné $x = uvw$, $|x| = n$, môžme popísať programom q

program p_v pre generovanie v
reťazec uw
separácia uw na u, w - nech $|u| \geq |w|$

$$\Big\} \; \mathbf{q = \overline{l(p_v)} p_v \overline{l(u)} uw}$$

$$l(q) = C(v) + 2l(C(v)) + 2l(u) + n - l(v) + 2$$

$$\mathbf{n - c \leq C(x) \leq l(q) + O(1)}$$

$$n - c \leq C(x) \leq C(v) + n - l(v) + 4 \log n + O(1)$$

$$\boxed{\mathbf{C(v) \geq l(v) - O(\log n)}}$$

Môžme predpokladať, že $C(v) \geq l(v) - O(1)$? *NIE*

Example

Nech $C(x) = l(p)$. Potom p je c-nestlačiteľné.

Sporom

- nech $\exists$ program $q : U(q) = p$ & $\mathbf{l(q) < l(p) - c}$.
- vezmime $V = T_i : V(q) = U(U(q))$
- potom $\mathbf{U(1^i 0 q) = x}$

$$\mathbf{C(x) \leq l(q) + i + 1 < l(p) - c + i + 1}$$

pre $c > i + 1$ SPOR ◇

Example

O vzťahu $C(x, y)$ k $C(x), C(y)$ — C nie je subaditívna.

horný odhad $C(x, y) \leq C(x) + C(y) + 2\min\{\log(C(x)), \log(C(y))\} + 1$

dolný odhad: Uvažujme $A = \{(x, y) \mid |xy| = n\}$.

- $d(A) = (n + 1)2^n$
- existuje $(x, y) : C(x, y) > \log d(A) - c = n + \log n - c$
- pritom $C(x) + C(y) \leq l(x) + l(y) + d = n + d$

$C(x, y) > n + \log n - c \geq C(x) + C(y) - d + \log n - c = C(x) + C(y) + \log n - e$

Example

C nie je monotónna na prefixoch.

$\exists\ n > m\ x = 1^n,\ y = 1^m$, y vlastný prefix x ale $C(y) > C(x)$
$n = 2^k :\rightsquigarrow C(1^n) \leq \log\log n + O(1)$
m: $A = \{1, 11, \ldots, 1^n\}$

Incompressibility Thm: aspoň $n/2 + 1$ prvkov má $C(z) \geq \log n - 1$

Zvolíme m: $n/2 \leq m < n$

$C(1^m) \geq \log n - 1$ ◇

KZ pri znalosti dĺžky – C(x|l(x))

$\rightsquigarrow C(x|l(x)) \leq C(x) + c$

n-string je reťazec $n0^{n-l(n)}$, pričom $|n0^{n-l(n)}| = n$

- Ak x je n-string, tak $C(x|n) \leq c$
- **ani $C(x|l(x))$ nie je na prefixoch monotónna**
  n nestlačiteľné: $C(n) \geq l(n)$; $x = n0^{n-l(n)}$ $C(x|l(x)) \leq c$

$$C(n|l(n)) \geq C(n) - C(l(n)) \geq \log n - 2\log\log n + O(1)$$

Čo ak vieme, že $x \in A$? // $C(x|A) \leq l(d(A)) + c_A$

Definition

Randomness deficiency x vzhľadom k množine A je

$$\delta(x|A) = l(d(A)) - C(x|A)$$

- $\delta(x|A) \geq -c_A$
- Čo hovorí veľké $\delta(x|A)$?
- Koľko je prvkov s $\delta(x|A) \geq k$?

$$\delta(x|A) = l(d(A)) - C(x|A) \geq k \quad \rightsquigarrow \quad C(x|A) \leq l(d(A)) - k$$

$$|x : C(x|A) \leq l(d(A)) - k| < 2^{l(d(A))-k+1} = \frac{2^{l(d(A))}}{2^{k-1}}$$

$$d(\{x : \delta(x|A) \geq k\}) \leq d(A)/2^{k-1}$$

randomness deficiency a náhodnosť

Nech: $\Phi = \Phi_r$, //Φ partial recursive, computable

$R = \{(x, y) : \Phi(i) = \langle x, y \rangle, i \geq 1\}$ //R recursively enumerable

Nech $\mathbf{A} = \{\mathbf{x} : (\mathbf{x}, \mathbf{y}) \in \mathbf{R}\}$ je konečná

$\hookrightarrow C(x|y) \leq \log d(A) + \log r + 2 \log \log r + O(1)$

$\hookrightarrow \delta(x|y) = \log d(A) - C(x|y)$

x je náhodný v A ak $\delta(x|y) = O(1)$

Nech $A = \{x : l(x) = n\}$, $R = \{(x, n) : l(x) = n\}$

$\hookrightarrow \delta(x|n) = \underbrace{n}_{\log d(A)} - C(x|n) + O(1)$

x je náhodný **iff** $\delta(x|n) = O(1)$

Theorem

Pre Kolmogorovskú zložitosť platí

1. C *je neohraničená*
2. $m(x) = \min\{C(y) : y \geq x\}$ *je neohraničená*
3. $\forall$ *PRF* Φ *monotónne rastúcu donekonečna (od nejakého* x_0*) platí* $m(x) < \Phi(x)$ *až na konečne veľa* x

!!! 2,3 pre $C(x|l(x))$ neplatia

// nekonečne veľa razy skočí na konštantu (n-stringy)

1. vyplýva z 2.
2. $\forall i \; \exists$ min. x_i také, že $\forall x > x_i : \; C(x) \geq i$

$$\rightsquigarrow \quad m(x) = i, \; x_i < x \leq x_{i+1}$$

3. $\forall$ PRF Φ monotónne rastúcu donekonečna (od nejakého x_0) platí $m(x) < \Phi(x)$ až na konečne veľa x

Nech $\Phi{=}\Phi_r$ monotónne neklesajúca,$\Phi(x) \leq m(x)$ pre nekonečne veľa x;

- $D(\Phi) = A = \{x : \ \Phi(x) < \infty\}$ nekonečná, RE, **preto** $\exists$ nekonečná $B, B \subseteq A$, s rekurzívnou χ_B
- $$\Psi(x) = \begin{cases} \Phi(x) & x \in B \\ \Phi(y) & y = \max\{z : z \in B, z < x\} \text{ inak} \end{cases}$$

  $$\Psi(x) \leq \Phi(x) \leq m(x)$$
- $M(a) = \max\{x : C(x) \leq a\}$ $M(a)+1 = \min\{x : m(x) > a\}$
  $\max\{x : \Psi(x) \leq a+1\} \geq \min\{x : m(x) > a\} = M(a)+1 > M(a)$
- $F(a) = \max\{x : \Psi(x) \leq a+1\}$ je totálna rekurzívna

  $\hookrightarrow$ $F(a) > M(a)$ pre nekonečne veľa a $\quad\hookrightarrow$ $C(F(a)) > a$

ALE $\quad \mathbf{C(F(a))} \leq C_F(F(a)) + O(1) \leq l(a) + c_F + O(1) = \mathbf{l(a) + O(1)}$

$\mathbf{a < C(F(a)) < l(a) + c}$ $\square$

intermezzo o rekurzívnych fciách $D(\Phi) = A = \{x : \Phi(x) < \infty\}$ nekonečná, RE,

preto má nekonečnú podmnožiny $B, B \subseteq A$, s rekurzívnou χ_B:

- f RE s oborom hodnôt A, g rekurzívna:

$$g(0) = f(0)$$
$$g(x+1) = f(y),\ y \text{ najmenšie také, že } f(y) > g(x)$$

B je obor hodnôt g

A nekonečná, preto B nekonečná
g vymenováva B v rastúcom poradí
} g **je rekurzívna**

- nekonečná A je rekurzívna $\iff$ ak je enumerovateľná v rastúcom poradí

$\Rightarrow$ postupne počítame charakteristickú fciu

$\Leftarrow$ $y \overset{?}{\in} A$ postupne počítame, až kým $\chi(x) \in B$ pre $x \geq y$

A je RE A je obor hodnôt totálnej rekurzívnej fcie
A je R χ_A je rekurzívna ◇

Theorem (C nie je rekurzívna...)

Funkcia $C(x)$ ***nie je rekurzívna.*** *Navyše, neexistuje čiastočne rekurzívna fcia Φ definovaná na nekonečnej množine, ktorá je na celom svojom definičnom obore s $C(x)$ totožná.*

SPORom

- Φ PRF definovaná na nekonečnej množine, **A** nekonečná rekurzívna podmnožina $D(\Phi)$
- $\Psi(m) = \min\{x : C(x) \geq m, x \in A\}$ je totálna rekurzívna

$$//\Phi(x) = C(x) \text{ na } A$$

- $C(\Psi(m)) \geq m$
- $C(\Psi(m)) \leq C_\Psi(\Psi(m)) + c_\Psi, \quad C_\Psi(\Psi(m)) \leq l(m)$

$$//\Psi(m) = x : C(x) \geq m$$

$$m \leq C(\Psi(m)) \leq C_\Psi(\Psi(m)) + c_\Psi \leq l(m) + c_\Psi$$

□

Theorem (C sa dá aproximovať..)

Existuje totálna rekurzívna fcia $\Phi(t,x)$ monotónne klesajúca v t taká, že $\lim_{t\to\infty}\Phi(t,x) = C(x)$

- $C(x) \leq l(x) + c$
- každý program p, $l(p) \leq l(x)+c$ necháme bežať t krokov a definujme

$$\Phi(t,x) = \begin{cases} \min\{l(p) : \ p \text{ zastal a vygeneroval } x\}, & \text{ak existuje} \\ l(x)+c & \text{inak} \end{cases}$$

□

Pre dané x, t NEVIEME rozhodnúť, či $\Phi(t,x) = C(x)$

Def. (limita $\Phi(t, x)$)

Nech $g = g_1, g_2, \ldots$ je postupnosť funkcií. Funkcia f je limitou postupnosti g ak $\mathbf{f(x) = \lim_{t\to\infty} g_t(x)}$

Limita je **rekurzívne uniformná**, ak $\exists$ totálna rekurzívna fcia $t(\epsilon)$

$$|\mathbf{f(x) - g_{t(\epsilon)}(x)}| \leq \epsilon \ \forall \mathbf{x}$$

$\Psi = \Psi_1, \Psi_2, \ldots \quad \Psi_t(x) = \Phi(t, x)$

- $C(x)$ **JE limitou** Ψ
- $C(x)$ **NIE JE rekurzívnou limitou** Ψ

  lebo z problému zastavenia

  $\forall \epsilon, t$ existuje nekonečne veľa x : $|C(x) - \Psi_t(x)| > \epsilon$

vlastnosti C

C je "spojitá" $\exists c : |C(x) - C(x \pm h)| \leq 2l(h) + c$

$\hookrightarrow$ ak máme program pre x, stačí povedať $\pm h$

C je "logaritmická"

$\hookrightarrow C(x) \leq l(x) + c$

$\hookrightarrow \forall k$ je počet x dĺžky n s $C(x) < \log x - k$ nanajvýš 2^{n-k}

"fluktuácia" C

$\forall x \; \exists x_1, x_2 : \quad |x_i - x| \leq \sqrt{x}$, pričom

1. $C(x_1) \geq l(x)/2 - c$
2. $C(x_2) \leq l(x)/2 + c$

2. nestlačiteľné $x = x''x'$; $l(x'') = l(x') = n/2 \quad \rightsquigarrow \quad x_2 = x''0^{n/2}$

$$C(x_2) \leq l(x)/2 + c \sim C(x)/2$$

1. stlačiteľné x; $C(x) = o(l(x)) \rightsquigarrow$ spodné bity nahradí nestlačiteľný
   $y = y_1 \ldots y_{n/2}$
   $x_1 = x''y \quad C(y) \geq l(y) + c = l(x)/2 + c$

Čo vieme o dlhých zložitých behoch...

1. $\forall c \exists d$: neexistuje d po sebe idúcich c-nestlačiteľných čísel

2. $\forall d \exists c$: existuje d po sebe idúcich c-nestlačiteľných čísel

Lemma

$\forall c \exists d$: *neexistuje* d *po sebe idúcich* c*-nestlačiteľných čísel*

Majme x $c-$ nestlačiteľné $C(x) > l(x) - c$
uvažujme $x, x+1, \ldots, x+\Delta$, $x + \Delta = i\underbrace{0 \ldots 0}_{j}$:

$$\left.\begin{array}{l} C(x+\Delta) \leq l(i) + l(j) + c \\ l(i) + l(j) + const < l(x) - c \end{array}\right\} \text{ stačí } l(j) + const < c$$

Lemma

$\forall d \exists c$: *existuje d po sebe idúcich c-nestlačiteľných čísel*

zoberme k-nestlačiteľné x $\quad C(x) > l(x) - k$
zo spojitosti: $\forall |x - y| \leq k : \ |C(x) - C(y)| \leq 2l(k) + \delta$

$$C(x) - 2l(k) - \delta \leq C(y) \quad \vee \ C(y) - 2l(k) - \delta \leq C(x)$$

$$l(x) - \underbrace{(k + 2l(k) + \delta)}_{c} < C(y)$$

□

keď náhodne zvolíme objekt očakávame, že má "typické" vlastnosti

S - priestor
P - pravdepodobnostné rozdelenie/distribúcia

test - predpis, kt. na hladine významnosti ε hovorí, pre kt. $x \in S$ hypotézu "x je typický" zamietame

kritické regióny

$\varepsilon = 2^{-m},\ m = 1, 2, \ldots,\ V \subseteq N \times S$

$V_m = \{x :\ (m, x) \in V\}$

$V_m \supseteq V_{m+1}$

$\sum_{\mathbf{x}} \{\mathbf{P}(\mathbf{x}|\mathbf{l}(\mathbf{x}) = \mathbf{n}) : \mathbf{x} \in \mathbf{V_m}\} \leq \varepsilon$

//ak $x \in V_m$, tak x neprešlo testom na hladine významnosti ε

komplement V_m je $1 - \varepsilon$ interval spoľahlivosti

Example

$x = 0.x_1 \dots x_n$ s nulami na začiatku nie je náhodný:

$$x = \sum x_i 2^{-i}$$

TEST: $V_0 = \langle 0, 1 \rangle)$
$V_1 = \langle 0, 1/2)$
$V_2 = \langle 0, 1/4)$
$\dots$

$\mathbf{x = 0 \dots 0 x_{m+1} \dots x_n} \in \langle 0, 2^{-m})$
neprejde testom na hladine významnosti $\varepsilon = 2^{-m}$ $\diamond$

Example

$x = x_1 \dots x_n$ - v náhodnom reťazci je počet 0 a 1 "podobný"

$\mathbf{f_n} = \sum_{i=1}^{n} x_i$

$\mathbf{g(n,m) = g}$, pričom g je najmenšie také, aby # binárnych reťazcov, pre ktoré to platí, bol nanajvýš 2^{n-m}

TEST : $\mathbf{V_m = \{x \in \{0,1\}^n : \ |2f_n - n| > g(n,m)\}}$

Definition (Efektívny test náhodnosti?)

Nech P je rekurzívne rozdelenie pravdepodobnosti na priestore $\mathbb{N}$. Totálna funkcia $\delta : \mathbb{N} \to \mathbb{N}$ je **P-test(Martin-Löf)**, ak

1. $V = \{(m, x) : \ \delta(x) \geq m\}$ **je rekurzívne vyčísliteľná**
2. $\sum_x \{P(x|l(x) = n), \ \delta(x) \geq m\} \leq 2^{-m} \quad \forall n$

$\hookrightarrow$ kritické regióny $V_m = \{x : \delta(x) \geq m\}$ sú vnorené a enumerovateľné

rovnomerné rozdelenie

$$\mathbf{L(x) = 2^{-2l(x)-1}}$$

$$\mathbf{L_n(x)} = \begin{cases} 2^{-n}, & l(x) = n \\ 0 & \text{inak} \end{cases}$$

2. $\rightsquigarrow \sum_{x \in V_m} L_n(x) \leq 2^{-m} \quad \rightsquigarrow \quad d(\{x : l(x) = n, x \in V_m\}) \leq 2^{n-m}$

Example

$x = 1x_2 1x_4 1x_6 \ldots$ **nie je náhodné vzhľadom k rovnomernému rozdeleniu**

$$\delta(x) = \begin{cases} \max\{i : x_1 = x_3 = x_5 = \ldots = x_{2i-1} = 1\}, & x_1 \neq 0 \\ 0, & x_1 = 0 \end{cases}$$

δ je TEST:

1. δ je rekurzívne vyčísliteľná
2. kritický región je malý (2^{-m})

   ak $\delta(x) = m, l(x) = n \geq 2m - 1$, tak máme

   $\begin{cases} 2^{m-1} & \text{možností (2m-1)dlhého prefixu} \\ 2^{n-(2m-1)} & \text{zvyškov} \end{cases}$

   $d(\{x : \delta(x) \geq m, l(x) = n\}) = 2^{m-1} \cdot 2^{n-(2m-1)} = 2^{n-m}$

Definition

univerzálny P-test je test $\delta_0(\cdot|P)$ taký, že ku každému P-testu δ existuje konštanta c_P taká, že $\forall x\ \delta_0(x|P) \geq \delta(x) - c_P$

binárny reťazec je náhodný vzhľadom k univerzálnemu P-testu, ak je (až na pár výnimiek) náhodný vzhľadom k ľubovoľnému testu

$\delta_0(\cdot|P)$ $U_m = \{(m,x) : \delta_0(x|P) \geq m\}$
δ $V_m = \{(m,x) : \delta(x) \geq m\}$

$$V_{m+c} \subseteq U_m$$

Ukážeme, že **univerzálny test existuje**

1. P-testy vieme enumerovať
2. iteratívna definícia univerzálneho P-testu

P-testy vieme enumerovať

zmena enumerácie $\Phi_1, \Phi_2, \ldots$ PRF na enumeráciu $\delta_1, \delta_2, \ldots$ P-testov

- $\Phi \rightsquigarrow \Psi$(indexy vynechané)

  Φ, Ψ majú rovnaké obory hodnôt
  ak je Ψ_n def $\rightsquigarrow \Psi_1, \ldots, \Psi_{n-1}$ tiež def

  "dovetail" - postupné simulovanie $\Phi(1), \Phi(2), \ldots$

  $$\begin{array}{ccc} \Phi(1) & & \\ \Phi(1) & \Phi(2) & \\ \Phi(1) & \Phi(2) & \Phi(3) \\ \vdots & \vdots & \vdots \end{array}$$

  $\Psi(1) = \Phi(i_1)$, kde $\Phi(i_1)$ je prvý výpočet, kt. zastal

  $\ldots$

  $\Psi(j) = \Phi(i_j)$, kde $\Phi(i_2)$ je j-ty výpočet, kt. zastal

- pomocou Ψ definujeme test δ aproximovaním zospodu

pomocou Ψ definujeme test δ aproximovaním zospodu

$\delta[1..\infty]$ - pole aktuálnych hodnôt $\delta(1), \delta(2), \ldots$

1. $\delta[x] = 0 \quad \forall x;\ i \leftarrow 0;$
2. $i \leftarrow i + 1$; vypočítaj $\Psi(i)$, nech $\Psi(i) = (m, x)$;
3. if $\delta(x) \geq m$ then goto 2 else $\delta[x] \leftarrow m$;
4. if $\exists k \in \{1, 2, \ldots, m\} \quad \sum\{P(y|l(y) = l(x)) : \delta(y) \geq k\} > 2^{-k}$
   then $\delta[x] \leftarrow 0$; halt
   else goto 2.

1. $\delta[x] = 0 \quad \forall x;\ i \leftarrow 0;$
2. $i \leftarrow i + 1$; vypočítaj $\Psi(i)$, nech $\Psi(i) = (x, m)$;
3. if $\delta(x) \geq m$ then goto 2 else $\delta[x] \leftarrow m$;
4. if $\sum\{P(y|l(y) = l(x)) : \delta(y) \geq k\} > 2^{-k}$ pre nejaké $k = 1, 2, \ldots, m$
   then $\delta[x] \leftarrow 0$; halt
   else goto 2.

