

Egzamin z Teorii informacji. Zadania

8 lutego 2019

1. Niech x będzie słowem binarnym zawierającym k jedynek i $n - k$ zer.

(a) Dowiedź, że zwykła złożoność Kołmogorowa $C(x)$ spełnia ograniczenie

$$C(x) \leq n\mathcal{H}\left(\frac{k}{n}\right) + 2\log\log n + 2\log n + c, \quad (1)$$

gdzie $\mathcal{H}(p)$ to entropia rozkładu dwupunktowego $(p, 1 - p)$, a stała c nie zależy od n i k .

(b) Dowiedź, że powyższe ograniczenie pozostaje prawdziwe dla bezprefiksowej złożoności Kołmogorowa, tzn. dla x jak wyżej

$$K(x) \leq n\mathcal{H}\left(\frac{k}{n}\right) + 2\log\log n + 2\log n + c, \quad (2)$$

gdzie stała c nadal nie zależy od n i k .

2. Niech $X_i : \Omega \rightarrow \{0, 1\}$, dla $i = 1, \dots, n$, będą niezależnymi binarnymi zmiennymi losowymi o identycznym rozkładzie. Rozważamy zmienną losową $X = (X_1, \dots, X_n)$ o wartościach w $\{0, 1\}^n$.

Korzystając z ograniczenia (2) dowiedź, że

$$0 \leq H(X_1) - \mathbf{E} \mathcal{H}\left(\frac{\sharp_1(X)}{n}\right) \leq \frac{2\log\log n + 2\log n + c}{n}, \quad (3)$$

gdzie $\sharp_1(x)$ oznacza liczbę jedynek w słowie x , $H(Y)$ to entropia zmiennej Y , a stała c nie zależy od n .

Uwaga. Mogą się przydać następujące zależności.

Nierówność Jensena $\mathbf{E} f(S) \leq f(\mathbf{E} S)$, dla zmiennej losowej $S : \Omega \rightarrow [0, 1]$ i wklęsłej funkcji f .

Nierówność z dowodu Tw. Shannona

$$\sum_{i \leq \lambda \cdot n} \binom{n}{i} \leq 2^{n \cdot \mathcal{H}(\lambda)}, \quad (4)$$

dla $0 \leq \lambda \leq \frac{1}{2}$.

Rozwiązanie zadania 1. Wykażemy od razu drugi punkt. Oczywiście wystarczy wskazać dowolną maszynę bezprefiksową, która dla dowolnego słowa x , gdzie $|x| = n \geq 1$ i $\sharp_1(x) = k \geq 0$ wygeneruje x z wejścia v_x o długości ograniczonej przez prawą stronę nierówności. Idea jest taka, że słowa długości n zawierające dokładnie k jedynek ustawiamy w ciąg (np. w porządku leksykograficznym) o $\binom{n}{k}$ elementach i elementem generującym słowo x jest

krotka (n, k, n_x) , gdzie n_x jest numerem x we wspomnianym ciągu. Dokładniej, niech $\text{bin}(m)$ oznacza standardową reprezentację binarną liczby $m \geq 1$ długości $|\text{bin}(m)| = (\lfloor \log m \rfloor + 1)$. Natomiast $\text{bin}'(m)$ oznacza „bezprefiksową” reprezentację m (znaną z ćwiczeń), np. w ciąg $\text{bin}(m)$ wkładamy co drugą cyfrę 0, a ostatnią 1. Słowo v_x , z którego nasza maszyna generuje x jest konkatencją słów następującej postaci:

- $\text{bin}'(|\text{bin}(n)|)$ długości $\leq 2(\log \log n) + 2$,
- $\text{bin}(n)$,
- $\text{bin}(k)$ w razie potrzeby dopełnione początkowymi zerami, tak by długość tego słowa była dokładnie $|\text{bin}(n)|$,
- $\text{bin}(n_x)$ w razie potrzeby dopełnione początkowymi zerami, tak by długość tego słowa była dokładnie $\left| \text{bin} \left(\frac{n}{k} \right) \right|$.