- ak je oborom hodnôt Ψ test $\rightsquigarrow$ neskončíme, ale zospodu ho aproximujeme
- ak Ψ diverguje$\rightsquigarrow$ neskončí a nemení vypočítané (je to enumerovateľná mn.)
- ak oborom hodnôt nie je test $\rightsquigarrow$ podmienka v kroku 4. sa niekedy splní a končíme

Ak zbehneme na $\forall\ \Phi_1, \Phi_2, \ldots$, "dostaneme" P-testy $\delta_1, \delta_2, \ldots$

Theorem (univerzálny P-test)

Nech $\delta_1, \delta_2, \ldots$ je enumerácia P-testov. Potom
$\delta_{\mathbf{0}}(\mathbf{x}|\mathbf{P}) = \max\{\delta_{\mathbf{y}}(\mathbf{x}) - \mathbf{y} : \ \mathbf{y} \geq \mathbf{1}\}$ *je univerzálny P-test*

- $\delta_0(x|P)$ je totálna
- $V = \{(m, x) : \delta_0(x|P) \geq m\}$ je enumerovateľná
- kritické regióny sú dosť malé

$$\sum_{l(x)=n} \{P(x|l(x) = n) : \delta_0(x|P) \geq m\}$$
$$\leq \sum_{y=1}^{\infty} \sum_{l(x)=n} \{P(x|l(x) = n) : \delta_y(x) \geq m + y\}$$
$$\leq \sum_{y=1}^{\infty} 2^{-m-y} = 2^{-m}$$

- δ_0 majorizuje aditívne

Theorem (konkrétny univerzálny test)

$\delta_0(x|L) = l(x) - C(x|l(x)) - 1$ *je univerzálny L-test pre rovnomerné rozdelenie L.*

že je to test

- $\{(m, x) : \delta_0(x|L) \geq m\}$ je enumerovateľná
- $d(\{x : \ \delta_0(x|L) \geq m\}) \leq 2^{l(x)-m} - 1$

že je univerzálny

- $\forall \delta \ \exists c \ \ \delta_0(x|L) \geq \delta(x) - c$

že $\delta_0(x|L) = l(x) - C(x|l(x)) - 1$ je univerzálny

$A = \{z : \ \delta(z) \geq \delta(x)\}$ $\quad x \in A(\text{fix});\ \delta = \delta_y$ v štandardnom enumerovaní

- $y, l(x), \delta(x) \quad \rightsquigarrow$ vieme enumerovať prvky z A
- plus j-index x v $A \quad \rightsquigarrow$ máme x $\qquad //j \leq l(d(A)) \leq l(x) - \delta(x)$
  $s = 0 \ldots 01j, \quad |s| = l(x) - \delta(x)$
- $l(x), l(s) \quad \rightsquigarrow$ vieme $\delta(x)$

$$C(x|l(x)) \leq \underbrace{l(x) - \delta(x)}_{l(s)} + \underbrace{2l(y) + 1}_{y} \qquad c = 2l(y) + 2 \qquad \square$$

Definition

Vezmime $\delta_0(x|L) = l(x) - C(x|l(x)) - 1$ ako referenčný univerzálny test vzhľadom k uniformnej distribúcii L. Reťazec nazveme c-náhodný, ak $\delta_0(x|L) \leq c$

$$C(x|l(x)) \leq C(x) \leq C(x|l(x)) + 2C(l(x) - C(x|l(x))) + O(1)$$

Definition

Nekonečná postupnosť $\in \{0,1\}^*$ je **normálna**, ak $\forall k$ je frekvencia výskytu bloku y dĺžky k v limite 2^{-k}

- normálna postupnosť nie je nutne náhodná – Champernow 12345678910111213 . . .
- náhodná postupnosť je normálna

v *konečnom prípade* — každý blok malej veľkosti sa v ňom vyskytuje približne rovnako

K() - prefixná KZ; nateraz

$$\mathbf{C(x|y) \leq K(x|y) \leq C(x|y) + 2\log(C(x|y)) + 1}$$

Definition

Trieda **deficiency funkcií** $\delta : N \to N$ takých, že

$$\mathbf{K(n, \delta(n)|n - \delta(n)) \leq c_1}$$

- $n - \delta(n)$ vieme rozobrať na $n, \delta(n)$ ale aj $n, \delta(n)$ vieme zakódovať do $n - \delta(n)$
- fixneme c_1 tak, aby vyhovovalo pre bežné deficiency: $\log n, \log\log n, \sqrt{n}, \ldots$

Bloky v reťazcoch

Theorem $\exists$ konštanta c taká, že $\forall$ deficiency funkciu δ, $\forall n$ a $x \in \{0,1\}^n$

ak $\mathbf{C(x) \geq n - \delta(n)}$ tak $\left|\#_1(\mathbf{x}) - \frac{\mathbf{n}}{2}\right| \leq \sqrt{\frac{3}{2}\frac{(\delta(\mathbf{n})+\mathbf{c})\mathbf{n}}{\log \mathbf{e}}}$

Theorem $\exists$ konštanta c taká, že $\forall n$ a $x \in \{0,1\}^n$

ak $\left|\#_1(\mathbf{x}) - \frac{\mathbf{n}}{2}\right| \leq \mathbf{2}^{-\delta(\mathbf{n})-\mathbf{c}}\sqrt{\mathbf{n}}$ tak $\mathbf{C(x) \leq n - \delta(n)}$

Theorem $\exists$ konštanta c taká, že $\forall n$ a $x \in \{0,1\}^n$

ak $\left|\#_1(\mathbf{x}) - \frac{\mathbf{n}}{2}\right| = \mathbf{j}$ tak $\mathbf{C(x|n) \leq n - 1/2 \log n + K(j|n) + c}$

Theorem Nech $l(y) = \ell$, $\ell \leq \log n$. $\exists$ konštanta c tak, že $\forall n \forall x \in \{0,1\}^n$

ak $\mathbf{C(x) \geq n - \delta(n)}$ potom $|\#_{\mathbf{y}}(\mathbf{x}) - \mathbf{pn}| \leq \sqrt{\alpha \mathbf{pn}}$

kde $\alpha = \mathbf{[K(y|n) + \log \ell + \delta(n) + c]3\ell / \log e}$

Theorem

$\exists$ *konštanta* c *taká, že* $\forall$ *deficiency funkciu* δ, $\forall n$ *a* $x \in \{0,1\}^n$

ak $\mathbf{C(x) \geq n - \delta(n)}$ *tak* $\left|\#_1(\mathbf{x}) - \frac{\mathbf{n}}{2}\right| \leq \sqrt{\frac{3}{2}\frac{(\delta(\mathbf{n})+\mathbf{c})\mathbf{n}}{\log \mathbf{e}}}$

Chernoff $\mathbf{Pr(|S_n - pn| > m) \leq 2e^{-m^2/3pn}}$ //mince $p = 1/2$

- $A = \{x \in \{0,1\}^n : \ |\#_1(x) - n/2| > m\}, \quad d(A) \leq 2^{n+1} e^{-2m^2/3n}$
- zvolíme m tak, aby $\frac{2m^2 \log e}{3n} = \delta(\mathbf{n}) + \mathbf{c}$
- **popis** x

  - selfdelimiting **s** $\quad n - \delta(n) \rightsquigarrow \delta(n), n \quad \mathbf{c_1}$

  - **i** je index x v A

$$l(i) \leq \log d(A) \leq (n+1) - (2m^2 \log e)/3n = n + 1 - \delta(n) - c$$
$$|0 \ldots 0\mathbf{si}| = n + 1 - \delta(n) - c + c_1$$

$$C(x) \leq C_T(x) + c_T \leq n + \mathbf{1} - \delta(n) - \mathbf{c} + \mathbf{c_1} + \mathbf{c_T}$$

pre $\mathbf{c = c_1 + c_T + 2}$ $\quad C(x) < n - \delta(n)$[Spor] $\qquad |\#_1(x) - n/2| \leq m \quad \square$

Theorem

∃ *konštanta* c *taká, že* $\forall n$ *a* $x \in \{0,1\}^n$

$$\textit{ak} \quad \left|\#_1(x) - \frac{n}{2}\right| \leq 2^{-\delta(n)-c}\sqrt{n} \quad \textit{tak} \quad C(x) \leq n - \delta(n)$$

- **Zvolíme** $m = 2^{-\delta(n)-c}\sqrt{n}$
- $A = \{x \in \{0,1\}^n : \ |\#_1(x) - n/2| \leq m\}$
  $d(A) \leq (m+1)\binom{n}{n/2} \leq c_2 \frac{2^n m}{\sqrt{n}}$ Stirling $n! \approx \sqrt{2\pi n}(n/e)^n$
- i je index x v A
  určenie $A : n, m, \delta(n) \quad \rightsquigarrow \quad n - \delta(n), c_1$

$$C_T(x) = \log(d(A)) + c_1 = n - \delta(n) - c_1 + \log c_2$$

$$C(x) \leq C_T(x) + c_T \leq n - \delta(n)$$

Theorem

Existuje konštanta c taká, že $\forall n$ a $x \in \{0,1\}^n$

$$\textit{ak} \quad \left|\#_1(x) - \frac{n}{2}\right| = j \quad \textit{tak} \quad C(x|n) \leq n - 1/2 \log n + K(j|n) + c$$

- $A = \{x \in \{0,1\}^n : \ |\#_1(x) - n/2| = j\}$
- $d(A) \leq 2\binom{n}{n/2} \leq c_3 \frac{2^n}{\sqrt{n}}$
- i je index x v A
  –enumerácia $A \quad \rightsquigarrow \quad j, n, \ d(A)$

$$C_T(x|n) \leq \log d(A) + K(j|n) \leq n - 1/2 \log n + \log c_3 + K(j|n)$$

$$C(x|n) \leq C_T(x|n) + c_T \leq n - 1/2 \log n + K(j|n) + c$$

□

Theorem

Nech $l(y) = \ell$, $\ell \leq \log n$. *Existuje konštanta* c *tak, že* $\forall n \forall x \in \{0,1\}^n$
ak $\mathbf{C(x) \geq n - \delta(n)}$ *potom* $|\#_{\mathbf{y}}(\mathbf{x}) - \mathbf{pn}| \leq \sqrt{\alpha \mathbf{pn}}$

kde $\alpha = \mathbf{[K(y|n) + \log \ell + \delta(n) + c]3\ell / \log e}$

Nech dĺžka $x, l(x) = n$, je násobkom dĺžky $y, l(y) = \ell$ $\mathbf{n = N\ell}$

- x ako cyklický kruh, ℓ rôznych delení $\#_{\mathbf{y}}(\mathbf{x}, \mathbf{i})$ neprekrývajúce sa výskyty
- $\mathbf{A} = \{x \in \{0,1\}^n : |\#_y(x,i) - pN| > m\}$
- $d(A) \leq 2^{n+1} e^{-m^2/3pN}$
  zvolíme m tak aby $\frac{\mathbf{m^2 \log e}}{\mathbf{3pN}} = \mathbf{K(\langle y, i \rangle | n) + \delta(n) + c}$
- **enumerácia A** $i, y, m \quad \rightsquigarrow \quad \mathbf{K(\langle y, i, \delta(n), n \rangle | n - \delta(n)) \leq K(\langle y, i \rangle | n) + c_1}$
- **index** i **prvku** x **v** A

$$\log d(A) \leq \log(2^{n+1} e^{-m^2/3pN}) = n+1 - \frac{m^2 \log e}{3pN} \leq n+1 - K(\langle y, i \rangle | n) - \delta(n) - c$$

$$C(x) \leq C_T(x) + c_T \leq n + 1 - \delta(n) - c + c_1 + c_T$$

pre $c = c_1 + c_T + 2 \quad C(x) < n - \delta(n)$ $\qquad |\#_y(x,i) - pN| \leq m$

$$|\#_y(x,i) - pN| \le m \qquad \frac{m^2 \log e}{3pN} = K(\langle y,i\rangle|n) + \delta(n) + \mathbf{c}$$

- $|\#_y(x,i) - pN| \le \sqrt{\frac{K(\langle y,i\rangle|n)+\delta(n)+c}{\log e} \cdot 3pN}$
- $K(\langle y,i\rangle|n) \le K(y|n) + K(i|n) + O(1); \quad K(i|n) \le \log \ell + O(1)$
- $|\#_y(x) - pn| = \sum_{i=0}^{\ell-1} |\#_y(x,i) - pN| \le \ell\sqrt{\frac{K(y,|n)+\log \ell+\delta(n)+c}{\log e} \cdot 3p\frac{n}{\ell}}$

$$|\#_{\mathbf{y}}(\mathbf{x}) - \mathbf{pn}| \le \sqrt{\alpha \mathbf{pn}}, \alpha = [\mathbf{K(y|n)} + \log \ell + \delta(\mathbf{n}) + \mathbf{c}]\mathbf{3}\ell/\log \mathbf{e} \quad \square$$

Bloky v reťazcoch

//opakovanie

Theorem $\exists$ konštanta c taká, že $\forall$ deficiency funkciu δ, $\forall n$ a $x \in \{0,1\}^n$

ak $\mathbf{C(x) \geq n - \delta(n)}$ tak $\left|\#_1(\mathbf{x}) - \frac{\mathbf{n}}{2}\right| \leq \sqrt{\frac{3}{2}\frac{(\delta(\mathbf{n})+\mathbf{c})\mathbf{n}}{\log \mathbf{e}}}$

Theorem $\exists$ konštanta c taká, že $\forall n$ a $x \in \{0,1\}^n$

ak $\left|\#_1(\mathbf{x}) - \frac{\mathbf{n}}{2}\right| \leq \mathbf{2}^{-\delta(\mathbf{n})-\mathbf{c}}\sqrt{\mathbf{n}}$ tak $\mathbf{C(x) \leq n - \delta(n)}$

Theorem $\exists$ konštanta c taká, že $\forall n$ a $x \in \{0,1\}^n$

ak $\left|\#_1(\mathbf{x}) - \frac{\mathbf{n}}{2}\right| = \mathbf{j}$ tak $\mathbf{C(x|n) \leq n - 1/2 \log n + K(j|n) + c}$

Theorem Nech $l(y) = \ell$, $\ell \leq \log n$. $\exists$ konštanta c tak, že $\forall n \forall x \in \{0,1\}^n$

ak $\mathbf{C(x) \geq n - \delta(n)}$ potom $|\#_{\mathbf{y}}(\mathbf{x}) - \mathbf{pn}| \leq \sqrt{\alpha \mathbf{pn}}$

kde $\alpha = \mathbf{[K(y|n) + \log \ell + \delta(n) + c]3\ell / \log e}$

bloky v reťazcoch

- Nech $l(y) = \ell$, $\ell \leq \log n$. Existuje konštanta c tak, že $\forall n \forall x \in \{0,1\}^n$
  ak $\mathbf{C(x) \geq n - \delta(n)}$ potom $|\#_{\mathbf{y}}(\mathbf{x}) - \mathbf{pn}| \leq \sqrt{\alpha \mathbf{pn}}$
  kde $\alpha = \mathbf{[K(y|n) + \log \ell + \delta(n) + c]3}\ell / \log \mathbf{e}$

AK sa blok y, v i-tom rozdelení objavuje s frekvenciou pN, $p = 1/2^\ell$ a $n = N\ell$, potom na určenie x stačí **n, y, i** a **poradie v množine** A takých reťazcov dĺžky n

$$d(A) = \binom{N}{pN}(2^\ell - 1)^{N-pN} = O\left(\frac{2^{N\ell}}{\sqrt{p(1-p)N}}\right)$$

$$C(x|\langle n, y\rangle) \leq n - 1/2 \log n + 1/2(\ell + 3 \log \ell) + O(1)$$

Theorem (Blok núl v náhodných reťazcoch)

Nech x dĺžky n spĺňa $C(x) \geq n - \delta(n)$. Potom pre dostatočne veľké n platí, že v x sa vyskytuje každý blok veľkosti

$$\ell = \log \mathbf{n} - \log\log \mathbf{n} - \log(\delta(\mathbf{n}) + \log(\mathbf{n})) - \mathbf{O(1)}$$

- Ak $\delta(n) = O(\log n)$, tak sa každý blok dĺžky $\log n - 2\log\log n - O(1)$ v x vyskytuje aspoň raz
- Ak $\delta(n) = O(\log\log n)$, tak $\forall \epsilon > 0$ a dosť veľké n každý reťazec x dĺžky n obsahuje $y = 0^{\ell}$, pričom

$$K(y|n) = O(\log n), \ell = \log n - (1+\epsilon)\log\log n + O(1)$$

//o kvalite odhadu bloku núl...

$x = uvw$ také, že $C(x) \geq n - \delta(n)$, popíšeme

- v ($K(v)$ bitov), uw ($n - l(v)$ bitov)
- popis $l(u)$ $\log n + \log\log n + 2\log\log\log n + O(1)$

$$C(x) \leq \underbrace{n - l(v)}_{uw} + K(v) + \log n + \underbrace{\log\log n + 2\log\log\log n}_{(1+o(1))\log\log n} + O(1)$$

pre $K(v) = o(\log\log n)$

$$n - \delta(n) \leq C(x) \leq n - l(v) + o(\log\log n) + \log n + (1 + o(1))\log\log n + O(1)$$

$$l(v) \leq \delta(n) + \log n + (1 + o(1))\log\log n$$

Zložitý reťazec $C(x) = n + O(1)$

- **neobsahuje** podreťazec núl dĺžky $\log \mathbf{n} + (\mathbf{1} + \epsilon)\log\log \mathbf{n}$ pre dostatočne dlhé a pravidelné n.
- **Musí obsahovať** reťazec núl dĺžky $\log \mathbf{n} - (\mathbf{1} + \epsilon)\log\log \mathbf{n} + \mathbf{O(1)}$

Vieme, že C

- neohraničene rastie
- rastie pomalšie ako každá monotónne rastúca neohraničená RE
- nie je rekurzívna, ale dá sa aproximovať

Theorem

O Kolmogorovskej zložitosti C platí

1. $A = \{(x, a) : \ C(x) \leq a\}$ *je RE ale nie R*
2. $\forall$ *čiastočne rekurzívna* $\Phi(x)$, *kt. je dolným odhadom* $C(x)$, *je ohraničená*
3. *Nech* $f(x)$ *je totálne rekurzívna,* $g(x) \leq f(x) \leq l(x)\ \forall x$ *a nejakú neohraničenú monotónnu funkciu* $g(x)$. *Potom*

$$B = \{x : \ C(x) \leq f(x)\}$$

je jednoduchá/simple (RE, B^C *nekonečná, kt. neobsahuje žiadnu nekonečnú RE podmnožinu)*

Každá PRF $\Phi(x)$, kt. je dolným odhadom $C(x)$, je ohraničená

Nech $D = \{x : \ \Phi(x) \leq C(x)\}$

- D konečná
- D nekonečná, $\Phi = \Phi_k$, $\Phi(x)$ neohraničená

  T enumeruje jej definičný obor bez opakovania

  $g(n) = \min\{x : \ \Phi(x) \geq n\}$ //totálna rekurzívna

  $n \leq \Phi_k(x) \leq C(x) \leq l(n) + l(\bar{k}) + O(1)$

$$n \leq l(n) + l(\bar{k}) + O(1) \qquad \textit{spor}$$

Nech $f(x)$ je totálne rekurzívna, $g(x) \leq f(x) \leq l(x)\ \forall x$ a nejakú neohraničenú monotónnu funkciu $g(x)$. Potom

$$B = \{x : \ C(x) \leq f(x)\}$$

je *jednoduchá/simple* (RE, B^C nekonečná, kt. neobsahuje žiadnu nekonečnú RE podmnožinu)

- B je RE:

  $f(x)$ rekurzívna: testom programov dĺžky $\leq f(x)$ vymenovávame $x \in B$
- B^C je nekonečná (nestlačiteľné reťazce)
- B^C neobsahuje nekonečnú RE podmnožinu

  NECH D je nekonečná RE, $D \subseteq B^C$

  – zúženie $f_D(x)$ je PRF

  – pre $x \in D \subseteq B^C$ $f_D(x) = f(x) < C(x)$, preto je $f_D(x)$ ohraničená

  – pritom $f(x)$ je neohraničená $\rightsquigarrow$ D je konečná

Example (Nerozhodnuteľnosť nestlačiteľnosťou)

Nech T je formálny systém. Potom existuje konštanta c_T tak, že tvrdenie typu $\mathbf{C(x) \geq c_T}$ je nedokázateľné

- T axiomatizovateľná $\rightsquigarrow$ k_T bitov na popis $\mathbf{C(T) \leq k_T}$
- T bezosporná
- $S_c(x) \sim x$ je lexikograficky najmenšie x dĺžky c pre ktoré $C(x) \geq c$

$$C(S_c(x)) \leq \log c + O(1) \qquad O(1) \sim s$$

- $\forall c\ \exists! x:\ S_c(x) = true$

popis x $\overline{\mathbf{T}}$ popis $\mathbf{S_c}$

$$C(x) \leq \mathbf{2k_T + 1} + \log \mathbf{c} + \mathbf{s} + O(1)$$

spor s tým, že $C(x) > c\ \forall\ c > c_T$

Nerozhodnuteľnosť nestlačiteľnosťou

Dôsl.

Nech B je simple (RE, B^C nekonečná bez nekonečnej RE podmnožiny). Potom množina $D = \{n \mid \text{dá sa dokázať, že } n \in B^C\}$ je konečná.

- D je enumerovateľná
- žiadna nekonečná podmnožina B^C – ani D – nie je RE

Dôsl.

- nekonečne veľa nedokázateľných formúl "$C(x) \geq c_T$"
- nemáme efektívnu procedúru na hľadanie c_T

Hoci vieme, že nerozhodnuteľných príkladov je veľa, nevieme ich hľadať

enumerovateľné vs. rekurzívne množiny

charakteristická postupnosť množiny A, $\chi = \chi_1\chi_2\ldots$, $\chi_i = \begin{cases} 1, & i \in A \\ 0, & i \notin A \end{cases}$

A aj A^C RE, potom $f(i) = \chi_i$ rekurzívna a $C(\chi_{1\ldots n}|n) \leq c_A$

Lemma (Barzdinova lema)

1. *ak A je RE, χ jej charakteristická postupnosť, tak* $C(\chi_{1\ldots n}|n) \leq \log n + c_A$
2. *existuje RE množina, pre ktorú* $C(\chi_{1\ldots n}) \geq \log n$

1. ak A je RE, χ jej charakteristická postupnosť, tak $C(\chi_{1\ldots n}|n) \leq \log n + c_A$

   enumerácia A pomocou PR funkcie Φ $\quad A = \{x \mid \Phi(x) < \infty\}$

   ukončenie pomocou počtu jednotiek $m \leq n$

2. existuje RE množina B, pre ktorú $C(\chi_{1\ldots n}) \geq \log n$

   existencia B diagonalizáciou(pomocou aditívne optimálnej Φ_0)

$$\chi_i = \begin{cases} 1, & \Phi_0(i,i) = 0 \\ 0, & \Phi_0(i,i) \neq 0 \text{ alebo } \Phi_0(i,i) = \infty \end{cases}$$

Ak $C(\chi_{1\ldots n}) < \log n$ potom $\exists$ program p : $l(p) < \log n$

problém $\Phi_0(p,p) = \chi_p$

Example (Diofantické rovnice)

$\Delta = \Delta_1\Delta_2\ldots$, $\Delta_i = 1$ ak i-ta rovnica má riešenie, inak 0

$$C(\Delta_{1\ldots n}|n) \leq \log n + O(1)$$

Example (nestlačiteľné reťazce cez zastavenie)

$$K_0 = \{\langle x, y\rangle\rangle : \ \Phi_x(y) < \infty\}$$

d – počet prvkov $\in K_0$, ktoré sú menšie ako 2^n

AK poznáme d, vieme nájsť $\forall$ výpočty $\Phi_x(y), |\langle x, y\rangle| < 2^n$, ktoré zastanú

AK χ je charakteristická pre K_0, tak $\chi_{1\ldots 2^n}$ vypočítame z d: l(p)≤l(d)+O(1)≤n+c

ALE $\exists$ enumerovateľná množina s χ, $C(\chi_{1\ldots m}) \geq \log m \ \forall m$

$\rightsquigarrow$ existuje konštanta c', že program p je c'-nestlačiteľný

Algoritmická informácia o y obsiahnutá v x

$$I_c(x:y) = C(y) - C(y|x)$$

- $C(x|x) = 0 \quad I_C(x:x) = C(x)$
- $\exists$ reťazce $C(x|n) > n$, je veľa $C(n) \geq l(n)$
- Ak $x, l(x) = n$ a n je náhodné číslo, tak

$$I_c(x:n) = C(n) - C(n|x) \geq l(n)$$
$$I_c(n:x) = C(x) - C(x|n) \leq n - n = 0$$

Theorem

$\forall x, y \in \mathbb{N} \quad C(x,y) = C(x) + C(y|x) + O(\log C(x,y))$

$\leq \qquad \log C(x,y) \geq \max\{\log C(x), \log C(y|x)\} \quad \checkmark$

pre $c \geq 0$, $\quad C(x,y) \boxed{\geq} C(x) + C(y|x) - c \log C(x,y)$

kvôli sporu: $\forall c \exists x,y: \; C(y|x) > C(x,y) - C(x) + c \log C(x,y)$

- $A = \{\langle u,z\rangle : \; C(u,z) \leq C(x,y)\}$

  ak vieme $C(x,y)$, tak A je enumerovateľná

- $A_x = \{z : \; C(x,z) \leq C(x,y)\}$

  ak vieme $C(x,y)$, tak A_x je enumerovateľná

- pre **popis** y stačí x a poradové číslo y v A_x , $C(x,y)$
  $C(y|x) \leq l(d(A_x)) + 2l(C(x,y)) + O(1)$
  $\rightsquigarrow \quad \mathbf{d(A_x) > 2^\ell}, \; \ell = \mathbf{C(x,y) - C(x) + (c-2)l(C(x,y))}$

- **krátky popis** x

$\left.\begin{matrix} C(x,y) \\ \ell \end{matrix}\right\}$ u je kandidát na x ak $A_u = \{z : C(u,z) \leq C(x,y)\} \& 2^\ell < d(A_u)$

U množina kandidátov, $x \in U$ $\qquad$ $\{\langle u, z\rangle : \ u \in U, z \in A_u\} \subseteq A$

$$d(A) \leq 2^{C(x,y)+O(1)} \quad \rightsquigarrow \quad d(U) < \frac{d(A)}{2^\ell} \leq \frac{2^{C(x,y)+O(1)}}{2^\ell}$$

x vieme zrekonštruovať z $C(x,y)$,ℓ, poradia x v U

$$C(x) < 2l(C(x,y)) + 2l(\ell) + C(x,y) - \ell + O(1)$$
$$C(x) < 2l(C(x,y)) + 2l(\ell) + C(x,y) - \mathbf{C(x,y) + C(x) - (c+2)l(C(x,y))} + O(1)$$

Pre veľké c dostaneme spor $C(x) < C(x)$ $\qquad$ □

Dôsledok

Až na aditívnych $O(\log C(x,y))$ *platí* $C(x) - C(x|y) = C(y) - C(y|x)$
Preto $|I_c(x:y) - I_c(y:x)| = O(\log(C(x,y)))$

Algoritmická prefixová zložitosť

čiastočne rekurzívna prefixová funkcia $\Phi : \{0,1\}^* \to \mathbb{N}$

AK $\Phi(p) < \infty \quad \& \quad \Phi(q) < \infty$ TAK p nie je vlastný prefix q

Enumerácia PR prefixových fcií

- $T_1, T_2, \ldots$ enumerácia PR $\Phi_1, \Phi_2, \ldots$ konečný vstup
- $T'_1, T'_2, \ldots$ enumerácia PRP $\Psi_1, \Psi_2, \ldots$ nekonečný vstup $b_1 b_2 \ldots$
- **halting input** T' zastane po dočítaní $b_1 \ldots b_m$, ale pred čítaním b_{m+1}

modifikácia T na T', vstup $b_1 b_2 \ldots$

1. $p \leftarrow \epsilon$
2. dovetail $T(pq)$ $\forall q \in \{0,1\}^*$; nech $\Phi(q)$ zastane prvé
3. ak $q = \epsilon$, na výstup $\Phi(p)$ a stop
   ak $q \neq \epsilon$, načítaj ďalšie b zo vstupu; $p \leftarrow pb$; choď na 2

- existuje *univerzálny* $T'_0 \quad \Psi_0$

$$C_{\psi_0}(x|y) \leq C_{\psi}(x|y) + c_{\psi}$$

Univerzálny prefix stroj $U(\langle y, \langle n, p\rangle\rangle) = T'_n(y, p)$

- $K(x, y) := K(\langle x, y\rangle)$

$$\mathbf{K(x, y) \leq K(x) + K(y) + O(1)}$$

- $\mathbf{C(x|y) \leq K(x|y) \leq C(x|y) + 2\log C(x|y) + O(1)}$

$$\begin{aligned}\mathbf{K(x|y)} &\leq \mathbf{C(x|y) + C(C(x|y))) + \cdots + O(1) + r}\\ &\leq \mathbf{C(x|y) + C(C(x|y)) + O(\log C(C(x|y))) + O(1)}\\ &\leq \mathbf{C(x|y) + l^*(C(x|y)) + O(1)}\end{aligned}$$

- $\mathbf{K(x) \leq \log^* n + n + l(n) + l(l(n)) + \cdots + O(1)}$

$$\begin{aligned}\mathbf{K(x)} &\leq \mathbf{K(x|n) + K(n) + O(1)}\\ &\leq \mathbf{K(x|n) + \log^* n + l(n) + l(l(n) + \cdots + O(1))}\end{aligned}$$

Theorem

Platí

1. $\forall n \max\{K(x) : \ l(x) = n\} = n + K(n) + O(1)$
2. $\forall r$ ***je počet slov*** x ***dĺžky*** n***, pre ktoré*** $K(x) \leq n + K(n) - r$***, nanajvýš*** $2^{n-r+O(1)}$

Dôkaz

1. $\boxed{\leq}\ K(n)x$

   $\boxed{\geq}\ 2^n$ reťazcov dĺžky n, podľa 2. $n + K(n) - r$ nestačí

2. **Nech** reťazec x spĺňa $\mathbf{K(x) \leq n + K(n) - r}$

   využijeme niečo, čo ešte nevieme ...

$$K(x) + K(n|x, K(x)) = K(n) + K(x|n, K(n)) + O(1)$$

Vezmeme $\mathbf{K(n|x,K(x))} = O(1)$ a pre $\mathbf{n=l(x)}$ potom

$$K(x|n, K(n)) \leq n - r + O(1)$$

Len $2^{n-r+O(1)}$ to môže spĺňať. □

$$\mathbf{C(x) = \min\{i : K(x|i) \leq i\} + O(1) = K(x|C(x)) + O(1)}$$

$\boxed{\geq}$

- x^* najkratší program pre x má dĺžku $C(x)$
- $K(x|C(x)) \leq C(x) + O(1)$
- $C(x) \geq \min\{i : K(x|i) \leq i\} + O(1)$

$\boxed{\leq}$

$C(x) \leq \min\{i : K(x|i) \leq i\} + O(1)$, resp. $K(x|i) \leq i$ potom $C(x) \leq i + O(1)$

$$\begin{cases} i - 1 > l(p), & \underbrace{0 \ldots 01p}_{i} \\ i - 1 \leq l(p), & 1p \end{cases}$$

$C(x) \leq C_T(x) + c_T \leq i + O(1)$

K, C sú nerekurzívne

ak f je rekurzívna, tak $K(f(x)|x) = O(1)$

ak $f(x) \in \{0,1\}$, tak $K(f(x)|x) = O(1)$

} $K(f(x)|x)$ hovorí niečo o f

Def.

Zložitosť **K(f)** *funkcie* f *je* **K(f(x)|x)**

K(K(x)|x) //zložitosť zložitosti

1. ak $l(x) = n$ tak $K(x) \leq n + 2\log\log n + O(1)$ $1^{l(n)}nx$

2. $\mathbf{K(K(x)|x) \leq K(K(x)|n) + O(1) \leq \log n + O(1)}$
   ak viem n, $|n - K(x)| < n$, preto stačí povedať $d = n - |n - K(x)|$

   $$\log d + 2\log\log d + O(1) \leq \log n + O(1)$$

3. **$\forall$n$\exists$x dĺžky n také, že K(K(x)|x) $\geq \log n - \log\log n + O(1)$**

Theorem

$\forall \mathbf{n} \exists \mathbf{x}$ ***dĺžky*** $\mathbf{n}$ ***také, že*** $\mathbf{K(K(x)|x) \geq \log n - \log\log n + O(1)}$

//analogicky pre $C(C(x)|x)$

Dôkaz

- U je referenčný stroj $\quad U(\langle y, \langle n, p\rangle\rangle) = T'_n(y, p)$
- fixneme dostatočne veľké n, všetky reťazce sú dĺžky n
- $s = \max_{l(x)=n} \min\{l(p) \mid U(p, x) = K(x)\}$

// maximálna hodnota K(K(x)|x)

Chceme $s \geq \log n - \log\log n + O(1)$ $\qquad$ $//K(K(x)|x) \leq s \leq \log n + O(1)$

vhodný program p pre x - ak pre nejaké q

- $l(p) \leq s$ určite existuje
- U počíta $l(q)$ z p, ak pozná x $l(p) = K(K(x)|x) \leq s$
- U počíta x z q $l(q) = K(x)$

M_i je množina tých x, že $\exists$ aspoň i vhodných programov pre x

$$\emptyset = M_{j+1} \subseteq M_j \subseteq \ldots \subseteq M_0 = \{0,1\}^n, \quad M_j \neq \emptyset$$

Ukážeme $l(d(M_i)) \leq l(d(M_{i+1})) + 5 \log n$, čo spolu s $j \leq 2^{s+1}$ dáva

$$\boxed{\mathbf{s \geq \log n - \log\log n + O(1)}}$$

Chceme $I(d(M_i)) \leq I(d(M_{i+1})) + 5 \log n$,

Ak vieme $\mathbf{i, s, n, d(M_{i+1}), I(d(M_i))}$

1. enumerujeme prvky z M_{i+1}
2. vygenerujeme dostatočne veľa prvkov z $M_i - M_{i+1}$

   $\forall z \in M_i - M_{i+1}$ nájdeme *všetkých* i vhodných programov

   vhodný program p, pre kt. $U(p, z)$ je minimálne, spĺňa $K(z) = U(p, z)$

   $\log d(M_i - M_{i+1}) \geq \log d(M_i) - 1$ inak platí, čo chceme, triviálne

   zoberieme z_{max} : $K(z_{max}) \geq I(d(M_i)) - 1$
3. nech $x = z_{max}$. Potom pre popis x stačí

   diskusia $\rightsquigarrow$ $O(1)$

   popis $d(M_{i+1})$ $\rightsquigarrow$ $I^*(n) + I(d(M_{i+1}))$

   popis $I(d(M_i))$ $\rightsquigarrow$ $I^*(n)$

   popis i, n, s $\rightsquigarrow$ $I^*(n), I^*(n), I^*(\log n + 2 \log \log n)$

$$\mathbf{I(d(M_i)) - 1 \leq K(x) \leq 4I^*(n) + O(I^*(\log n)) + I(d(M_{i+1})) + O(1)}$$

$$\mathbf{I(d(M_i)) \leq I(d(M_{i+1})) + 5 \log n + O(1)}$$

Vieme,

- $C(C(x)) \leq \log n + O(1)\ \forall x$
- $C(C(x)|x) \geq \log n - \log\log n + O(1)$ pre niektoré x – pre aké?

$x, C(x) \geq n - k$

- $C(C(x)|x) \leq C(k) + O(1)$, $\mathbf{C(k) \leq \log k + O(1)}$
- ak platí veta 42, tak

$$\mathbf{C(k)} \geq C(C(x)|x) \geq \log \mathbf{n} - \log\log \mathbf{n} + \mathbf{O(1)}$$

$$\Downarrow$$

$$k = \Omega(\frac{n}{\log n}) \rightsquigarrow \boxed{C(x) \leq n - \Omega(\frac{n}{\log n})} \quad \textbf{x nie je náhodné}$$

Nech $\mathbf{c'}$ je konštanta.

$$C(C(x)|x) \geq \log n - \log\log n + O(1)$$

$$\Downarrow$$

$$\forall k \leq n : C(k|x) \leq c',\ C(x) \notin [k - \delta, k + \delta],\ \delta = O(n/\log n)$$

- $g(<< y, z >, k >) =< p/q >$ píšeme ako $g(y/z, k) = p/q$
- (čiastočná) funkcia $f : \mathbb{Q} \to \mathbb{R}$ je upper semicomputable/co-enumerable ak existuje rekurzívna $g(x, k)$ nerastúca v k tak, že $f(x) = \lim_{k\to\infty} g(x, k)$

  f je enumerovateľná/lower semicomputable ak $-f$ je upper semicomputable; semicomputable ak je lower alebo upper semicomputable
- funkcia $f : \mathbb{Q} \to \mathbb{R}$ je rekurzívna ak existuje rekurzívna $g(x, k)$ tak, že $|f(x) - g(x, k)| < 1/k$
- enumerovateľná f je univerzálna, ak existuje enumerácia $f_1, f_2, \ldots$ enumerovateľných tak, že

$$f(i, x) = f_i(x) \qquad //f(i, x) = f(\langle i, x \rangle)$$

Lemma

Existuje univerzálna enumerovateľná funkcia. $f(i,x) = f_i(x)$

enumerácia $\Phi_1, \Phi_2, \ldots$ PRF $\quad \Phi(\langle x,k\rangle) = \langle p,q\rangle \iff \Phi(x,k) = p/q$

transformácia enumerácie Φ na enumeráciu f:

$$f_i(x) = \begin{cases} \sup_{k\in N}\{\Phi_i(x,k)\}, & \text{ak existuje} \\ 0, & \text{ak neexistuje} \end{cases}$$

univerzálna $\Phi_0(i,\langle x,k\rangle) = \Phi_i(x,k)$ pritom $\exists j\ \Phi_0(i,\langle x,k\rangle) = \Phi_j(\langle i,x\rangle,k)$

$$f_j(< i,x >) = f_i(x)$$

Lemma

Nech $f(x,y)$ je co-enumerovateľná. Potom

$$\forall x,y\ C(x|y) \leq f(x,y) + O(1) \Leftrightarrow d(\{x : f(x,y) \leq m\}) = O(2^m) \forall y, m$$

$\Rightarrow$ AK $\forall c\ \exists y, m :\ d(\{x : f(x,y) \leq m\}) > c2^m$
TAK $\exists x : m \geq C(x,y) \geq m + \log c$

$\Leftarrow$ g aproximuje f zhora; $m := f(x,y)$ a $A := \{x : f(x,y) \leq m\}$

popis x - znalosť g, y, m a indexu do A $\quad \rightsquigarrow C(x|y,m) \leq m + O(1)$
pre h : $C(x|y,m) = m - h$

$$C(x|y) \leq C(x|y,m) + 2logh + O(1) \leq f(x,y) - h + 2\log h + O(1)$$

- **diskrétna semimiera** $P : N \to \mathbb{R} \quad \sum_{x \in N} P(x) \leq 1$

//Ak rovnosť, pravdepodobnosť

- $\mathcal{M}$-trieda diskrétnych semimier. P_0 je univerzálna pre $\mathcal{M}$ ak
  - $P_0 \in \mathcal{M}$
  - $\forall P \in \mathcal{M}\ \exists c_P\ \forall x \in N\ c_p P_0(x) \geq P(x)$

  P_0 multiplikatívne dominuje

Theorem

Existuje univerzálna enumerovateľná diskrétna semimiera **m**

1. enumerovateľné diskrétne semimiery sa dajú vymenovávať $P_1, P_2, \ldots$
2. $P_0(x) = \sum_{n \geq 1} \alpha(n) P_n(x), \quad \sum \alpha(n) \leq 1, \quad \alpha(n) > 0$ je enumerovateľná

!!

Lemma

Trieda rekurzívnych semimier nemá univerzálny prvok.

- x_i najmenšie také, že $P_0(x_i) < 2^{-i}/i$
- $Q(x) = \begin{cases} 2^{-i}, & x = x_i \\ 0, & \text{inak} \end{cases}$

$\sum_x Q(x) = \sum_i 2^{-i} = 1$ Q je semimiera

$\forall i \; Q(x_i) = Q(x) = 2^{-i} > iP_0(x)$

Theorem

m nie je rekurzívna a $\sum_x m(x) < 1$

!!

Univerzálne pravdepodobnostné rozdelenie

- spočitateľný priestor S; $S = N \cup \{u\}$
- $P : S \to \mathbb{R}$ také, že $\sum_{x \in S} P(x) \leq 1$ //zvyšok priradíme $u \notin N$
- P je enumerovateľná, ak $\{(x, y) \mid x \in \mathbb{N}, y \in \mathbb{Q}, y \leq P(x)\}$ je RE

  P je zdola aproximovateľná
- Levin:
  - enumerovateľné rozdelenia vieme vymenúvať: $P_1, P_2, \ldots$
  - existuje univerzálne rozdelenie m

$$\forall k \in \mathbb{N}\ \exists c \in \mathbb{Q}\ \forall x \in \mathbb{N}\ [c \cdot m(x) \geq (P_k(x))]$$

$$m(x) = 2^{-K(x)}$$

$$c = 2^{K(P_k)+O(1)} = 2^{K(k)+O(1)} = O(k \log^2 k)$$

Priemerná zložitosť

$x \in N$, $l(x)$– dĺžka binárneho zápisu x, $t(x)$–čas algoritmu na vstupe x

$$T(n) = \max\{t(x) \mid l(x) = n\} \quad \text{najhorší prípad}$$

$$T^P_{av}(n|P) = \frac{\sum_{l(x)=n} P(x)t(x)}{\sum_{l(x)=n} P(x)} \quad \text{priemerný prípad}$$

- rovnomerné rozdelenie $L(x) = 2^{-2(l(x))-1}$, $L(x|l(x) = n) = 2^{-n}$
- univerzálne rozdelenie $m(x)$

Quicksort

- $T^L_{av}(n) = \Theta(n \log n)$
- $T^m_{av}(n) = \Omega(n^2)$

Theorem

Nech A je algoritmus, ktorý na každom vstupe zastane, vstupy podľa m. Potom priemerná zložitosť je rádovo rovná najhoršej.

Nech A je algoritmus, ktorý na každom vstupe zastane, vstupy podľa m. Potom priemerná zložitosť je rádovo rovná najhoršej.

- Definujme $P(x)$

  $\forall n = 1, 2, \ldots \quad a_n = \sum_{l(x)=n} m(x)$

  ak $l(x) = n$ a x je lexikograficky najmenšie také, že $t(x) = T(n)$,

  tak $P(x) = a_n$, inak $P(x) = 0$

- P je enummerovateľné
- dodefinujeme $P(u) = m(u)$
- keďže $\sum_{x \in S} P(x) = \sum_{x \in S} m(x)$, tak $\mathbf{c_P \cdot m(x) \geq P(x)}$

$$T^m_{av}(n) = \frac{\sum_{l(x)=n} \mathbf{m(x)} t(x)}{\sum_{l(x)=n} m(x)} \geq \frac{1}{c_P} \frac{\sum_{l(x)=n} P(x) t(x)}{\sum_{l(x)=n} m(x)}$$

$$= \frac{1}{c_P} \frac{\mathbf{a_n T(n)}}{\sum_{l(x)=n} m(x)} = \frac{1}{c_P} T(n) \geq \frac{1}{c_P} T^m_{av}(n) \quad \diamond$$

$$P \rightsquigarrow P_k, \text{ potom } c_P \leq k \log^2 k \qquad T^m_{av}(n) \geq \frac{T(n)}{k \log^2 k}$$

KZ s ohraničením

model TS— read-only vstup, write-only výstup, viacpáskový
$\Phi_1, \Phi_2, \ldots$ enumerácia rekurzívnych fcií T_Φ - TS, kt. počíta Φ

$C_\Phi^{t,s}(x|y) = \min\{l(p) : \Phi^{t,s}(p,y) = x\}$ ak $y = \epsilon$, potom $C_\Phi^{t,s}(x)$
$\hookrightarrow$ TS počíta v čase $t(n)$ a priestore $s(n)$

Theorem (invariance theorem)

Existuje univerzálna čiastočne rekurzívna funkcia Φ_0 taká, že $\forall$ čiastočne rekurzívnu fciu Φ existuje konštanta $c, c = c(\Phi)$, pričom $C_{\Phi_0}^{ct\log t, cs}(x|y) \leq C_\Phi^{t,s}(x|y)$

- $T_1, T_2, \ldots$; Φ_i- PRE fcia, kt. počíta T_i, $\langle y,p\rangle = 1^{l(p)}0yp$
- Φ_0 – k UTS s dvomi páskami; $U(y,\langle i,p\rangle) = T_i(y,p)$
- $T_i(y,p) = x$, $t(n)$, $s(n)$

  prechod na dve pásky $c't\log t, c's$

  ťahanie kódu, c''

  $\Big\}\ c \quad \Phi_0^{ct\log t, cs}(\langle y,\langle i,p\rangle\rangle) = \Phi_i^{t,s}(y,p)$

zložitostné triedy

- $C^{t,s}(x|y) = \min\{l(p) : \ U(\langle y, p\rangle) = x$ v čase $t(n)$ a priestore $s(n)\}$
- predikát – fcia s hodnotami $0, 1$

  $\Psi_1, \Psi_2, \ldots$ PRF enumerácia; $T_\Psi(x) \in \{0, 1\}$; $T_\Psi^{t,s}$

- $x, y, p \in \mathbb{N}$; Ψ PRF predikát; CD-zložitosť x vzhľadom k Ψ, y

  $CD_\Psi^{t,s}(x|y) = \min\{l(p) : \ \forall v \Psi^{t,s}(v, p, y) = 1$ iff $v = x\}$

- $S \subseteq \mathbb{N}$, $C^{t,s}(x|S)$, $CD^{t,s}(x|S)$ TS s orákulom S

- zložitostné triedy

$$C_\Phi[f(n), t(n), s(n)] = \{x : \ C_\Phi^{t,s}(x) \leq f(n), n = l(x)\}$$

$$CD_\Phi[f(n), t(n), s(n)] = \{x : \ CD_\Phi^{t,s}(x) \leq f(n), n = l(x)\}$$

Theorem (vzťah $C^p - CD^q, K^p - KD^q$)

Nech p, q sú polynómy, $A \in NPU$.
(i.) $\forall p\ \exists q\ CD^q(x) \leq C^p(x) + O(1)$
(ii.) $\forall p\ \exists q\ C^q(x|A) \leq CD^p(x) + O(1)$

Dôkaz

(i.) vygeneruj a porovnaj ✓

(ii.) postupne generujeme bity $x_1, x_2, ..., x_n$, $x = x_1 \dots x_n$
TS T, $T(v) = 1$ iff $x = v$ v čase $p(n)$. K nemu skonštruujeme T^A
$x_1 \dots x_i \quad \rightsquigarrow \quad x_1 \dots x_i x_{i+1}$?akceptuje T slovo s prefixom $x_1 \dots x_i 0$?

$$\{\langle T, y, 1^t, 1^n : \ T \text{ axceptuje } yz \text{ v čase } t \text{ pre } z : l(yz) = n\rangle\} \in NP$$

□

analogicky pre $K^{t,s}$, $KD^{t,s}$

majorant KZ co-enumerovateľná funkcia Φ taká, že $C(x) < \Phi(x) + c \ \ \forall x; \quad c = c(\Phi)$

Example

Ak t je totálna rekurzívna, potom C^t je totálny rekurzívny majorant C.