Dla oszacowania długości tego ostatniego słowa wykorzystamy nierówność (4). Otrzymujemy z niej $\binom{n}{k} \leq 2^{n \cdot \mathcal{H}(\frac{k}{n})}$, dla $\frac{k}{n} \leq \frac{1}{2}$, ale ponieważ $\binom{n}{k} = \binom{n}{n-k}$ i $\mathcal{H}(\frac{k}{n}) = \mathcal{H}(\frac{n-k}{n})$, mamy tę nierówność dla każdego $k \leq n$ i w konsekwencji $\log \binom{n}{k} \leq n \cdot \mathcal{H}(\frac{k}{n})$. Długość całego słowa da się więc oszacować

$$|v_x| \leq 2 \log \log n + 2 \log n + \mathcal{H} \left(\frac{k}{n} \right) + \text{const},$$

co daje nierówność z zadania.

Jedynie słowa postaci v_x dla pewnego x są akceptowane, co gwarantuje bezprefiksowość.

Rozwiązanie zadania 2. Zauważmy, że uśredniona złożoność Kolmogorowa spełnia nierówność

$$H(X) \leq \sum_{x \in \{0,1\}^n} p(X=x) K(x). \quad (5)$$

Jest tak dlatego, że programy dla bezprefiksowej maszyny uniwersalnej poświadczające złożoność Kolmogorowa (tzn. $U(y) = x$) tworzą kod bezprefiksowy dla wartości zmiennej X ; możemy więc skorzystać z ogólnej zależności między średnią długością kodu a entropią.

Skorzystamy teraz z nierówności (2), przy czym zastosujemy skróty

$$\begin{aligned} \ell(n) &= 2 \log \log n + 2 \log n \\ p(X=x) &= p(x) \end{aligned}$$

Jeśli po prawej stronie nierówności (5) zgrupujemy słowa o tej samej liczbie jedynek $k = 0, 1, \dots, n$, i dla każdego k zastosujemy (2), otrzymamy po oczywistym

przekształceniu

$$\begin{aligned} \sum_{k=0}^n \sum_{\sharp_1(x)=k} p(x)K(x) &\leq \sum_{k=0}^n \underbrace{\left(\sum_{\sharp_1(x)=k} p(x) \right)}_{p(\sharp_1(X)=k)} n\mathcal{H}\left(\frac{k}{n}\right) + \ell(n) + c \\ &= \mathbf{E} n\mathcal{H}\left(\frac{\sharp_1(X)}{n}\right) + \ell(n) + c \end{aligned}$$

gdzie ostatnia równość wynika wprost z definicji wartości oczekiwanej zastosowanej do zmiennej $n\mathcal{H}\left(\frac{\sharp_1(X)}{n}\right)$.

Lewa strona nierówności (5) wynosi oczywiście $H(X) = \sum_{i=1}^n H(X_i) = nH(X_1)$ (z założenia o niezależności i jednakowym rozkładzie zmiennych X_i). Zatem z obu nierówności, po podzieleniu przez n , otrzymujemy

$$H(X_1) \leq \mathbf{E} \mathcal{H}\left(\frac{\sharp_1(X)}{n}\right) + \frac{\ell(n)n + c}{n},$$

czyli drugą nierówność z zadania. Pierwszą nierówność (czyli $0 \leq H(X_1) - \mathbf{E} \mathcal{H}\left(\frac{\sharp_1(X)}{n}\right)$) otrzymamy, korzystając z nierówności Jensena dla funkcji $\mathcal{H}$

$$\mathbf{E} \mathcal{H}\left(\frac{\sharp_1(X)}{n}\right) \leq \mathcal{H} \mathbf{E}\left(\frac{\sharp_1(X)}{n}\right) = \mathcal{H}(p(X_1 = 1)) = H(X_1),$$

gdzie przedostatnia równość wynika z addytywności wartości oczekiwanej

$$\mathbf{E} \sharp_1(X) = \mathbf{E} (X_1 + \dots + X_n) = n\mathbf{E} X_1.$$

Ł.Dębowski, T.Gogacz, D.Niwiński