- UTS beží na programoch dĺžky nanajvýš $l(x) + c$, pričom spraví nanajvýš $t(l(x))$ krokov

  $C^t(x)$ — dĺžka najkratšieho, ktorý vygeneruje x $C^t(x) \geq C(x) - O(1)$
- je to totálne rekurzívne

$C(x)$, $C^t(x)$ sa môžu exponenciálne líšiť

Theorem (blow-up)

Existuje RE množina A s charakteristickou postupnosťou χ, že pre každú totálne rekurzívnu t a $\forall n$ $C^t(\chi_{1:n}|n) \geq c_t n$; $0 < c_t < 1$.

//Barzdinova lema: Ak A je rekurzívne enumerovateľná, potom $C(\chi_{1:n}|n) \leq \log n + O(1)$

- štandardné číslovanie $T_1, T_2, \ldots$ zmeníme na $M = M_1, M_2, \ldots$ tak, že
  $M_i = T_k, i = 2^{k-1} + j2^k \ \forall k \geq 1,\ j \geq 0$ $\quad$ k max.také, že 2^{k-1} delí i

1	2	1	3	1	2	1	4	1	2	1	3	1	2	1	5		
1		1		1		1		1		1		1		1			2^1
	2				2				2				2				2^2

- konštrukcia A, resp. χ diagonalizáciou

  $i = 1, \chi_1 = 0$
  for all $i > 1$ **do** dovetail nasledujúce výpočty
  $n \leftarrow 2^{i-1}$; $n' \leftarrow 2^{2^k} n$ pre $M_i = T_k$
  if $M_i(n')$ zastane s výstupom $t(n')$ **then**
  UTS simuluje $\forall$ programy $p, l(p) \leq n-1$ počas $t(n')$ krokov
  zvolíme $\chi_{n+1:2n}$ tak, že nie je na výstupe žiadneho z tých programov
  else $\chi_{n+1:2n} = 0^n$

Fakt ($\hookrightarrow$)

1. $A \subseteq \mathbb{N}, x \in A \Leftrightarrow \chi_x = 1$. *Potom A je RE.*
2. *χ je nestlačiteľná pre (wlog) monotónne rastúcu totálnu rekurzívnu funkciu t*

$c_t = 1/2^{2^k+1}$ pre index stroja k, kt. počíta t
$C^{t(n)}(\chi_{1:n}) \geq n/2^{2^k+1}$

□

$A \subseteq \mathbb{N}$, $x \in A \Leftrightarrow \chi_x = 1$. Potom A je RE.

$\forall x$ dovetail

- nájdi i : $n+1 < x \leq 2n$, $n = 2^{i-1}$
- vypočítaj $\chi_{n+1:2n}$ simuláciou $M_i(n')$

if $M_i(n')$ nezastane **then** $\chi_{n+1:2n} = 0^n, x \notin A$ & $M_i(n') = \infty$
else $x \in A$ iff $\chi_x = 1$

□

χ je nestlačiteľná pre (wlog) monotónne rastúcu totálnu rekurzívnu fciu t

- $t = T_k, M_{k_0} : k_0 = 2^{k-1}$ po prvýkrát a potom stále ďalej
  $I = \{k_0, k_0 + 2^k, k_0 + 2 \cdot 2^k, k_0 + 3 \cdot 2^k, \ldots\}$, $S = \{n |\ n = 2^{i-1}, i \in I\}$
- pre dosť veľké $n \in S$ existuje $m \in S$: $n/2^{2^k} \lll m \leq n/2$, pričom žiaden z programov dĺžky $m-1$ nedá ako výstup $\chi_{1:2m}$ v čase $t(m') = t(2^{2^k} m) \ggg t(n)$
  $C^{t(m')}(\chi_{1:2m}) \geq m$ (*)
- pre SPOR $C^{t(n)}(\chi_{1:n}|n) \leq n/2^{2^k+1}$
  ak poznáme n, v čase $t(n) \lll t(m')$ vypočítame $\chi_{1:n}$ z programu dĺžky max $m/2$
  pridáme $\log m$ bitov a vieme m, $\chi_{1:2m}$ v lineárnom čase — spor s (*)

□

Theorem

Nech t, f sú neohraničené totálne rekurzívne fcie. Potom existuje rekurzívna postupnosť χ taká, že $C^t(\chi_{1:n}|n) \geq n - f(n)$ nekonečne veľa krát.

//analogicky pre K

(wlog)$f(n) < n$, $\lim_{n\to\infty} f(n) = \infty$. χ skonštruujeme diagonalizáciou

- $g(1) = 1,\ g(n+1) = \min\{m :\ f(m) > g(n)\}$

  g total. rekurzívna, neklesajúca, neohraničená

- $\chi_1 = 0$

  for all $n = 2, 3, \ldots$ **do** vypočítaj $\chi_{g(n-1)+1:g(n)}$ takto:
  simuluj $\forall$ programy dĺžky $< g(n) - g(n-1)$ počas $t(g(n))$ krokov
  $\chi_{g(n-1)+1:g(n)}$ zvolíme tak, že $\chi_{1:g(n)}$ sa nevyskytuje ako prefix žiadneho z výstupov

$$\left.\begin{array}{l} C^t(\chi_{1:g(n)}|g(n)) \geq g(n) - g(n-1) \\ \\ f(g(n)) > g(n-1) \end{array}\right\} C^t(\chi_{1:n}|n) \geq n - f(n)$$

□

Theorem

Nech χ je rekurzívna, $t(n) = \Omega(n)$ neohraničený totálne rekurzívny čas. Potom existuje neohraničená totálna f taká, že $C^t(\chi_{1:n}|n) \leq n - f(n)$

- χ rekurzívna — T : $T(n) = \chi_n$
- krátky program q pre UTS $U(\langle q, n\rangle) = \chi_{1:n}$
  1. simuluje T so vstupom $0, 1, \ldots$ v celkovom čase n krokov; nech dopočítal $T(0), T(1), \ldots, T(m)$ $\rightsquigarrow$ U pozná $\chi_{1:m}$
  2. povieme $\chi_{m+1:n}$ $\max n - m + 1$ bitov
- $l(q) \leq n - m + 1 + c$ c je popis stroja (1-2)
- U počíta v lineárnom čase

□

Theorem (hierarchia ohraničenej KZ)

Nech f, g sú neohraničené totálne rekurzívne funkcie také, že $f(n) + g(n) \leq n$ a $\forall k$ vieme vypočítať najmenšie n také, že $f(n) = k$ v priestore $s(n) \geq \log n$ a čase $t(n) - n$. Potom pre dostatočne veľké n

$$C[f(n) + g(n), t(n), s(n)] - C[f(n), \infty, \infty] \neq \emptyset$$

Ukážeme, že $\forall UTS\ U\ \exists c = c(U), \underbrace{C[f(n) + c, t(n), s(n)]}_{A} - \underbrace{C[f(n), \infty, \infty]}_{B} \neq \emptyset$

Vezmeme $x,\ l(x) = f(n) + c/2,\ C(x) \geq l(x)$. Ukážeme, že $x0^{n-l(x)} \in A - B$

($\in A$) program $p = qx$ počíta $x0^{n-l(x)}$:

- nájde najmenšie n : $f(n) + c \geq l(p)$ — $t(n) - n, s(n)$
- vypíše $x0^{n-l(x)}$ — čas $n - l(x)$

$l(p) = l(x) + O(1) \leq f(n) + c/2 + O(1) \leq f(n) + c$

($\notin B$) **SPOROM** - nech by $x0^{n-l(x)} \in C[f(n), \infty, \infty]$. Potom rekonštrukcia x

- $x0^{n-l(x)}$ programom dĺžky $f(n)$
- oddelenie x: $l(x) = f(n) + c/2$, $f(n)$ poznáme, program dĺžky d vypočíta n

$l(programu) = f(n) + 2\log c + d + 1 < f(n) + c/2$

Theorem

$s'(n) \geq 2n + s(n) + c$, $s(n)$ neklesajúca, počitateľná v priestore $s'(n)$, $f(n) \leq n$ neohraničená neklesajúca počitateľná v priestore $s'(n) - \log n$. Pre veľké n

$$C[f(n), \infty, s'(n)] - C[n-1, \infty, s(n)] \neq \emptyset$$

Budeme $\underbrace{\text{hľadať prvý reťazec}}_{p}$, ktorý nepatrí do $C[n-1, \infty, s(n)]$

- p vypočíta najmenšie n : $f(n) > l(p)$ priestor $s'(n) = (s'(n) - \log n) + \log n$
- vyhradenie priestoru $s(n)$ v $s'(n)$ vypočítame $s(n)$
- $\forall q$ dĺžky $l(q) \leq n-1$ simuluje $UTS(q)$; x je lexikograficky najmenší dĺžky n, kt. žiaden z nich nevygeneroval //$2n + s(n)$ diagonalizácia

stačí priestor $s'(n) \geq 2n + s(n) + c$ program p má dĺžku $l(p) < f(n)$

$\in C[f(n), \infty, s'(n)]$ ale $\notin C[n-1, \infty, s(n)]$

Theorem (kompresia pasky)

K ľubovoľnej totálne rekurzívne $g(n) \geq n$ existuje totálne rekurzívna $s(n)$ taká, že $\forall f(n) < n$ a dosť veľké n

$$C[f(n), \infty, s(n)] = C[f(n), \infty, g(s(n))]$$

- $s(n)$ je minimálne i také, že $\neg\exists p,\ l(p) \leq n-1$, ktorý použije priestor s_p : $i+1 \leq s_p < g(i)$ //vieme to počítať
- $x \in C[f(n), \infty, g(s(n))] \quad \rightsquigarrow$ p vypíše x v priestore $g(s(n))$
- $f(n) < n$, p nepoužije priestor s_p : $s(n)+1 < s_p < g(s(n))$

$$\Downarrow$$

p spotrebuje priestor nanajvýš $s(n)$

□

Q — množina podmnožín množiny $\{0,1\}^*$. Množina A je Q-imúnna, ak je nekonečná a žiadna jej nekonečná podmnožina nepatrí do Q

Lemma

$f(n) < n$ je neohraničená totálna rekurzívna funkcia. Potom $\underbrace{\{0,1\}^* - C[f(n),\infty,\infty]}_{\mathbb{X}}$ je RE-imúnna.

NECH $\exists$ nekonečná RE množina akceptovaná TS T.
Uvažujme TS T'

- T' simuluje T a vymenováva prvky množiny $\mathbb{X}$;
  hľadá x aby $f(l(x)) > l(T')$
  $x \in \mathbb{X}$, $x \in C[f(n),\infty,\infty]$ a súčasne $\mathbb{X} \subseteq (C[f(n),\infty,\infty])^C$

SPOR

Theorem

Nech $\lim_{n\to\infty} s(n)/s'(n) \to 0$, *U je UTS,* $s'(n) \geq n$ *neklesajúca. Nech* $f(n)$ *je neohraničená neklesajúca počitateľná v priestore* $s(n)$. *Potom*

$$\{0,1\}^* - C[f(n), \infty, s'(n)] \text{ je } DSPACE(s(n))\text{-imúnna}$$

NECH

- nekonečná $A \in DSPACE(s(n)) \bigcap \{\{0,1\}^* - C[f(n), \infty, s'(n)]\}$
- T_A akceptuje A v priestore $O(s(n))$

Majme program p pre UTS, ktorý

- nájde najmenšie i také, že $f(i) > l(p)$ v $s(n)$
- nájde prvé $x \in A$ také, že $l(x) \geq i$ A nekonečná
  $x \in A$ a súčasne $x \in C[f(n), \infty, s'(n)]$ spor

kompresia jazykov

- $\Sigma = \{0,1\}$. $\mathbf{f} : \Sigma^* \to \Sigma^*$ je kompresia jazyka $L \subseteq \Sigma^*$ ak je one-to-one na L a až na konečne veľa $x \in L$ platí $l(f(x)) < l(x)$
- Funkcia f, $f^{-1} : f(L) \to L$ pričom $\forall x \in L\ f^{-1}(f(x)) = x$. Jazyk je komprimovateľný v čase $T(n)$, ak existuje kompresia f pre L počitateľná v čase $T(n)$ ktorej inverzná f^{-1} je tiež počitateľná v $T(n)$
- Kompresia f optimálne komprimuje L ak $\forall x \in L$ dĺžky n,

$$l(f(x)) \leq \lceil \log \sum_{i=0}^{n} d(L^{=i}) \rceil$$

- prirodzená kompresia — ranking

$$r_L : L \to N \qquad \forall x \in L : r_L(x) = \text{ index } x \text{ v } L$$

Theorem

Nech A množina. Potom existuje konštanta c a polynóm p tak, že

$$\forall x \in A^{=n} \quad CD^p(x|A^{=n}) \leq 2\log d(A^{=n}) + 2\log n + c$$

$A, d = d(A^{=n})$

Lemma

Nech $S = \{x_1, \ldots, x_d\} \subseteq \{0, \ldots, 2^n - 1\}$. Potom $\forall x_i \in S$ existuje prvočíslo $p_i \leq 2dn$ tak, že $\forall j \neq i \; x_i \neq x_j \mod p_i$. ↪

- $A = S$, $\forall x \in A$ máme p_x
- CD program p pre x

  vstup y

  if $y \notin A^{=n}$ reject y else if $y = x \mod p_x$ then accept y else reject y

  $l(p) = l(p_x) + l(x \mod p_x) + O(1)$

  $p \leq 2\log d(A^{=n}) + 2\log n + O(1)$ rovnako dlhé $p_x, x \mod p_x$

  akceptuje **len x** Thm.

Dôsledok

$A \in P$. Potom existuje polynóm p $\forall x \in A^{=n}$

$$CD^p(x|n) \leq 2\log d(A^{=n}) + 2\log n + O(1)$$

$S = \{x_1, \ldots, x_d\} \subseteq \{0, \ldots, 2^n\text{-}1\}$. $\forall x_i \in S$ existuje prvočíslo $p_i \leq 2dn$ tak, že $\forall j \neq i$ $x_i \neq x_j \mod p_i$.

- $N = 2^n$
- $\forall x_i \neq x_j \in S$ existuje nanajvýš $\log_c N = \log N / \log c$ rôznych prvočísel p takých, že $c \leq p \leq 2c$ a $x_i = x_j \mod p$
- len $d-1$ dvojíc obsahuje x_i, preto medzi $\frac{(d-1)\log N}{\log c} + 1$ prvočíslami $\exists p_i$, že $\forall j,\ i \neq j,\ x_i \neq x_j \mod p_i$
- PrimesNumberThm: $\pi(n) \backsim \frac{n}{\ln n}$,
  resp. $\pi(n) = \frac{n}{\ln n} + \frac{n}{(\ln n)^2} + \frac{2!n}{(\ln n)^3} + \frac{3!n}{(\ln n)^4} + O(\frac{n}{(\ln n)^5})$

  $\hookrightarrow$ medzi $c, 2c$ je aspoň $\frac{c}{\log c}$ prvočísel

  $\hookrightarrow$ ak $c > (d-1)\log N$ tak $p_i \leq 2d \log N = 2dn$ $\square$

Theorem

$\exists$ polynóm $p(n)$ taký, že $\forall A$ a dostatočne veľké $n \in \mathbb{N}$, ak $x \in A^{=n}$, tak

$$CD^p(x|A^{=n}, s) \leq \log d(A^{=n}) + \log\log d(A^{=n}) + O(1)$$

pričom $A^{=n}$ je dané ako orákulum, dĺžka $s \sim n \log d(A^{=n})$

$\hookrightarrow$

$h : \Sigma^n \to \Sigma^m$ lineárna transformácia daná náhodnou binárnou maticou $R = \{r_{i,j}\}$
$\forall x \in \Sigma^n \; Rx = y \in \Sigma^m : \quad y_i \equiv (\sum_j r_{i,j} x_j) \mod 2$

H..množina kódovacích funkcií.

- Nech $B, C \subseteq \Sigma^n, x \in \Sigma^n$. h separuje x v B ak $\forall y \in B, y \neq x, h(y) \neq h(x)$
- h separuje C v B ak $\forall y \in C$ h separuje y v B
- H separuje C v B ak $\forall x \in C \; \exists h \in H$ tak, že h separuje x v B.

Lemma (o kódovaní)

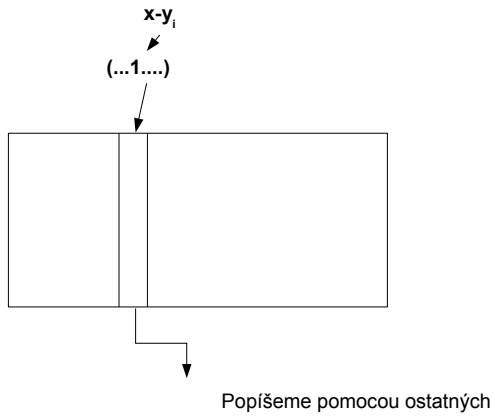
$B \subseteq \Sigma^n, d(B) = k, m = 1 + \lceil \log k \rceil$. Existuje množina H m náhodných lineárnych transformácií $\Sigma^n \to \Sigma^m$ takých, že H separuje B v B.

$B \subseteq \Sigma^n, d(B) = k, m = 1 + \lceil \log k \rceil$. Existuje množina H m náhodných lineárnych transformácií $\Sigma^n \to \Sigma^m$ takých, že H separuje B v B

fixneme náhodný reťazec s dĺžky nm^2 : $C(s|B, P, m, n) \geq l(s)$,
P program na popis s; //s reprezentuje H

H separuje B v B SPOROM:
$\exists x \in B$ že žiadne $h \in H$ neseparuje x v B:

$\exists y_1, \ldots, y_m \in B \; h_i(x) = h_i(y_i) \;\Rightarrow\; h_i(x - y_i) = 0$



krátky popis s

- index x v B $\log k$
- indexy $y_1, \ldots, y_m$ $m\lceil \log k \rceil$
- $h_1, \ldots, h_m$ bez "redundantných" stĺpcov $nm^2 - m^2$

$$C(s|B, P, m, n) \leq m^2 n - m^2 + m\lceil \log k \rceil + \lceil \log k \rceil \leq m^2 n - 1$$

◇

$\exists$ polynóm $p(n)$ taký, že $\forall A$ a dostatočne veľké $n \in \mathbb{N}$, ak $x \in A^{=n}$, tak

$$CD^p(x|A^{=n}, s) \leq \log d(A^{=n}) + \log\log d(A^{=n}) + O(1)$$

pričom $A^{=n}$ je dané ako orákulum, dĺžka $s \sim n \log d(A^{=n})$

— H z kódovacej lemy, s popisuje H; $m = 1 + \lceil \log d(A^{=n}) \rceil$

— $\forall z \in A^{=n} \exists h_i \in H$ separujúce z v $A^{=n}$

— program, kt. z $ih_i(z)$ akceptuje z $\qquad$ $A^{=n}, s$ poznáme

- over, či $x \in A^{=n}$
- ak $x \in A^{=n}$, $ih_i(z) \rightsquigarrow i, h_i(z)$, nájde h_i
- vypočíta $h_i(x)$
- akceptuje $z \quad \Leftrightarrow \quad h_i(x) = h_i(z)$

$$CD^p(z|A^{=n}, s) \leq m + \log m + O(1) = \log d(A^{=n}) + \log\log d(A^{=n}) + O(1)$$

Theorem

Existuje polynóm p, že $\forall A$ a dostatočne veľké n ak $x \in A^{=n}$, tak

1. $C^p(x|A^{=n}, s, NP^A) \leq \log d(A^{=n}) + \log\log d(A^{=n}) + O(1)$
2. $C^p(x|A^{=n}, \Sigma_2^{P,A}) \leq \log d(A^{=n}) + \log\log d(A^{=n}) + O(1)$

keď $A^{=n}$, NP^A-úplná mn., $\Sigma_2^{P,A}$-úplná mn. sú orákulá, s je reťazec dĺžky $\sim n \log d(A^{=n})$

$$\Sigma_2^p = NP^{NP} \begin{cases} \Sigma_1^P = & NP \\ \Sigma_{i+1}^P = & NP^{\Sigma_i^P} \end{cases}$$

(1.) $C^p(x|A^{=n}, s, NP^A) \leq$ $CD^q(x) + O(1)$
$CD^p(x|A^{=n}, s) \leq \log d(A^{=n}) + \log\log d(A^{=n}) + O(1)$

(2.) pomocou orákula **generujeme s**

je $s_1 \ldots s_i 1$ prefix H kt. separuje $A^{=n}$ v $A^{=n}$?

□

riedka množina A — $d(A^{=n}) \leq n^c + c$

Theorem

Existuje riedky jazyk L taký, že ak L skomprimujeme pomocou pravdepodobnostného polynomiálneho stroja s orákulom pre L, tak kompresná funkcia mapuje reťazce dĺžky n na reťazce dĺžky $n - O(\log n)$.

- L obsahuje jediné nestlačiteľné slovo dĺžky $2^{2^{\cdots^2}}$, $\mathbf{I(x) \leq C(x)}$
- T pravdepodobnostný orákulovský stroj, kt. rozpoznáva L v čase n^k
- kvôli SPORu: existuje kompresia f : $l(f(x)) \leq n - c_1 - (k + c_1)\log n$. Ukážeme, že T nemôže zrekonštruovať x z $f(x)$ s pravdep. aspoň $1/2$

  Ak áno:

  – R je množina rozhodnutí korektnej odpovede — $1/2$ z 2^{n^k}

  – $r \in R$ s $\mathbf{C(r|x) \geq I(r) - 1}$ — existuje!

  – $\mathbf{C(x) \geq I(x)}$, $C(r) \leq l(r) + O(1)$, symetria KZ — $\mathbf{C(x|r) \geq n - c_2 \log n}$

krátky popis $x|r$

informácia potrebná pre odpovede orákula keď T počíta pri vstupe $f(x)$

$$(k+2)\log n + c_1$$

simulácia T so vstupom $f(x)$ pri znalosti výsledkov hádzania r

- diskusia
- $2\log n$ bitov pre slová dĺžky $< n$
- T sa pýta $?x \in L?$ - $k\log n$ bitov na určenie času, keď sa to pýta

$$C(x|r) \leq c_1 + 2\log n + k\log n + l(f(x))$$

pre $c_1 > c_2 + 2$ dostaneme spor $//l(f(x)) \leq n - c_1 - (k + c_1)\log n$

□

Theorem

Ak $L \in 1DLOG$, potom r_L môžme vypočítať v P.

- T rozpoznáva L
- $L_x = \{y \in L : y \leq x\}$; zrejme $r_L(x) = d(L_x)$
- x do vnútornej pamäte a rozpoznávame L_x jednosmerne v DLOG s polynomiálnym počtom stavov — $\mathbf{T_x}$
- $G_x = (V_x, E_x)$ — graf konfigurácií pre T_x;
  $V_x = \{[hlava, stav, paska]\}$
  E_x— zachytáva možné prechody

$\hookrightarrow$ $r_L(x)$ je počet akceptujúcich ciest v G_x

honest funkcia$f$ — $f : \{0,1\}^* \to \{0,1\}^*$ ak $\exists k \in \mathbb{N}$, že $\forall x \in \{0,1\}^*$

$$l(f(x)) \leq l(x)^k + k, \; l(x) \leq l(f(x))^k + k$$

Lemma

Nech f je honest, počitateľná strojom T v polytime. $\forall t \geq 1$ a takmer $\forall n$ (až na konečne veľa)

$$f(C[\log\log n, n^t, \infty]) \subseteq C[\log n, n^{\log\log n}, \infty]$$

$\forall x \in C[\log\log n, n^t, \infty]$ $\exists y$ dĺžky $l(y) \leq \log\log n$, že $T'(y) = x$ v polytime

$f(x)$ počítame

- na základe y T T' $\qquad C(f(x)) \leq l(y) + O(1) \leq \log n$
- čas
  - $y \rightsquigarrow x$ v n^t
  - $x \rightsquigarrow f(x)$ v polytime

  $\hookrightarrow$

  $n^{t_1} + n^{t_2} = O(n^{\log\log n})$

□

Theorem

Existuje rekurzívne orákulum A také, že $P^A \neq NP^A$ *//Baker-Gill-Solovay*

- $f(1) = 2$
  $f(k) = 2^{f(k-1)}$
- zoberme $B \subseteq \{1^{f(k)} : k \geq 1\}$ a $B \in DTIME[n^{\log n}] - P$

  $B \in NP^A, B \notin P^A$
- skonštruujeme A
  — $\forall k : \ 1^{f(k)} \in B$

  $A \leftarrow$ prvý reťazec dĺžky $f(k) \in C[\log n, n^{\log n}, \infty] - C[\log n, n^{\log \log n}, \infty]$

$\hookrightarrow$ $T^?$ sa pýta na slová mimo A alebo nie dlhšie ako $\log n$

$\hookrightarrow$ membership pre $l(y) \leq \log n$ v čase $O(\log n)^{\log \log n} = O(n^k)$

$B \in P^A \Rightarrow B \in P \Rightarrow$ spor s $B \notin P$

□

Množiny sú P-izomorfné, ak existuje polynomiálne počitateľná bijekcia medzi nimi

Bermann -Hartmanis predpoklad: $\forall$ NPÚ sú P-izomorfné

Theorem

Ak existuje $L \in P$ *taký, že* $L \subset SAT$ *a*

$$C[\log n, n^{\log n}, \infty] \cap SAT \subseteq L$$

Potom $SAT - L \in NPÚ$, *ale NIE JE P-izomorfný so SAT*

- $SAT - L \in NPU \rightsquigarrow$ redukciou $SAT \rightsquigarrow SAT - L$
  - $\forall x \in L \quad x \rightsquigarrow \omega \in SAT - L$
  - $\forall x \notin L \quad x \rightsquigarrow x$
- keby existoval P-izomorizmus h medzi $SAT, SAT - L$, bol by honest fcia v P
  $\hookrightarrow \quad h(SAT \cap C[\log\log n, n^3, \infty]) \subseteq C[\log n, n^{\log\log n}, \infty]$
- keďže $SAT \cap C[\log\log n, n^3, \infty] \neq \emptyset$, $h(SAT \cap C[\log\log n, n^3, \infty]) \notin SAT - L$

□

množina A je exponentially low ak $E^A = E,\ E = \cup_{C \in \mathbb{N}} DTIME(2^{cn})$

Theorem

Existuje exponentially low riedka A, ktorá nie je v P.

Nech $B = C[n/2, 2^{3n}, \infty], \overline{B} = B^C$

$A = \{x \mid x \text{ je lexikograficky prvé z } \overline{B} \text{ také, že } l(x) = 2^{\left.2^{\cdots^2}\right\}^m}, m > 0\}$

- $A \in E$
- $A \notin P$ SPOROM
  ak $T,\ L(T) = A \subseteq \overline{B}$ v polytime
  – $\log n$, popis T,
  hľadáme $x \in \{0,1\}^n, n > 2(l(T) + \log n)$, aby $x \in A$ $\hookrightarrow$ spor, $x \in B$
- $E^A = E$ nech T^A počíta v 2^{cn}

Fakt

Pre $c' > 3c + 3$ T^A sa nemôže pýtať na "$y \overset{?}{\in} A$ ak $l(y) \geq c'n, y \in A$" $\hookrightarrow$

A vieme simulovať BEZ orákula — dotaz na slovo

- dlhšie ako $c'n$ – NIE
- kratšia ako $c'n$ – prehľadávaním zistíme odpoveď $\square$

Pre $c' > 3c + 3$ T^A sa nemôže pýtať na "$y \stackrel{?}{\in} A$" ak $l(y) \geq c'n,\ y \in A$

SPORom — Nech y je prvý v $A : l(y) > c'n$, na ktorý sa T^A pýta //pri vstupe $x, l(x) = n$

že $y \in B$: $//B = C[n/2, 2^{3n}, \infty], \overline{B} = B^C$

- zápis $A^{<c'n}$ v $2 \log c'n$
- rekonštrukcia y, na ktoré sa pýtame v čase $t < 2^{cn}$, simuláciou $T^A(x)$ po čas t

- diskusia + popis T^A	$O(1)$
- x	n
- popis $A^{<l(y)} = A^{<c'n}$	$2 \log c'n$
- popis t	cn

$$c' \frac{n}{2} > (c+1)n + 2 \log c'n + O(1)$$

- čas $\quad 2^{cn} < 2^{c'n} \leq 2^{l(y)} \quad y \in B = C[l(y)/2, 2^{3l(y)}, \infty]$

$y \in B \Rightarrow\ y \notin \overline{B}$, pričom $A \subseteq \overline{B}$ ◇

L je **P-printable** ak $\exists k \in \mathbb{N}$ také, že $\forall$ prvky z $L^{\leq n}$ vypíšeme v čase $\mathbf{n^k + k}$

Theorem

$L \subseteq \{0,1\}^*$. *Nasledujúce podmienky sú ekvivalentné*

1. *L je P-printable*
2. *L je riedka a má P-time vypočitateľný ranking* //$r_L(x) = d(\{y; y \leq x\})$
3. *L je P-izomorfný s nejakou tally množinou* $\in P$ //tally $\subseteq \{a\}^*$
4. $L \in C[k \log n, n^k, \infty]$ *a* $L \in P$

$1 \Rightarrow 2$ ✓

$2 \Rightarrow 3$

- L polytime ranking r_1, $d(L^{\leq n}) \leq p(n)$
- $\mathbf{r_2(x) = 1x - r_1(x)}$ je ranking pre L^c
- $T = \{0^{np(n)+i} : r_1(1^{n-1}) < i \leq r_1(1^n))\}$ je tally v P
- $\mathbf{r_3}$ ranking pre $\{0\}^* - T$
- p mapuje x na $\begin{cases} 0^{np(n)+\mathbf{r_1}(x)}, & x \in L \\ r_3^{-1}(r_2(x)) & x \notin L \end{cases}$

  p je P-izomorfizmus
  – v P-time
  – $x \in L \rightsquigarrow$ jednoznačná hodnota v T; $0^{np(n)+\mathbf{r_1}(x)} \in T \rightsquigarrow$ jednoznačne $x \in L$

$3 \Rightarrow 4$ $//L$ je P-izomorfný s nejakou tally množinou $\in P \Rightarrow$ $L \in C[k \log n, n^k, \infty]$

- P izomorfizmus f, f a f^{-1} počitateľné v n^c, $L \leftrightarrow$tally T, $T \in P$ $\rightsquigarrow L \in P$
- f počitateľná v n^c, $l(f(x)) \leq n^c, n = l(x)$
  - $\hookrightarrow$ binárna reprezentácia $f(x)$ $c \log n$
  - $\hookrightarrow$ namiesto x povieme $f(x)$ $C[k \log n, n^k, \infty]$

$4 \Rightarrow 1$ $//L \in P$, $L \in C[k \log n, n^k, \infty]$ potom L je P-printable

pre vstup $x, l(x) = n$

- simuluj n^k krokov $\forall$ programu dĺžky $k \log n$
- ak stihol vypísať y, over či $y \in L, l(y) = n$; ak áno, vypíš y $//y = x$
- P-time ✓

□

čiastočný program p — $0,1,\perp$;
charakteristická funkcia množiny A $\quad A(x) = 1 \Leftrightarrow x \in A$;
funkcia p je konzistentná s A ak $p(x) \neq \perp \Rightarrow p(x) = A(x)$

$time_T(p, y)$– čas stroja T, keď podľa programu p spracováva vstup y

Def.

(t-ohraničená) instance zložitosť reťazca x vzhľadom k T, A

$$ic^t_T(x : A) = \min\{l(p) : \; T(p, x) \neq \perp \; a \; \forall y T(p, y) \neq \perp \Rightarrow time^t_T(p, y) \leq t(l(y)), T(p, y) = A(y)\};$$
resp. ∞, ak p neexistuje

Fakt (invariance)

Existuje univerzálny TS U, že $\forall T \; \exists c \; \forall A, t, x$
$ic^{t'}_U(x : A) \leq ic^t_T(x : A) + c$ *kde* $t'(n) = ct(n) \log t(n) + c$

Example

- $CD^t(x) = ic^t(x : \{x\})$
- $t'(n) = ct(n) \log(n) + c$
  $ic^{t'}(x : A) \leq C^t(x) + c$
  $ic^{t'}(x : A) \leq CD^t(x) + c$

domnienka Nech $t(n)$ je vypočitateľná v $t'(n) = ct(n) + c$, A je rekurzívna množina. Ak $A \notin DTIME(t(n))$, potom $\exists c$ a nekonečne veľa x, že

$$ic^t(x : A) \geq C^{t'}(x) - c$$

Lemma

$A \in P \Leftrightarrow \exists$ polynóm t a konštanta c, $\forall x\ ic^t(x : A) \leq c$

- $\Rightarrow$ $A \in P \Rightarrow ic^t(x : A) \leq c$ pre polynóm t ✓
- $\Leftarrow$ B=$\{p \mid l(p) \leq c$, p je konzistentný s A, čas nanajvýš $t(n)\}$

  $x \overset{?}{\in} A$ simuláciou $\forall p \in B$ □

Def. (polynomiálne jadro)

Nech A je rekurzívna. Nekonečná množina C je polynomiálne jadro A ak každý totálny program p, ktorý rozhoduje A, a polynóm t, $time_p(x) > t(l(x))$ $\forall x \in C$ až na konečne veľa.

//C nemusí byť podmnožina A!

Lemma

C je polynomiálne jadro $A \Leftrightarrow \forall$ polynóm t a konštantu c

$$ic^t(x : A) > c \text{ až na konečne veľa } x \in C$$

$\Rightarrow$

- nech by nekonečne veľa $x \in C : ic^t(x : A) \leq c$; t polynóm
- B- programy p konzistentné s A, $l(p) \leq c, time_p() \leq t(n)$

  $\hookrightarrow$ nekonečne veľa prvkov v C identifikovaných simulovaním konečnej množiny programov SPOR

$\Leftarrow$

- ak C NIE JE poly.jadro A, tak $\exists$ program

  p, polynóm t, že pre nekonečne veľa $x \in C$ $time_p(x) \leq t(l(x))$ // p konzistentný s A
- modifikácia p – po $t(n)$ krokoch HALT

$$ic^{t'}(x : A) \leq c \text{ pre nekonečne veľa } x \in C$$

□

$IC[\log, poly]$ množiny A, pre kt. $ic^t(x : A) \leq c \log l(x) + c$, kde t je polynóm
SAT $\Phi(x_1, \ldots, x_k)$ v KNF, $l(\Phi(x_1, \ldots, x_k)) = n$; k, n polynomiálne ekv.

Theorem

$SAT \in IC[\log, poly] \Rightarrow NP = P$

- $SAT \in IC[\log, poly] \Rightarrow ic^t(\Phi : SAT) \leq c \log l(\Phi) + c$
- pre Φ v KNF existuje p_Φ : $l(p_\Phi) \leq c \log(l(\Phi)) + c$, že korekte rozhodne $\Phi \stackrel{?}{\in} SAT$ a konzistentný so SAT
- $\Phi(x_1, \ldots, x_k)$, $l(\Phi) = n$
- $A = \{p : l(p) \leq c \log n + c\}$ $\quad$ $d(A) \leq q(n)$

Nech p_0 korektne rozhoduje $\Phi \stackrel{?}{\in} SAT$ potom procedúra SAT v polytime

– alebo korektne rozhodne $\Phi \stackrel{?}{\in} SAT$
– alebo z A vyhodí $p \neq p_0$ nekonzistentný so SAT

procedúra SAT

krok 1 simuluj všetky $p \in A$ nanajvýš $t(n)$ krokov so vstupom $\Phi(x_1, \ldots, x_k)$
ak žiaden program nezamieta, tak return(akceptuj)
ak žiaden program neakceptuje, tak return(zamietaj)

krok 2 //existujú aj akceptujúce aj zamietajúce programy
pre $i = 1, 2, \ldots, k$ postupne:
nech už máme $b_1, \ldots b_{i-1}$, $p \in A$ akceptuje $\Phi = \Phi(b_1, \ldots, b_{i-1}, x_i, \ldots, x_k)$

- simuluj $q \in A$ so vstupmi
  $\Phi_0 = \Phi(b_1, \ldots, b_{i-1}, 0, x_{i+1}, \ldots, x_k)$ a
  $\Phi_1 = \Phi(b_1, \ldots, b_{i-1}, 1, x_{i+1}, \ldots, x_k)$
- ak žiaden z programov neakceptuje, tak
  – $\Phi(b_1, \ldots, b_{i-1}, x_i, \ldots . x_k) \notin SAT$
  – $p \neq p_0$ nie je konzistentný so SAT $A \leftarrow A - p$
- ak jeden zo vstupov JE akceptovaný, tak $b_i \leftarrow x_i$

koniec po kroku 1 ✓
koniec s $b_1, \ldots, b_k$:
if $\Phi(b_1, \ldots, b_k) = 1$ **then** return($\Phi \in SAT$)
else $A \leftarrow A - p$

□

$age(x) = \min_p\{2^{l(p)}t : U(p) = x \text{ v priebehu t krokov}\}$

Def. (Kt zložitosť)

UTS monotonne skenuje začiatok p kým vypíše x, $t(p, x)$je počet krokov, kým vypísal x

$$Kt_U(x) = \min_p\{l(p) + \log t(p, x)\}$$

$T_1, T_2, \ldots$ enumerácia prefixných TS, $\Phi_1, \Phi_2, \ldots$ príslušné čiastočne rekurzívne funkcie.

Φ čiastočne rekurzívna a $\Phi(y) = x$ potom y je svedok pre x

// $h(\Phi)$ - obor hodnôt Φ

Def.

Algoritmus A invertuje problém Φ ak pre dané $x \in h(\Phi)$ vypočíta Φ svedka y pre x a overí, že $\Phi(y) = x$. Pre $x \notin h(\Phi)$ diverguje.

Example

- splniteľná formula – chceme spĺňajúce priradenie
- k-zafarbiteľný graf – chceme to zafarbenie
- ...

Lemma

Ak existuje algoritmus A, ktorý invertuje Φ v čase $t(n)$, potom existuje algoritmus, ktorý invertuje Φ v čase $c_A t(n)$.

SIMPLE simuluje

- T_1 v každom druhom
- T_2 v každom druhom z toho, čo ostalo
- T_3 v každom druhom z toho, čo ostalo
- ...

1	⊔	1	⊔	1	⊔	1	⊔	1	⊔	1	⊔	1	⊔	1	⊔	1
1	2	1	⊔	1	2	1	⊔	1	2	1	⊔	1	2	1	⊔	1
1	2	1	3	1	2	1	⊔	1	2	1	3	1	2	1	⊔	1
1	2	1	3	1	2	1	4	1	2	1	3	1	2	1	⊔	1

T_k invertuje Φ v čase t, potom SIMPLE to spraví v

$$2^k t + 2^{k-1}$$
$$c_k = 2^{k+1}$$

Theorem

Ak existuje algoritmus A, ktorý invertuje Φ v čase $t(n)$, potom metóda SEARCH invertuje Φ v čase $c_A t(n)$.

$Kt'(w|x,\Phi) = \min\{l(p) + \log t(p,x)\}$ – pre dané x program p vypíše w a otestuje, či $\Phi(w) = x$ v čase $t(p,x)$

SEARCH — univerzálny prefixový U beží postupne na programoch p dĺžky $< i$ počas $2^i 2^{-l(p)}$ krokov a hľadá $\Phi(w) = x$

Fakt

$\forall w : Kt'(w|x,\Phi) \leq k$ *sú testované v čase* 2^{k+1} $\hookrightarrow$

$\hookrightarrow$ $m = \min\{Kt'(w|x,\Phi) : w$ je Φ-svedok pre $x,\ l(x) = n\}$.

$\hookrightarrow$ prefixný T invertuje Φ v čase $t(n)$ $\quad m \leq Kt'(T|x,\Phi)$

$\hookrightarrow$ SEARCH invertuje počas 2^{m+1} krokov, $Kt'(T|x,\Phi) \leq K(T) + \log t(n)$

$\hookrightarrow 2^{K(T)+1} t(n)$ krokov

$$c = 2^{K(T)+1}$$

$\forall w : Kt'(w) \leq k$ sú testované v čase 2^{k+1}

- $Kt'(w|x,p) \leq i \hookrightarrow l(p) + \log t(p,x) \leq i \hookrightarrow t(p,x) \leq 2^{i-l(p)}$
- Kraft $\sum 2^{-l(p)} \leq 1$

$$\hookrightarrow \sum_{1 \leq i \leq k} \sum_{0 < i-l(p)} 2^{i-l(p)} \leq \sum_{U(p)<\infty} 2^{-l(p)} \sum_{1 \leq i \leq k} 2^i \leq 2^{k+1}$$

□

Example

T_k invertuje v čase $t(n)$

SIMPLE $2^{k+1}t(n)$

SEARCH $2^{K(T)+O(1)}t(n) = O(k(\log k)^2)t(n)$

ak $K(T_k) = \log\log k$, čas pre SEARCH je len $O((\log k)t(n))$

SEARCH ? SIMPLE

$t(n)$ časovo konštruovateľná v $O(t(n))$

- $P^*(x)$= súčet pravdepodobností prvkov $\leq x$; počitateľná v $t(n)$ ak $\exists T : \forall x, k$ vypočíta y v $t(l(x)+k)$: $|P^*(x) - y| \leq 1/2^k$
- $P(x) = P^*(x) - P^*(x-1)$
- $K^t(x) = \min\{l(p) : \ U(p) = x$ v t(n) krokoch $\}$
- $m^t(x) = 2^{-K^t(x)}$, $m^{*t}(y) = \sum_{y \leq x} m^t(x)$

Theorem

Rozdelenie $m^{t'}$, kde $t'(n) = O(nt(n)\log(nt(n)))$, je univerzálne pre pravdepodobnostné mass funkcie P v čase t počítateľného rozdelenia P^*; $m^{t'}$ multiplikatívne dominuje P v nasledujúcom zmysle:
$\exists c_p : c_p m^{t'}(x) \geq P(x)$, pričom $\log c_p = K^{t'}(P^*) + O(1)$ závisí od P^*,ale nie t, x.

Tvrdenie Ak P^* je počitateľné v $t(n)$, tak existuje c_P

$$K^{t'}(x) \leq \log \frac{1}{P(x)} + \log c_P$$

- $[0,)$ rozdelíme na podintervaly tak, že kód $c(x)$ zaberá $I_x = [P^*(x-1), P^*(x))$
- binárny interval určený reťazcom r je $\Gamma_r = 0.r, 0.r + 2^{-l(r)}$
- ak Γ_r je najväčší binárny interval obsiahnutý v I_x, tak $c(x) \leftarrow r$
- $|\Gamma_r| \geq \frac{|I_x|}{4} \Longrightarrow l(c(x)) \leq \log 1/P(x) + 2$

dekódovanie

$k \leftarrow 1$

doubling

repeat $k \leftarrow 2k$ until $\Gamma_{c(x)} \subseteq [P^*(k-1), P^*(k))$

$l \leftarrow k/2;\ u \leftarrow k$

binarySearch

$m \leftarrow (u+l)/2$

if $\Gamma_{c(x)} \subseteq [P^*(m-1), P^*(m))$

then $x \leftarrow m$

else $\begin{cases} u \leftarrow m, & \Gamma_{c(x)} \text{ naľavo od } P^*(x-1); \\ l \leftarrow m, & \Gamma_{c(x)} \text{ napravo od } P^*(x) \end{cases}$

$\sum_{i=0}^{n} O(t(i)) = O(nt(n))$

rekonštrukcia x

- diskusia $O(1)$
- program q na výpočet $P^*(x)$ v $t(n)$
- $c(x)$

čas

$O(nt(n))$ //univerzálny TS v $t'(n) = O(nt(n)\log(nt(n)))$

KZ

- $K^{t'}(x) \leq l(c(x)) + \underbrace{l(q) + O(1)}_{\log c_P}$
- $K^{t'}(x) \leq \log 1/P(x) + \log c_P$
- q v $t(n)$, q' v $t'(n)$ $\quad l(q') \leq l(q)$, $\mathbf{l(q') = K^{t'}(P^*)}$
- $\log c_P = l(q') + O(1) = K^{t'}(P^*) + O(1)$ □

Logická hĺbka – počet krokov na ceste od pôvodu k objektu $\leftrightsquigarrow$ čas algoritmu na generovanie objektu z kratšieho popisu

x^* najkratší samoodeľujúci popis

1. počet krokov na výpočet x z x^* nie je stabilné

   pár bitov navyše, podstatné zníženie času

2. relaxujeme na minimum $\rightsquigarrow$ skoro minimálny program
   x má hĺbku d s toleranciou 2^{-b} ak x vypočítame v d krokoch z p, $|p| \leq |x^*| + b$

   $\hookrightarrow \quad \frac{2^{-l(p)}}{2^{-K(x)}} \geq 2^{-b}$

   stabilné, nevyhovujúce

$Q_U(x) = \sum_{U(p)=x} 2^{-l(p)}$

$$\log \frac{1}{Q_U(x)} = \log \frac{1}{m(x)} = K(x)$$

hĺbka reťazca x na hladine významnosti $\varepsilon = 2^{-b}$ je

$$depth_\varepsilon(x) = \min \left\{ d : \frac{Q_U^d(x)}{Q_U(x)} \geq \varepsilon \right\}$$

$Q_U^d(x) = \sum_{U^d(p)=x} 2^{-l(p)}$ $\quad$ $U_d(p) = x$ U po nanajvýš d krokoch zastane.

x je (d,b)-deep, ak $d = depth_\varepsilon(x)$, $\varepsilon = 2^{-b}$

b-nestlačiteľný reťazec: $K(x) > l(x) - b$

Theorem

Reťazec x je (d,b)-deep(b s presnosťou $K(d) + O(1)$) $\iff$ d je minimálny počet krokov potrebný na generovanie x b-nestlačiteľným programom

↪

chceme $\frac{1}{2^{b+K(d)+O(1)}} \underbrace{\leq}_{\Longleftarrow} \frac{Q_U^d(x)}{Q_U(x)} \underbrace{\leq}_{\Longrightarrow} \frac{1}{2^{b-O(1)}}$

d - čas potrebný na generovanie x b-nestlačiteľným programom

//menej ako d krokov, b-stlačiteľný program p

- $\forall p\ \exists p' : U(p') = p,\ l(p') \leq l(p) - b$
- q: $U(p') \rightsquigarrow p, U(p) \rightsquigarrow x$
- $l(q) \approx l(p) - b + O(1)$
- $\alpha = Q_U(x) - \sum_{U(q)=x} 2^{-l(q)} \geq 0$

$$\Longrightarrow \quad \frac{Q_U^d(x)}{Q_U(x)} = \frac{\sum_{U^d(p)2^{-l(p)}}}{\alpha + \sum_{U(q)=x} 2^{-l(q)}}$$

$$\leq \frac{\sum_{U^d(p)2^{-l(p)}}}{\sum_{U(q)=x} 2^{-l(q)}}$$

$$\leq \frac{\sum_{U^d(p)2^{-l(p)}}}{\sum_{U(q)=x} 2^{-(l(p)-b+O(1))}} \leq \frac{1}{2^{b-O(1)}}$$

$\Longleftarrow$ SPORom

$// \frac{1}{2^{b+K(d)+O(1)}} > \frac{Q_U^d(x)}{Q_U(x)}$

- x^*, d^*
  $\rightsquigarrow$ vymenovávame A = samoodd. programy generujúce x v čase max d
  $l(q) = K(x) + K(d) + O(1)$

- $\sum_{p \in A} 2^{-l(p)} = Q_U^d(x)$

- platí $Q_U(x) = 2^{-K(x)+O(1)}$

(*) $\sum_{p \in A} 2^{-l(p)} = Q_U^d(x) < \frac{Q_U(x)}{2^{b+K(d)+O(1)}} = \frac{2^{-K(x)+O(1)}}{2^{b+K(d)+O(1)}} = 2^{-K(x)-b-K(d)-O(1)}$

(**) B prefix-free s $\sum_{x \in B} 2^{-l(x)} < 2^{-m}$, enumerovateľná programom s. Potom B komprimovaná o $m - l(s) - O(1)$ bitov $\hookrightarrow$

- $B = A, s = q, m = K(x) + b + K(d) + O(1)$, (*) , (**)
  $\forall p \in A$ môže byť komprimovaný o $K(x) + b + K(d) + O(1) - l(q) - O(1) > b$ bitov
  SPOR s x je (d, b)-deep

B prefix-free s $\sum_{x \in B} 2^{-l(x)<2^{-m}}$, enumerovateľná programom s. B komprimovaná o $m - l(s) - O(1)$ bitov

$\sum_{x \in B} 2^{-l(x)} 2^m < 1 \quad \rightsquigarrow$ Shannon-Fano: $|c(x)| \leq l(x) - m + 2 + \quad l(s) + O(1)$

$\diamond$

reťazec x je d-shallow ak nie je $(d+1, b)$-deep. $n + O(1)$ shallow je shallow.

- náhodný reťazec je shallow
- 1^n je shallow

Theorem

Nech $R = \{x : C(x) \geq l(x) - \log^2 l(x)\}$. *Potom* $BPP^R = P^R$

konštrukcia x ∈ R, l(x)=m //induktívne

$x = \epsilon$

repeat $\forall y, l(y) = \log m$

if $xy \in R$ then $x \leftarrow xy$

until $l(x) \geq m$‘

y existuje

$x \in R, C(y|x) \geq l(y)$, potom symetria informácie:

$$\begin{aligned} C(xy) &\geq C(x) + C(y|x) - \text{clog}(C(xy)) \\ &\geq C(x) + C(y|x) - \text{clog } l(xy) \\ &\geq l(x) - \log^2 l(x) + l(y) - \text{clog } l(xy) \\ &\geq l(xy) - \log^2 l(xy) \end{aligned}$$

n^2 krát BPP výpočet dĺžky n + väčšina

$\rightsquigarrow$ pravdepodobnosť chyby $\leq 2e^{-O(n^2)}$

$\rightsquigarrow \leq 2^M 2^{-O(n^2)}$ reťazcov s chybou

$n^2 \leq M \leq n^3$ je počet "chybných" postupností

$\rightsquigarrow$ KZ chybných postupnost$\leq M$-$O(n^2)$

Lemma

Na rozpoznanie $\mathbf{xx^R}$ *treba rádovo* $\mathbf{n^2}$ *krokov*

- x dĺžky n tak, že $\mathbf{C(x|T,n) \geq n}$
- $\mathbf{l(T)}$, $\mathbf{l(pp)}$ dĺžka popisu stroja a prechodovej postupnosti

- výpočet na $x0^{2n}x^R$
  ak $\forall l(pp(n+1)), \ldots, l(pp(2n)) \geq \frac{n}{2l(T)}$ ✓
  ak $\exists n < i_0 \leq 2n : \mathbf{l(pp(i_0)) < \frac{n}{2l(T)}}$, spravíme **kratší popis** x**:**
  $\forall y \in \{0,1\}^n$ skúšame $\mathbf{y0^{2n}}$
  $C(x|T,n) \leq l(pp(i_0)) + O(1) \leq n/2 + O(1)$ **? pozícia ?**

Náhodný graf s vysokou pravdepodobnosťou nemá izolované vrcholy

štandardne

konkrétny vrchol je izolovaný s pravdepodobnosťou $\frac{1}{2^{n-1}}$
nejaký vrchol je izolovaný s pravdep. nanajvýš $\frac{n}{2^{n-1}}$
graf nemá izolovaný vrchol s pravdep. $\geq 1 - \frac{n}{2^{n-1}}$

nestlačiteľnosťou

$\chi_1 \dots \chi_e$, $e = \binom{n}{2}$ -charakteristická postupnosť množiny hrán $\sim$ grafu

graf G s $C(G|n) \geq \binom{n}{2} - \delta(n)$

graf s izolovaným vrcholom i - nie susedia i, áno identifikácia i

$\binom{n}{2} - \delta(n) \leq C(G|n) \leq \binom{n}{2} - (n-1) + \log n \quad \rightsquigarrow \boxed{\delta(n) \geq n - \log n - 1}$

nanajvýš $2^{-(n-\log n-1)}$-tina má izolovaný vrchol, preto s pravdep. aspoň $1 - \frac{n}{2^{n-1}}$
náhodný graf nemá izolovaný vrchol

Example

$n, r, s \in N$, $2 \log n \leq r, s \leq n/4$, s párne; potom $\forall n$ existuje $n \times n$ matica nad $GF(2)$ taká, že $\forall$ podmatica rozmeru s riadkov a $n - r$ stĺpcov má hodnosť $\geq s/2$.

náhodné x dĺžky n^2 tvorí štvorcovú maticu R; $\mathbf{C(x) \geq n^2}$

ak existuje podmatica s iba $s/2 - 1$ lineárne nezávislými riadkami, tak skrátenie popisu x

- prvky odpovedajúce podmatici
+ charakteristická postupnosť s riadkov vzhľadom k nezávislým
+ zoznam nezávislých riadkov
+ závislé určené koeficientami nezávislých
+ samoodeľujúco n, r, s
+ identifikácia podmatice

ak existuje podmatica s iba $s/2-1$ lineárne nezávislými riadkami, tak skrátenie popisu x

-	prvky odpovedajúce podmatici	$(\mathbf{n}-\mathbf{r})\mathbf{s}$
+	charakteristická postupnosť s riadkov vzhľadom k nezávislým	$\mathbf{s}$
+	zoznam nezávislých riadkov	$(\mathbf{s/2-1})(\mathbf{n}-\mathbf{r})$
+	závislé určené koeficientami nezávislých	$(\mathbf{s/2-1})\times(\mathbf{s/2+1})$
+	samoodeľujúco n, r, s	$\mathbf{3}\log\mathbf{n}+\mathbf{6}\log\log\mathbf{n}+\mathbf{O(1)}\leq\mathbf{4}\log\mathbf{n}$
+	identifikácia podmatice	$\mathbf{2n}$

$$n^2-(n-r)s+s+(s/2-1)(n-r)+(s/2-1)\cdot(s/2+1)+4\log n+2n$$

$2\log n \leq s, r \leq n/4$ -pre veľké n to klesne pod n^2

vlastnosť s veľkou pravdepodobnosťou $\leftrightsquigarrow$ vlastnosť objektov s malou deficiency

1. všetky objekty zahŕňajú aj objekty s malou (Kolmogorovskou) zložitosťou
2. objekt s malou deficiency je prvkom množiny, na ktorej je založená platnosť vlastnosti s veľkou pravdepodobnosťou
3. Ak P, Q platia s pravdepodobnosťou aspoň $1 - \epsilon$, tak $P \wedge Q$ s pravdepodobnosťou aspoň $1 - 2\epsilon$

$\Updownarrow$

Ak P, Q platia na objektoch s $\delta(n)$, tak $P \wedge Q$ platia na objektoch s $\delta(n)$

$\rightsquigarrow$ situácia s n rôznymi vlastnosťami

veľká pravdepodobnosť a nestlačiteľnosť nie sú vo všeobecnosti totožné, ale za nejakých predpokladov skoro áno

Lemma

Nech $S \subseteq \{0,1\}^n$

1. *ak* P *spĺňajú* $\forall x \in S$ *s* $\delta(x|S) \leq \delta(n)$, *potom* P *platí aspoň na* $(1 - 1/2^{\delta(n)})$*-tine prvkov* S
2. n, S *fixované,* P *vlastnosť, kt. platí aspoň na* $(1 - 1/2^{\delta(n)})$*-tine prvkov* S*. Potom existuje* c*, že každá taká vlastnosť* P *platí* ***súčasne*** $\forall x \in S$ *s* $\delta(x|S) \leq \delta(n) - K(P|S) - c$

Dôsledok

Objekty s random deficiency $\leq \delta(n)$ *majú všetky vlastnosti* P*, ktoré platia aspoň s pravdepodobnosťou* $1 - 2^{-\delta(n)-K(P|n)-O(1)}$

Dôsledok

Všetky rekurzívne vlastnosti P *s* $K(P|n) = O(1)$*, ktoré samostatne platia s pravdepodobnosťou* $\to 1$ *pre* $n \to \infty$*, platia súčasne s pravdepodobnosťou* $\to 1$ *pre* $n \to \infty$*.*

Nech $S \subseteq \{0,1\}^n$

1. ak P spĺňajú $\forall x \in S$ s $\delta(x|S) \leq \delta(n)$, potom P platí aspoň na $(1 - 1/2^{\delta(n)})$-tine prvkov S

   # programov dĺžky $\leq \log d(S) - \delta(n) \leq \sum_{i=0}^{\log d(S) - \delta(n)} 2^i$

2. n, S fixované, P vlastnosť, kt. platí aspoň na $(1 - 1/2^{\delta(n)})$-tine prvkov S. Potom existuje c, že každá taká vlastnosť P platí **súčasne** $\forall x \in S$ s $\delta(x|S) \leq \delta(n) - K(P|S) - c$

   Sporom:... nech existuje $x \in S : \delta(\mathbf{x}|\mathbf{S}) \leq \delta(\mathbf{n}) - \mathbf{K}(\mathbf{P}|\mathbf{S}) - \mathbf{c}$ a P **neplatí**. Potom rekonštrukcia x:

   - popis P
   - index j v množine M objektov, pre kt. P neplatí $|M| \leq \frac{d(S)}{2^{\delta(n)}}$

$$\exists c_1, c_2 \quad K(x|S) \leq \log j + c_1 \leq \log d(S) - \delta(n) + c_2$$
$$\delta(\mathbf{n}) - \mathbf{K}(\mathbf{P}|\mathbf{S}) - \mathbf{c} \geq \delta(\mathbf{X}|\mathbf{S}) = l(d(S)) - K(X|S)$$
$$\mathbf{c_2} - \mathbf{c} \geq \mathbf{K}(\mathbf{P}|\mathbf{S})$$

Γ - množina turnajov na $\{1, 2, \ldots, n\}$ — zorientovaný úplný graf
– $E : T \to \{0,1\}^{n(n-1)/2}$
– tranzitívny turnaj $T(i,j), T(j,k) \Rightarrow T(i,k)$ — $T(i,j) \leftrightsquigarrow i < j$

$v(n)$ - max. číslo také, že $\forall T \in \Gamma$: T obsahuje tranzitívny podturnaj o $v(n)$ vrcholoch

Theorem

$$v(n) \leq 1 + \lfloor 2 \log n \rfloor$$

Fixnime
- **T** $\in \Gamma$: $C(E(T)|n,p) \geq n(n-1)/2$
- **S** tranzitívny podturnaj na $v(n)$ vrcholoch

popis T

+ zoznam vrcholov S podľa dominancie
- odstránime hrany patriace S z $E(T)$

$$n(n-1)/2 \leq C(E(T)|n,p) \leq n(n-1)/2 + v(n)\lfloor \log n \rfloor - v(n)(v(n)-1)/2$$

□

counting

tranzitívny podturnaj-pôvodne

- Γ' - turnaje, ktoré obsahujú tranzitívny podturnaj na $v = 2 + 2\lfloor \log n \rfloor$ vrcholoch
- $A \subseteq \{1, 2, \ldots, n\}$ $d(A) = v$ σ je permutácia na A
- $d(\Gamma_{A,\sigma}) = 2^{\binom{n}{2}-\binom{v}{2}}$

$$d(\Gamma') \leq \sum_{A,\sigma} d(\Gamma_{A,\sigma}) = \binom{n}{v} v! 2^{\binom{n}{2}-\binom{v}{2}} < 2^{\binom{n}{2}} = d(\Gamma)$$

existuje $T \in \Gamma - \Gamma'$, ktorý nemá tranzitívny podturnaj na v vrcholoch

pravdepodobnostné metódy

$\mathbb{T}$ náhodná premenná s hodnotami z Γ, $\mathbf{Pr}(\mathbb{T} = \mathbf{T}) = \mathbf{2}^{-\binom{n}{2}}$, $\mathbf{v = 2 + 2\lfloor \log n \rfloor}$

$$\sum_{\mathbf{A}} \sum_{\sigma} \mathbf{Pr}(\mathbf{T}|\mathbf{A} \text{ generované podľa } \sigma) = \binom{\mathbf{n}}{\mathbf{v}} \mathbf{v}! \mathbf{2}^{-\binom{v}{2}} < \mathbf{1}$$

existuje taká hodnota T, že T nemá tranzitívny podturnaj na v vrcholoch

turnaj s k-dominátormi

vlatnosť S(k)- $\forall A \subseteq N, d(A) = k$, existuje v $N - A$ hráč x, ktorý dominuje všetkým v A

$s(k)$ - minimálny počet hráčov turnaja T s vlastnosťou S(k)

Theorem

$s(k) \leq 2^k k^2(\ln 2 + o(1))$

Zoberme

$n = 2^k k^2(\ln 2 + o(1))$, pričom $C(E(T)|n, k) \geq n(n-1)/2$

turnaj T, v ktorom existuje $A, d(A) = k$, bez dominujúceho v $N - A$

kratší popis T

+	vrcholy A	$d(A) \log n$
-	odstránime popis hrán $A \times (N - A)$	$k(n-k)$
+	popíšeme ich indexom j v množine možností	$\log(2^k - 1)^{(n-k)}$

$$n(n-1)/2 \leq C(E(T)|n, k, p) \leq n(n-1)/2 + \underbrace{k \log n - (n-k)k + \log(2^k - 1)^{(n-k)}}_{<0}$$

veľmi pravdepodobné vlastnosti

c konštanta, y fixné. Potom každá konečná množina A veľkosti m má aspoň $m(1-2^{-c})+1$ prvkov s $C(x|y) \geq \log m - c$

$\hookrightarrow$ aspoň $2^{n(n-1)/2}(1-1/n)$ turnajov s $C(E(T)|n,p) \geq n(n-1)/2 - \log n$

Dôsledok

Takmer všetky n-vrcholové turnaje (aspoň $(1-1/n)$tina) majú najväčší tranzitívny podturnaj veľkosti nanajvýš $1+2\lfloor 2\log n \rfloor$ $n \to \infty$

$C(E(T)|n,k,p) \geq n(n-1)/2 - \log n$

Dôsledok

Pre dostatočne veľké k existuje n, $n \leq 2^k k^2(\ln 2 + o(1))$ tak, že aspoň $(1-1/n)$-tina turnajov má vlastnosť $S(k)$

$\mathcal{D} = \{D_1, \ldots, D_j\}$, $D_i \subseteq N = \{1, \ldots, n\}$ rozlišuje N, ak $\forall M, M' \subseteq N, M \neq M'$, existuje $i, 1 \leq i \leq j$ tak, že $d(D_i \cap M) \neq d(D_i \cap M')$

$f(n)$ - minimálne $d(\mathcal{D})$ (Vie sa) $f(n) = \frac{2n}{\log n} + O\left(\frac{n \log\log n}{\log^2 n}\right)$

Theorem

$f(n) \geq (2n/\log n)[1 + O(\log\log n/\log n)]$

$M, M \subseteq N$ tak, že $C(E(M)|\mathcal{D}) \geq n$; $d_i = d(D_i)$, $m_i = d(D_i \cap M)$

- s_i je podpostupnosť $E(M)$ odpovedajúca jednotkám v $E(D_i)$; $l(s_i) = d_i$, $\#_1(s_i) = m_i$, $C(s_i|\mathcal{D}) \geq d_i - O(\log i)$

- $d_i/2 - O(\sqrt{d_i \log i}) \leq m_i \leq d_i/2 + O(\sqrt{d_i \log i})$

$$C(x) \geq n - \delta(n), \text{ tak } \underbrace{|\#ones(x) - n/2| \leq \sqrt{\frac{3}{2}(\delta(n) + c)n/loge}}_{A}$$

ak vieme d_i, m_i určíme poradím v A $\rightsquigarrow$ $C(m_i|D_i) \leq \frac{1}{2}\log d_i + O(\log\log i)$

$$n \leq C(E(M)|\mathcal{D}) \leq C(m_1, \ldots, m_j|\mathcal{D}) \leq \sum_{i=1}^{j}\left(\frac{1}{2}\log n + O(\log\log n)\right)$$

Kolmogorovsky náhodný graf - graf je $\delta(n)$-náhodný ak má randomness deficiency $\leq \delta(n)$

$$C(E(G)|n,\delta) \geq n(n-1)/2 - \delta(n)$$

Lemma

1. *Aspoň* $(1 - 1/2^{\delta(n)})$*-tina grafov je* $\delta(n)$*-náhodná.*
2. *Ak* G *je* $\delta(n)$*-náhodný,* d *stupeň ľubovoľného vrchola, potom*
$$|d - (n-1)/2| = O\left(\sqrt{(\delta(n) + \log n)n}\right)$$

1. počet programov
2. - $C(E(G)|n,\delta) \geq n(n-1)/2 - \delta(n)$
   - i je vrchol s $|d - (n-1)/2| > k$
   - $V_i = \{j \in V - \{i\} \mid (i,j) \in E\}$

+ pôvodné kódovanie $E(G)$ $C(E(G)|n,\delta) \geq n(n-1)/2 - \delta(n)$
+ identifikácia i, k $2\log n$
− odstránime pôvodný bitový vektor pre V_i $-n$
+ namiesto bitového vektora pre V_i index do množiny množín susedov mohutnosti m

$$\log \underbrace{\left[\sum_{|d-(n-1)/2|>k} \binom{n-1}{d}\right]}_{m} \leq \log\left[2 \cdot 2^{n-1} e^{-2k^2/3(n-1)}\right] = n - (2k^2/3(n-1)) \log e$$

$$\log m + 2\log n + n(n-1)/2 - n + O(1) \geq n(n-1)/2 - \delta(n)$$

$$n - (2k^2/3(n-1)) \log e \geq \log m \geq n - 2\log n - \delta(n) - O(1)$$

$$k \leq \sqrt{\tfrac{3}{2}\left(\delta(n) + 2\log n + O(1)\right)(n-1)/\log e} = O\left(\sqrt{(\delta(n) + \log n)n}\right)$$

□

Lemma

Všetky $o(n)$-náhodné očíslované grafy majú medzi ľubovoľnou dvojicou vrcholov $n/4 + o(n)$ rôznych ciest dĺžky 2.

1. priemer 1 majú len K_n
2. $i, j \in V$, spojené r cestami dĺžky 2; potom popis G

   popis $O(1)$

   $i, j, i < j$ $\mathbf{2 \log n}$

   $E(G)$ bez bitov popisujúcich hrany $(i,k),(j,k)$ $\binom{\mathbf{n}}{\mathbf{2}} - \mathbf{2(n-2)}$

   najkratší popis zoznamu $e_{i,j}$ odstránených hrán $C(e_{i,j}|n)$

   $$O(\log n) + n(n-1)/2 - 2(n-2) + C(e_{i,j}|n) \geq n(n-1)/2 - o(n)$$

   $$C(e_{i,j}|n) \geq l(e_{i,j}) - o(n) = 2(n-2) - o(n)$$

3. frekvencia 11 v $e_{i,j}$ podľa $|\#y - \frac{n}{2^{l(y)}}| \leq \sqrt{\alpha p n},\ \alpha = K(y|n) + \log l(y) + \delta(n)$

$$\boxed{n/4 + o(n)}$$

$\rightsquigarrow$ $o(n)$-náhodné grafy sú $n/4 + o(n)$-súvislé.

Lemma

$G = (E, V)$ má random.def. $\delta(n) = O(\log n)$. Potom maximálna klika, ktorú G obsahuje, má veľkosť nanajvýš $\lfloor 2 \log n \rfloor + O(1)$.

$$-\frac{k(k-1)}{2} + k \log n + O(1)$$

Lemma

Nech G je aspoň $c \log n$-náhodný očíslovaný, $i \in V$. $\forall i \neq j$ alebo $(i,j) \in E$ alebo existuje k spomedzi $(c+3)\log n$ najmenších $(i,k) \in E$, pričom $(k,j) \in E$

A je množina $(c+3)\log n$ najmenších spomedzi susedov i;
SPOROM: Nech k nie je spojený so žiadnym z $A \cup \{i\}$; potom krátky popis G:

- popis hrán incidentných s i, k $\quad -(2n-2)$
- \+ popis i $\quad \log n$
  bitový vektor o susedoch i $\quad (n-1)$
  popis k+ s ním incidentné hrany $\quad \log n + (n-2) - (c+3)\log n$

$$n(n-1)/2 + 2\log n + 2n - 3 - (c+3)\log n - 2n + 2 \geq n(n-1)/2 - c \log n$$

$$c \log n \geq (c+1)\log n + 1$$

$G = (V, E)$ $H = (V_H, E_H), V_H = \{1, \ldots, k\}$ $G_k = (V_k, E_k)$ podgraf G s k vrcholmi

G_k je výskyt H, ak H dostaneme z G_k premenovaním vrcholov $i_j \rightsquigarrow j$
$//V_k = i_1 < i_2 < \ldots < i_k$

- dvojvrcholové podgrafy - izolované vrcholy, hrana; počet výskytov v $\delta(n)$-náhodnom grafe

$$\frac{n(n-1)}{4} \pm \sqrt{\frac{3}{4}(\delta(n) + O(1))n(n-1)/\log e}$$

- $\#H(G)$ - počet výskytov H ako (usporiadaného pomenovaného) podgrafu G (vrátane prekryvov)
- $p = 2^{-k(k+1)/2}$ pravdepodobnosť, že H dostaneme hádzaním fair mince

Theorem

$G = (V, E)$, k delí n, H očíslovaný graf na $k \leq \sqrt{2 \log n}$ vrcholoch. Potom

$$\left|\#H(G) - \binom{n}{k} p\right| \leq \binom{n}{k} \sqrt{\alpha(k/n)p},$$

$\alpha = \left(K(H|n) + \delta(n) + \log \binom{n}{k}/(n/k) + O(1)\right) 3/\log e$

$G = (V, E)$, k delí n, H očíslovaný graf na $k \leq \sqrt{2 \log n}$ vrcholoch. Potom

$$\left| \#H(G) - \binom{n}{k} p \right| \leq \binom{n}{k} \sqrt{\alpha (k/n) p},$$

$\alpha = \left(K(H|n) + \delta(n) + \log \binom{n}{k} / (n/k) + O(1) \right) 3 / \log e$

- pokrytie grafu G je mn. $C = \{S_1, \ldots, S_N\}, N = n/k$ disjunktných mn. $V = \bigcup_{i=1}^{N} S_i$
- **(Baranayi)**existuje rozklad $\binom{n}{k}$ rôznych k-vrcholových mn. na $h = \binom{n}{k}/N$ rôznych pokrytí G, každé pozostávajúce z $N = n/k$ podmnožín $\mathbf{C_0, C_1, \ldots, C_{h-1}}$
- $i \in \{0, \ldots, h-1\}$, H k-vrcholový graf $\mathbf{\#H(G, i)}$ počet (neprekrývajúcich sa) výskytov podgrafu H grafu G v C_i
- N pokusov $\rightsquigarrow$ k-vrcholové podgrafy pokrytia
  – reťazec s_i dĺžky $N\binom{k}{2}$
  – rekonštrukcia G z $i, s_i, n + \binom{n}{2} - N\binom{k}{2} \rightsquigarrow C(s_i|n) \geq l(s_i) - \delta(n) - \log h$
  – počet výskytov $\#H(G, i) = Np \pm \sqrt{Np\alpha}$

$$\left| \#H(G) - p\binom{n}{k} \right| = \sum_{i=0}^{h-1} |\#H(G, i) - Np| \leq \binom{n}{k} k \sqrt{\alpha (k/n) p}$$ □

automorfizmus - permutácia vrcholov π taká, že $(\pi(u), \pi(v)) \in E \Leftrightarrow (u, v) \in E$

- grupa (skladanie, identita)
- $G, \pi(G)$ majú rovnaké štandardné kódovanie $E(G) = E(\pi(G))$

rigid/stabilný graf $\leftrightsquigarrow$ jediným automorfizmom je identita

$g(n)$ počet neočíslovaných n-vrcholových grafov

$\mathcal{G}_n$ trieda neorientovaných grafov na $\{0, 1, \ldots, n-1\}$
$\mathcal{G}_n = \mathcal{G}_n^0 \cup \mathcal{G}_n^1 \cup \mathcal{G}_n^n$ v $\mathcal{G}_n^m$ sa každým automorfizmom pohne m vrcholov

$S(n)$ grupa permutácií n-prvkovej množiny

$\overline{G}$ trieda grafov izomorfných s G

$Aut(G)$ grupa automorfizmov grafu G

Theorem

Počet neočíslovaných n-vrcholových grafov $g_n \sim \frac{2^{\binom{n}{2}}}{n!}$

- $g_n = \sum_{G \in \mathcal{G}_n} \frac{1}{d(\overline{G})}, \quad d(\overline{G}) = \frac{d(S_n)}{d(Aut(G))} = \frac{n!}{d(Aut(G))}$
- $g_n = \sum_{G \in \mathcal{G}_n} \frac{d(Aut(G))}{n!} = \frac{2^{\binom{n}{2}}}{n!} E_n$, $E_n = \sum_{G \in \mathcal{G}_n} Pr(G) d(Aut(G))$; $E_n > 1$; $\mathcal{G}_n^1 = \emptyset$

  F1 $\forall G \in \mathcal{G}_n^m,\ d(Aut(G)) \leq \binom{n}{m} m! \leq n^m = 2^{m \log n}$ ✓

  F2 $Pr(G \in \mathcal{G}_n^m) \leq 2^{-m(\frac{n}{2} - \frac{3m}{8} - 2\log n)}$!!KZ

- $E_n = \sum_{m=0}^{n} Pr(G \in \mathcal{G}_n^m) Avg_{G \in \mathcal{G}_n^m} d(Aut(G)) \leq 1 + \sum_{m=2}^{n} 2^{-m(\frac{n}{2} - \frac{3m}{8} - 2\log n)}$
  $\leq 1 + 2^{-n+4\log n+2}$
- $\frac{2^{\binom{n}{2}}}{n!} \leq g(n) \leq \frac{2^{\binom{n}{2}}}{n!} \left(1 + \frac{4n^4}{2^n}\right)$

$F2 \quad Pr(G \in \mathcal{G}_n^m) \leq 2^{-m(\frac{n}{2} - \frac{3m}{8} - \log n)}$

- π, m vrcholov $i_1 < \ldots i_m$ sa pohne $\quad \pi(\mathbf{i_1}), \ldots, \pi(\mathbf{i_m}) \quad m \log n$
- $\pi \rightsquigarrow k$ cyklov veľkosti $c_1, \ldots, c_k$, v každom najmenší zvolený
- nezvolený
  - netreba hrany do stabilných $\quad -(n-m)$
  - netreba niektoré hrany do nezvolených $\quad -\frac{m-k}{2}$

$$\sum_{i=1}^{k}(c_i - 1)(n - m + \frac{m-k}{2}) = (m-k)(n - \frac{m+k}{2})$$

$$m \log n - (\mathbf{m} - \mathbf{k})(\mathbf{n} - \frac{\mathbf{m}+\mathbf{k}}{\mathbf{2}}) = \underbrace{-\frac{m}{2}(n - \frac{3m}{4} - 2\log n)}_{\delta(n,m)} \qquad //k = m/2$$

efektívna enumerácia neočíslovaných grafov

$\mathcal{G} \leftarrow \emptyset$

for all očíslované grafy G–binárne reťazce **do**

  if G nevieme získať premenovaním $H \in \mathcal{G}$ **then**

    $\mathcal{G} \leftarrow \mathcal{G} \cup G$

$$\Downarrow \textit{odhad } g(n) + \textit{Stirling}$$

$$C(E(G)|n) \leq \binom{n}{2} - n \log n + O(n)$$

Theorem

Nech G je očíslovaný n-vrcholový, G_0 jeho neočíslovaná verzia. Potom existujú graf G' a permutácia π tak, že $G' = \pi(G)$ a (až na konštanty)

$$C(E(G')) = C(E(G_0)) \qquad C(E(G)|n) = C(E(G_0), \pi|n)$$

Theorem

ROUTOVANIE po najkratších cestách v $O(\log n)$-náhodných grafoch/sieťach sa dá spraviť s lokálnou pamäťou $6n$, celkovou $6n^2$ bitov.

- nanajvýš dĺžka 2; $u \rightsquigarrow v$ cez $\log n$ najmenších susedov u $\quad \underline{O(n \log\log n)}$
- $v_i, \ldots, v_m$ je $O(\log n)$ najmenších susedov u; $A_0 \subseteq V$ nesusedia u.
- $A_t \stackrel{def}{:=} \{w \in A_0 - \bigcup_{s=1}^{t-1} A_s : (v_t, w) \in E\}$
- $\mathbf{m_0} = d(A_0),\ \mathbf{m_{t+1}} = m_t - d(A_{t+1})$
- Nech ℓ je min. také, že $m_\ell < n/\log\log n$
- konštrukcia lokálnej routovacej funkcie $F(u)$
  - tabuľka pre A_0:

$$\text{pre } w \begin{cases} \in \bigcup_{s-1}^{t} A_s & \text{unárne } v0\text{: } (u,v) \in E, (v,w) \in E \\ \notin \bigcup_{s-1}^{t} A_s, & 0 \end{cases}$$

  - zvyšok (0 v prvej tab) - nanajvýš $m_\ell < n/\log\log n$ explicitne binárne

$O(\log n)$-náhodný graf
$v_i, \ldots, v_m$ je $O(\log n)$ najmenších susedov u;
$A_0 \subseteq V$ nesusedia u; $A_t \stackrel{def}{:=} \{w \in A_0 - \bigcup_{s=1}^{t-1} A_s : (v_t, w) \in E\}$
$\mathbf{m_0} = d(A_0)$, $\mathbf{m_{t+1}} = m_t - d(A_{t+1})$

Nech ℓ je min. také, že $m_\ell < n/\log\log n$. Potom $d(A_t) > m_{t-1}/3$ pre $1 \leq t \leq \ell$

Sporom – AK $\exists t, |d(A_t) - m_{t-1}/2| > m_{t-1}/6$ POTOM krátky popis G

+	u, v_t	$2\log n$
+	charakteristické postupnosti $A_0, \ldots, A_{t-1}$	$r = n-1+\cdots+n-(t-1)$
+	samoodeľujúco A_t v $A_0 - \bigcup_{s=1}^{t-1} A_s$	$m_{t-1} - \frac{2}{3}(1/6)^2 m_{t-1}\log e + O(\log m_{t-1})$
-	netreba hrany $\{v_t\} \times (A_0 - \bigcup_{s=1}^{t-1} A_s)$	$-m_{t-1}$
-	netreba hrany incidentné s $u, v_1, \ldots, v_t$	$-r$

$$n(n-1)/2 - O(\log n) \leq C(E(G)) \leq n(n-1)/2 + O(\log n) + m_{t-1} - \frac{2}{3}(1/6)^2 m_{t-1}\log e - m_{t-1}$$

$$m_{t-1} \leq O(\log n) \quad \bigwedge \quad m_{t-1} > n/\log\log n$$

$A_0 \subseteq V$ nesusedia u; $A_t = \{w \in A_0 - \bigcup_{s=1}^{t-1} A_s : (v_t, w) \in E\}$

ℓ je min. také, že $m_\ell < n/\log\log n$, $d(A_t) > m_{t-1}/3$ pre $1 \leq t \leq \ell$

konštrukcia lokálnej routovacej fcie $F(u)$

- tabuľka pre A_0:

$$\text{pre } w \begin{cases} \in \bigcup_{s-1}^{t} A_s & \text{unárne } v0\text{: } (u,v) \in E, (v,w) \in E \\ \notin \bigcup_{s-1}^{t} A_s, & 0 \end{cases}$$

$$n + \sum_{s=1}^{\ell} \frac{1}{3}\left(\frac{2}{3}\right)^{s-1} sn \leq n + \sum_{s=1}^{\infty} \frac{1}{3}\left(\frac{2}{3}\right)^{s-1} sn \leq 4n$$

- zvyšok (0 v prvej tab) - nanajvýš $m_\ell < n/\log\log n$ explicitne binárne

$\log\log n + O(n)$ bitov na určenie poradia (mimo $v_1, \ldots, v_m$) $2n$

$$|F(u)| \leq 6n$$

Theorem

Routovanie po najkratších cestách v o(n)-náhodných grafoch/sieťach VYŽADUJE lokálnu pamäť $n/2 - o(n)$, celkovo $n^2/2 - o(n^2)$ bitov.

Uvažujme nasledujúci popis

- \+ u — $\log n$
- \+ samooddeľujúco $F(u)$ — $(F(u)) + 2\log(d(F(u)))$
- \+ hrany medzi u a zvyškom
- \− odstránime $(v, w) \in E$, keď $F(u, w) = v$
  $|d - \frac{(n-1)}{2}| = O(\sqrt{(\delta(n) + \log n)n})$ — ušetríme aspoň $n/2 - o(n)$

$$\frac{n(n-1)}{2} + O(1) + \underbrace{O(\log n) + 2\log(d(F(u)))}_{O(\log n)} + d(F(u)) - \frac{n}{2} + o(n) \geq C(E(G)) \geq \frac{n(n-1)}{2} - o(n)$$

$$d(F(u)) \geq \frac{n}{2} - o(n)$$

heapsort

$A[1..n] \rightsquigarrow$ **halda**-otec väčší ako synovia $\rightsquigarrow$ triedenie $\rightsquigarrow A[1] < \ldots A[n]$

Heapify

for $i = \lfloor n/2 \rfloor$ downto 1 **do**
kľúč pôvodne $A[i]$ padá na svoje miesto cez väčšieho syna

Sort

for $i = n$ downto 2 **do**
vymeň $k_i = A[1] \leftrightarrow A[i] = k$ a **preusporiadaj** $A[1..i-1]$ na haldu

Williams— porovaním oboch synov s výmenou cez väčšieho syna padá na svoje miesto

2d

Floyd — porovnaním synov presun (bez výmeny) cez väčšieho syna do listu; potom hore (bez výmeny) na miesto j, ktoré treba vymeniť s koreňom ($A[j] \leq k$); výmena a posun všetkých na ceste z j do koreňa o 1 hore

$\mathbf{d} + \mathbf{2}\delta; \delta = \log \mathbf{n} - \mathbf{d}$

Theorem

Heapsort robí v priemere

$$\begin{array}{lll} \textit{presunov:} & n\log n + O(n) & \\ \textit{porovnaní:} & \textit{Wiliams} & 2n\log n - O(n) \\ & \textit{Floyd} & n\log n + O(n) \end{array}$$

n kľúčov, $n!$ listov // $n^n e^{-n}\sqrt{2\pi n}$

existuje permutácia p : $C(p|n) \geq n\log n - 2n$ (*)

Fakt

Nech h je halda po Heapify pri vstupe p podľa (*). Potom

$$C(h|n) \geq n\log n - 6n$$

(SPOROM) Nech by $C(h|n) < n\log n - 6n$. Krátky popis p:

- popis Heapify, kt. z p vyrobí h - popis cesty na svoje miesto

  $$n\log 3 \sum_{i=1}^{\log n} i/2^{i+i} \leq 2n\log 3$$

- $C(p|n) \leq C(h|n) + 2n\log 3 + O(1) < n\log n - 6n + 2n\log 3 + O(1) < n\log n - 2n \diamond$

H popis činnosti v časti sort:
nová pozícia "koreňa" - cesta v strome: s_i H zreťazenie s_i
$\hookrightarrow$ samooddeľujúco $\delta_i = \log n - l(s_i)$ a za to s_i $l(s_i) + 2\log\delta_i$
$\hookrightarrow$ h vieme zrekonštruovať z H
(reverzným postupom zrekonštruujem alebo pri dostatku času nájdem permutáciu p)
$\diamond$

$C(h|n) \leq l(H) + O(1)$

$$n\log n - 6n \leq C(h|n) \leq l(H) + O(1) \quad \Rightarrow \quad \boxed{l(H) \geq n\log n - 6n}$$

$$l(H) = \sum_{i=1}^{n}(l(s_i) + 2\log\delta_i) = \sum_{i=1}^{n}\log n - \sum_{i=1}^{n}(\delta_i - 2\log\delta_i) \geq n\log n - 6n$$

$\sum_{i=1}^{n}\delta_i = O(n) \Rightarrow$ súčet priemerných ciest je $n\log n - nc$ // # presunov v sort

$\Rightarrow$ # porovnaní Williams $2n\log n - O(n)$
$\Rightarrow$ # porovnaní Floyd $n\log n + O(n)\square$

Najdlhší spoločný podreťazec

- $s = s_1 \dots s_m$, $t = t_1 \dots t_n$
  s je podreťazec t ak existuje $i_1 < \dots < i_m : \ s_j = t_{i_j}$

S-množina reťazcov

- $SCS(S)$ - najkratší reťazec s: $\forall s' \in S$ s' je podreťazec s
- $LCS(S)$ - najdlhší reťazec, ktorý je podreťazcom každého z S
  $\Sigma = \{a_1, \dots, a_k\}$, $lcs(S)$-dĺžka $LCS(S)$, $S \subseteq \Sigma^*$

Algoritmus Long-Run

1. nájdi maximálne m také, že (pre nejaké $a \in \Sigma$) a^m je podreťazec s $\forall s \in S$
2. return a^m

Theorem

Nech S je n-prvková množina reťazcov dĺžky n. Algoritmus Long-Run vypočíta spoločný podreťazec dĺžky $lcs(S) - O(lcs(S)^{1/2+\epsilon})$ pre aspoň $(1 - 1/n^2)$-tinu vstupov, a teda v priemere.

Algoritmus Long-Run

1. nájdi maximálne m také, že (pre nejaké $a \in \Sigma$) a^m je podreťazec s $\forall s \in S$
2. return a^m

Algoritmus Long-Run vypočíta spoločný podreťazec dĺžky $lcs(S) - O(lcs(S)^{1/2+\epsilon})$ pre aspoň $(1 - 1/n^2)$-tinu vstupov

$S = \{x_1, \ldots, x_n\} \quad \rightsquigarrow \quad x = x_1 \ldots x_n : \; C(x) \geq (n^2 - 2\log n)\log k$

Fakt

Ak $|\#_a(x_i) - n/k| > n^{1/2+\epsilon}$, *tak existuje* $\delta > 0$ *tak, že*

$$C(x_i|k) \leq (n - \delta n^{2\epsilon})\log k$$

//ak $C(x) \geq n - \delta(n)$ tak $|\#y(x) - pn| \leq \sqrt{\alpha pn}$, $\alpha = (K(y|n) + \log l(y) + \delta(n) + c)3l/loge$
resp. index do množiny veľkosti $\binom{n}{m}(k-1)^{n-m}$

Fakt ($\hookrightarrow$)

$$lcs(S) < \frac{n}{k} + n^{1/2+\epsilon}$$

$lcs(S) < n/k + n^{1/2+\epsilon}$ a
$\hookrightarrow$ každé $a \in \Sigma$ v každom x_i aspoň $(n/k) - O(n^{1/2+\epsilon})$
$\hookrightarrow$ a^m dĺžky $(n/k) - O(n^{1/2+\epsilon})$ je spoločný podreťazec
$\hookrightarrow$ $lcs(S) - l(m) = O(n^{1/2+\epsilon}) \quad l(m) = lcs(S) - O(lcs(S))^{1/2+\epsilon}$

$S = \{x_1, \ldots, x_n\} \quad \rightsquigarrow \quad x = x_1 \ldots x_n : \; C(x) \geq (n^2 - 2\log n)\log k$

$s = lcs(S)$, kvôli sporu $l(s) > \frac{n}{k} + n^{1/2+\epsilon}$ //chceme $lcs(S) < \frac{n}{k} + n^{1/2+\epsilon}$

krátky popis x :

- $s = s_1 \ldots s_p, s_i \in \Sigma$, fixujme x_i
- vložíme s do x_i tak, že volíme najľavejší možný výskyt; $x_i = \underbrace{\alpha_1 s_1 \alpha_2 s_2 \ldots \alpha_p s_p}_{y} z$

  $\Rightarrow \alpha_j \in (\Sigma - s_j)^*$
- zameňme v y $s_i \to a_k$ a výskyt a_k v α_j nahradíme s_j

  $y' = \alpha'_1 a_k \alpha'_2 a_k \ldots \alpha'_p a_k$ $\qquad \boxed{\#_{a_k}(y') \geq (n/k) + n^{1/2+\epsilon}}$

- $\#_{a_k}(y') \geq (n/k) + n^{1/2+\epsilon}$, preto $C(y'|k) \leq (l(y') - \delta n^{2\epsilon})\log k$
- $C(x_i|k, s) \leq (n - \delta n^{2\epsilon})\log k + O(\log n)$
- $C(x|k) \leq \sum C(x_i|k, s) + l(s) + \underbrace{n \log n}_{\text{oddeľovače}}$
- $C(x|k)/\log k < n^2 - 2\log n$ spor $\diamond$

Nech $L \subseteq \Sigma^*$ je regulárny, $L_x = \{y \mid xy \in L\}$. Potom existuje konštanta c, že $\forall x$: ak y je n-té slovo v L_x, tak $C(y) \leq C(n) + c$

- stav DKA po dočítaní x
- n
- diskusia ◇

$\Sigma^* = \{y_1, y_2, \ldots\}$;
pre $L \subseteq \Sigma^*, x \in \Sigma^*$ je $\mathbb{X} = X_1 X_2 \ldots$ charakteristická postup. L_x: $X_i = 1 \Leftrightarrow xy_i \in L$.

Theorem

Nech $L \subseteq \Sigma^*$. Existuje konštanta c_L, že nasledujúce podmienky sú ekvivalentné

1. *L je regulárny*
2. $\forall x \in \Sigma^* \forall n \in \mathbb{N}: \; C(\mathbb{X}_{1..n}|n) \leq c_L$
3. $\forall x \in \Sigma^* \forall n \in \mathbb{N}: \; C(\mathbb{X}_{1..n}) \leq c_L + C(n)$
4. $\forall x \in \Sigma^* \forall n \in \mathbb{N}: \; C(\mathbb{X}_{1..n}) \leq \log n + c_L$

Ak $\forall x \in \Sigma^* \forall n \in \mathbb{N}: \; C(\mathbb{X}_{1..n}) \leq \log n + c_L$ tak L je regulárny

Lemma

$\forall c$ *existuje konečne veľa postupností* $w \in \{0,1\}^*$ *takých, že*
$\forall n \; C(w_{1..n}) \leq \log n + c$

$\hookrightarrow$ pravá kongruencia konečného indexu $\sim$ $\quad x \sim x' \Leftrightarrow \mathbb{X} = \mathbb{X}'$
$\hookrightarrow$ konečný automat □

k dôkazu lemy:

- $A_n = \{x \in \{0,1\}^n : C(x) \leq \log n + c\}$;
  $A = \{w \in \{0,1\}^\infty : \forall n \; C(w_{1..n}) \leq \log n + c\}$
- ak $d(A_n) \leq c'$ pre nekonečne veľa n, tak $d(A) \leq c'$
- fixnime $\ell \in \mathbb{N}$: y dĺžky $2\ell + c + 1$, $C(y) \geq 2\ell + c + 1$
- vezmime i maximálne také, že $y = mn$, $l(m) = i$, $m \leq d(A_n)$
  nech $y = sr$, $l(s) = i + 1$
- pomocou programov $l(p) \leq \log n$ vymenovávanie A_n; y_0- m-tý v A_n;
  rekonštrukcia $y = mn$ z $y_0 \hookrightarrow n = l(y_m), m : y_0 = y_m$

$$l(n) + l(m) = 2\ell + c + 1 \leq C(y) \leq C(y_0) + O(1) \leq \log n + c + O(1)$$

on-line TS - pred pohybom hlavy na vstupe akcept/reject na príslušný prefix

$L = \{y\#x_1 * x_2 * \ldots x_k : \ \exists i \ y = ux_i^R v\}$

Theorem

Viacpáskový TS akceptuje L online v čase $\Omega(n^2/\log n)$

Lemma

Nech $n = l(x)$, p program. Ak $C(x|n,p) \geq n$, potom žiaden podreťazec dĺžky $> 2\log n$ sa v x nevyskytuje viac ako raz.

Lemma

Ak sa v reťazci neopakuje reťazec dĺžky m, tak je jednoznačne určený množinou svojich podreťazcov dĺžky $m+1$.

- $\hookrightarrow$ $m = 3\log n \quad l(x_l) = m$
- $\hookrightarrow$ pre y, $l(y) = n$, $C(y) \geq n$ vyrobíme vstup taký, že žiadne x_l nie je reverzom podreťazca v y ale jeho spracovanie vyžaduje $n\epsilon$ krokov;
  pre $k = \Omega(n/\log n)$ vynútime $\Omega(n^2/\log n)$ krokov

Nech $n = l(x)$, p program. Ak $C(x|n,p) \geq n$, potom žiaden podreťazec dĺžky $> 2\log n$ sa v x nevyskytuje viac ako raz.

- $x = uvw$, $l(v) > \log n$, v sa v uv vyskytuje presne dvakrát
- povieme i, j– indexy začiatku v v x $2\log n$
- povieme uw $//l(v) = n - l(uw)//$ $l(uw)$

$$C(x|p,n) \leq n - 2\log n + \log n(n-1) < n$$

Ak sa v reťazci neopakuje reťazec dĺžky m, tak je jednoznačne určený množinou $S_{x,m+1}$ svojich podreťazcov dĺžky $m+1$.
$a, b \in \{0,1\}, u, v, w \in \{0,1\}^*$

- prefix ua je jednoznačne určený podmienkou $\forall b$, $bu \notin S_{x,m+1}$
- $\forall$ prefix vw, $l(w) = m$ existuje **práve jedno** $b \in \{0,1\}$: $wb \in S_{x,m+1}$ a vwb je prefix x

$\hookrightarrow$ rekonštrukcia x zo znalosti $S_{x,m+1}$

$m = 3 \log n \quad l(x_l) = m$; pre y, $l(y) = n$, $C(y) \geq n$ vyrobíme vstup taký, že žiadne x_l nie je reverzom podreťazca v y ale jeho spracovanie vyžaduje $n\epsilon$ krokov

majme to po $y\#x_1 * \ldots * x_{i-1}*$

1. ak $i - 1 = k$ ✓
2. ak také x_i neexistuje, vyrobíme krátky popis na generovanie $S_{y,m+1}$ //a teda y
   - diskusia $O(1)$
   - obsah pracovných pások do vzdialenosti $n\epsilon$ od hláv v čase t_0, keď T ukončil spracovanie prefixu $y\#x_1 * \ldots x_{i-1}*$ $O(n\epsilon)$
   - polohy hláv v čase t_0 $O(\log n\epsilon)$
   - popis $T, n, stav$ v čase t_0 $O(1) + \log n$

$$C(y) \leq O(1) + O(n\epsilon) + O(\log n\epsilon) + logn < n$$

□

Pre FA je $(k+1)$ hláv viac ako k

$$L_b = \{w_1\# \cdots \#w_b\#\#w_b\# \cdots \#w_1 \mid w_i \in \{0,1\}^*\}, b = \binom{k}{2} + 1$$

1. $L_b \in (k+1)DFA$
2. $L_n \notin (k)NFA$ SPOROM:
   - vezmeme nestlačiteľné w, $C(w) \geq l(w), l(w) \ggg \log l(w)$
   - $w = w_1 \dots w_b \quad \rightsquigarrow \quad$ korektný vstup I
   - NKA kontroluje w_i ak súčasne jedna hlava na ľavej a jedna na pravej kópii w_i

Lemma

M musí skontrolovať každé w_i

- $[q, p(h_1), \dots, p(h_k)]$ — postupnosť hlavových konfigurácií v okamihoch, keď jedna z hláv prvýkrát prichádza na p $\quad O(k \log(l(I)) + l(M)) = O(\log l(w))$
- r_l, r_r-L a R pozícia pravej kópie w_i, $p(r_l), p(r_r)$ príslušné prechodové postupnosti

$\hookrightarrow$ rekonštrukcia w_i z

- popis $w - w_i$
- popis $p(r_l), p(r_r)$

$$C(w_i|w - w_i) \leq O(\log l(w))$$

$$C(w) \leq l(w) - l(w_i) + O(\log l(w)) < l(w)$$

$L = \{x_1 * x_2 * \cdots * x_k \# y_1 * y_2 * \cdots * y_\ell \# \# 0^i 1^j,\ x_i = y_j\}$

rozpoznávanie L

- na 2-páskovom on-line v lineárnom čase
- na 1-páskovom on-line vyžaduje čas $\Omega(n^{3/2}/\log n)$

↪ Simulácia lineárneho 2-páskového on-line na 1-páskovom on-line vyžaduje čas $\Omega(n^{3/2}/\log n)$

1páskový on-line T:

h_1, h_2 – pozícia hlavy na vstupe, resp. pracovnej páske
S – súvislý segment vstupnej pásky
R – súvislý segment pracovnej pásky

T zobrazuje S DO R, ak h_2 neopustí R kým je h_1 v S
T zobrazuje S NA R, ak h_2 prečíta R kým je h_1 v S

$x\# = x_1 * x_2 * \cdots * x_k\#, \ \forall i \ l(x_i) = l(x)/k$

R je taký segment na p.páske, že T zobrazuje $\forall$ segment z $S = \{x_{i_1}, \ldots, x_{i_\ell}\}$ do R

Lemma

Obsah pracovnej pásky v čase, keď h_1 dorazila na $\#$ vieme zrekonštruovať, keď poznáme

- $\overline{S} = \{x_i \mid 1 \leq i \leq k\} - S$
- *výsledný obsah segmentu R*
- *prechodové postupnosti okolo R* l_R, r_R
- *popis T +diskusia*

miesta blokov z S necháme prázdne

— zrekonštruujeme pásku naľavo od r_L

— zrekonštruujeme pásku napravo od r_R

NECH 1-páskový T rozpoznáva L v čase $T_1(n) < c^{-5} n^{3/2} / \log n$

- vezmime $x \in \{0,1\}^*,\ l(x) = n,\ C(x) \geq n,\ x = x_1 \ldots x_k,\ k = \sqrt{n},\ |x_i| = \sqrt{n} \forall i$
- uvažujme vstup $x_1 * x_2 * \ldots * x_k \#$ hlava na $\#$ v čase $t_\#$

$\Rightarrow$ ak sa viac ako $k/2$ z x_i sa zobrazí NA segment veľkosti aspoň n/c^3, tak pre čas T

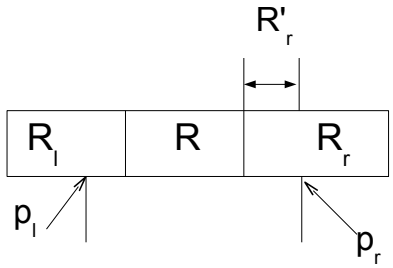
$$T_1 = \Omega(n/c^3 \cdot \underbrace{\sqrt{n}/2}_{k/2}) = \Omega\left(n^{3/2}\right)$$

$\Rightarrow$ teda S obsahuje $k/2$ blokov x_i, ktoré sa zobrazia DO segmentu veľkosti $\leq n/c^3$
x_m - medián pri usporiadaní podľa ľavých okrajov segmentov, do kt. sa zobrazia

1. veľa $x_i \in S$ sa zobrazí do malého R //lema
2. neexistuje taký R a segmenty sú rozložené "rovnomerne"
   //y_j umiestnime ďaleko od x_i

1. existuje k/c blokov $x_i \in S$ a segment R dĺžky n/c^2, že sa všetky zobrazia do $R \rightsquigarrow \overline{S}$

$|R_l| = |R| = |R_r|$,
p_l, p_r pozície PP v R_l, R_r, kt. sú najkratšie



AK $l(PP(p_r)), l(PP(p_l)) > \sqrt{n}/(c^2 \log n)$, potom čas aspoň $\sqrt{n}/(c^2 \log n) \cdot n/c^2$

PRETO $l(PP(p_r)), l(PP(p_l)) \leq \sqrt{n}/(c^2 \log n)$:

krátky program pre x

- diskusia + popis T_1 — $O(1)$
- hodnoty n, k, c, p_l, p_r — $O(\log n)$
- zreťazenie $\{x_i, \ldots, x_k\} - \overline{S}$ — $n - n/c$
- stav T a pozícia h_2 v čase, keď h_1 sedí na $\#$ — $O(\log n)$
- 2 prechodové postupnosti na pozíciách p_l, p_r v čase $t_\#$ — $2\sqrt{n}(l(T) + O(\log n))$
- obsah $R'_l R R'_r$ v $t_\#$ — $3n/c^2 + O(\log n)$

overíme, že kandidát $y = x$

- $l(x) = l(y)$
- rekonštrukcia pamäťovej pásky, keď h_1 prvýkrát vstúpila na $\#$
- $y \quad \hookrightarrow \quad y_1 * \ldots * y_k$;
  $\forall 0^i 1^i$ zbehneme simuláciu T_1 od $t_\#$
  akceptuje len ak $x = y$

$$C(x) \leq n - \frac{n}{c} + \underbrace{\frac{3n}{c^2} + O(\sqrt{n}\log n) + O(\log n)}_{\leq \frac{n}{c}} \leq \gamma n; \quad 0 < \gamma < 1$$

spor s $C(x) \geq n$

2. pre $\forall$ blok R veľkosti $l(R) = n/c^2$ existuje nanajvýš k/c blokov $x_i \in S$, kt. sa zobrazia do neho

- do R_m sa zobrazí medián x_m
- aspoň $k/6$ napravo od $R_m \quad \hookrightarrow \quad S_r = \{x_{i_1}, \ldots, x_{i_{k/6}}\}$
  analogicky S_l naľavo od $R_m \hookrightarrow \quad S_l = \{x_{j_1}, \ldots, x_{j_{k/6}}\}$
- $y_1 = x_{i_1}, y_2 = x_{j_1}, \ldots$ $\boxed{x_{j_s} = y_{2s}, \quad x_{i_s} = y_{2s-1}}$
- pre vstup $I = x_1 * x_2 * \cdots * x_k \# y_1 * \cdots y_{k/3} \#$ existuje dvojica $y_{2s-1} * y_{2s}$, že sú namapované do segmentu menšieho ako $n/(4c^2)$ //inak čas aspoň $\frac{k}{6} \cdot \frac{n}{4c^2} = \frac{n^{3/2}}{24c^2}$
- y_{2s-1}, y_{2s} vo vzdialenosti aspoň n/c^3 od x_{i_s} alebo x_{j_s}, w.l.o.g. od x_{i_s}
  $\hookrightarrow x_{i_s} \ldots R \ldots y_{2s-1}$,
- suffix $0^{i_s} 1^{2s-1}$
- $l(R) = n/c^3$, p pozícia najkratšej prechodovej postupnosti v R, kt. dĺžka je $\max \sqrt{n}/(c^2 \log n)$

krátky popis x

- diskusia + popis T_1 $O(1)$
- n, k, c, pozícia p $O(\log n)$
- $S - \{x_{i_s}\}$ $n - \sqrt{n}$
- index i_s $O(\log n)$
- prechodová postupnosť na pozícii p $\leq \sqrt{n}/c$
  // dľžka max $\sqrt{n}/(c^2 \log n)$

$$C(X) \leq n - \sqrt{n} + \sqrt{n}/c + O(\log n) \leq n - \gamma\sqrt{n}$$

spor s $C(x) \geq n$

PRAM - concurrent read, priority write(minimálny uspeje)

Theorem

Sčítanie(násobenie) n celých čísel dĺžky aspoň n^ϵ pre fixované $\epsilon > 0$, vyžaduje $\Omega(\log n)$ paralelných krokov.

$P = n^d$, kvôli SPORu $T < (\epsilon/2) \cdot \log n$ //súčet

- nech $n^{\epsilon/2} d(\log n)(\log\log n) < n^\epsilon/2$
- program poskytneme orákulu $(i, \ell) \hookrightarrow A \hookrightarrow$ pôvodný úsek programu dĺžky ℓ pre procesor $P(i)$
- fixneme nestlačiteľný reťazec $\mathbf{x}$, $|\mathbf{x}| = \mathbf{n^{1+\epsilon}}$, $\mathbf{C(x|A,T,P,n) \geq n^{1+\epsilon}}$

$$\mathbf{x = x_1 \ldots x_n},\ |\mathbf{x_i}| = \mathbf{n^\epsilon}, \mathbf{c(i) \leftarrow x_i}$$

procesor je živý v kroku t ak

- píše výstup, alebo
- zapísal niečo v kroku $t' \geq t$, čo neskôr $t'' \geq t'$ prečíta v tom čase t'' živý procesor

vstup je užitočný, ak je v t čítaný procesorom, ktorý je v tom čase živý

$\hookrightarrow$ počet užitočných vstupov, resp. živých procesorov je nanajvýš 2^T

USEFUL – užitočné vstupy

ALIVE – $\{(P(i), t_i) : \ p_i$ je živý až do času $t_i > 0\}$

$T < (\epsilon/2) \log n \ \Rightarrow \quad 2^T < n \ \Rightarrow$ existuje kus vstupu x_{i_0}, ktorý nie je užitočný

rekonštrukcia x

- pomocou $A, ALIVE, USEFUL$ zrekonštruujeme $\sum x_i$

$$2^T (\log P)(\log T) \leq n^{\epsilon/2}(d \log n)(\log \log n) < n^\epsilon/2$$

- pomocou A, selfdelimiting indexu i_0 a zreťazenia $x_i, i \neq i_0$ zrekonštruujeme zo $\sum x_i$ prvok x_{i_0}

$$\mathbf{n^{1+\epsilon} - n^\epsilon + 2 \log n + 1}$$

- v správnom poradí zapíšeme x_i

$$C(x) \leq n^{1+\epsilon} - n^\epsilon + 2 \log n + n^\epsilon/2 + O(1) < n^{1+\epsilon}$$

SPOR s $C(x|A, T, P, n) \geq n^{1+\epsilon}$

□

$C_{AB}(x \leftrightarrow y)$=min$\{|p| : A(p, y) = x, B(p, x) = y\}$ informačná vzdialenosť

- $C_{AB}(x \leftrightarrow y) \geq \max\{C(x|y), C(y|x)\} + O(\log \max\{C(x|y), C(y|x)\})$

Dôkaz pomocou online farbenia grafov (susediace hrany majú rôznu farbu)

hra A a B

A generuje hrany pri zachovaní $d(G) \leq d$

B určuje farby
greedy - najmenšia možná farba $\rightsquigarrow 2d - 1$ farieb stačí

konštrukcia grafu G_k pre $k = \max\{C(x|y), C(y|x)\}$

vrcholy binárne reťazce dĺžky max k

hrany $(x, y) \in E \Leftrightarrow (C(x|y) \leq k \wedge C(y|x) \leq k)$

počet hrán, $d(G_k) \leq 2^{k+1} - 1$, počet farieb $\leq 2^{k+2} - 3$

algoritmus T: vstup k, výstup: postupnosť hrán grafu G_k a ich farieb

algoritmus A

- pozná y, vstup $\bar{k}i$
- simulácia $T(k)$, výstup $x :< (x, y), i >$ vo výstupe T

$$\begin{aligned}(x \leftrightarrow y) \leq C_{AA}(x \leftrightarrow y) &\leq |\bar{k}i| = 2\log k + 1 + k \\ &= O(\log \max\{C(x|y), C(y|x)\} + \max\{C(x|y), C(y|x)\})\end{aligned}$$

□