Egzamin z Teorii informacji. Rozwiązanie zadania o kanale.

10 lutego 2017

Zadanie

Rozważmy kanał Γ_n, w którym alfabet wejściowy Σ_A oraz wyjściowy Σ_B są równe i wynoszą $\Sigma_A = \Sigma_B = \{0,1\}^n$, natomiast przesłanie przez kanał pojedynczego symbolu $x = (x_1, x_2, \ldots, x_n)$ powoduje obrót cykliczny* o k pozycji, $k < n$. Z prawdopodobieństwem $\frac{1}{2}$ obrót ten wynosi 0 (czyli x pozostaje bez zmian), a z prawdopodobieństwem $\frac{1}{2(n-1)}$ obrót wynosi $k \in \{1, \ldots, n-1\}$.

*Obrót cykliczny ciągu $x = (x_1, x_2 \ldots, x_n)$ o k pozycji daje ciąg $(x_{k+1}, x_{k+2}, \ldots, x_n, x_1, \ldots, x_k)$.

Dokładniej, niech dla słów $x, y \in \{0,1\}^n$,

$$cycle(x,y) \quad = \quad \{k \in \{1,\ldots,n-1\} : x \text{ przesunięte cyklicznie o } k \text{ pozycji daje } y\}$$

Wówczas

$$\mathbb{P}(B = y | A = x) \quad = \quad \begin{cases} \frac{1}{2(n-1)} \cdot |cycle(x,y)| & \text{gdy } x \neq y \\ \\ \frac{1}{2(n-1)} \cdot |cycle(x,y)| + \frac{1}{2} & \text{gdy } x = y. \end{cases}$$

Przykład: dla $n = 4$ mamy
$\mathbb{P}(B = 0000 | A = 0000) = \frac{1}{6} \cdot 3 + \frac{1}{2} = 1$,
$\mathbb{P}(B = 1010 | A = 1010) = \frac{1}{6} \cdot 1 + \frac{1}{2} = \frac{2}{3}$, $\mathbb{P}(B = 0101 | A = 1010) = \frac{1}{6} \cdot 2 = \frac{1}{3}$, $\mathbb{P}(B = 0011 | A = 1010) = 0$.

Udowodnij, że przepustowość tego kanału wynosi co najmniej $n - 1 - \frac{1}{2} \cdot \log(n-1)$.

Wskazówka. Rozważyć najpierw przypadek, gdy n jest liczbą pierwszą. (Jak wtedy może wyglądać zbiór $cycle(x,y)$?) Rozwiązania ograniczone do tego przypadku też będą punktowane.

Rozwiązanie 1 — rachunkowe z definicji. Oszacujemy z dołu wartość $I(A;B) = H(B) - H(B|A)$.

Na początek pokażemy, że zmienna wyjściowa B może mieć rozkład jednostajny i tym samym entropię równą n; dzieje się tak w szczególności, gdy A ma rozkład jednostajny. Pamiętając, że $p(B = y) = \sum_x p(A = x) \cdot P(x \to y)$, wystarczy sprawdzić, że suma każdej kolumny w macierzy kanału daje 1. Istotnie, dla różnych x, x', zbiory $cycle(x,y)$ i $cycle(x',y)$ są rozłączne i $\bigcup_x cycle(x,y) = \{1,\ldots,n-1\}$ (obrót o daną liczbę pozycji jest bijekcją na słowach), skąd otrzymujemy

$$\sum_x |cycle(x,y)| \quad = \quad n-1. \tag{1}$$

Zatem

$$\sum_x P(x \to y) = \left(\frac{1}{2} + \frac{|cycle(y,y)}{2(n-1)}\right) + \sum_{x \neq y} \frac{|cycle(x,y)|}{2(n-1)} = \frac{1}{2} + \sum_x \frac{|cycle(x,y)|}{2(n-1)} = 1. \tag{2}$$

Pozostaje oszacować z góry $H(B|A)$. Mamy

$$H(B|x) \quad = \quad \sum_y -p(B = y|x) \cdot \log p(B = y|x) \tag{3}$$

$$\quad = \quad -\left(\frac{1}{2} + \frac{|cycle(x,x)|}{2(n-1)}\right) \log \left(\frac{1}{2} + \frac{|cycle(x,x)|}{2(n-1)}\right) + \sum_{y \neq x} \frac{|cycle(x,y)|}{2(n-1)} \log \frac{2(n-1)}{|cycle(x,y)|}. \tag{4}$$

Zauważmy najpierw, że dla x równego 0^n lub 1^n, mamy $cycle(x,x) = \{1,\ldots,n-1\}$ i powyższa entropia wynosi 0.

W dalszych szacowaniach wykorzystamy własność dualną do równości (1), którą uzasadniamy analogicznie: dla dowolnego x,

$$\sum_y |cycle(x,y)| \quad = \quad n-1. \tag{5}$$

Rozważmy najpierw sugerowany we Wskazówce przypadek, gdy n jest liczbą pierwszą. Dla $x \neq 0^n, 1^n$ mamy wówczas $cycle(x,x) = \emptyset$ i $(\forall y)\ |cycle(x,y)| \leq 1$. (Jest tak dlatego, że minimalny obrót nakładający słowo na siebie musi być dzielnikiem n.) Wtedy z równości (3) i (5) otrzymujemy

$$\begin{aligned} H(B|x) &= \frac{1}{2} + (n-1) \cdot \frac{1}{2(n-1)} \log 2(n-1) \\ &= 1 + \frac{1}{2} \log(n-1) \end{aligned}$$

i w konsekwencji $H(B|A) \leq 1 + \frac{1}{2}\log(n-1)$, skąd otrzymujemy tezę zadania. (Zauważmy, że nierówność jest ostra z powodu słów $0^n, 1^n$.)

Gdy n jest liczbą złożoną, rachunek jest nieco bardziej skomplikowany, ale szacowania działają na naszą korzyść. (Jest to zgodne z intuicją – nietrywialny obrót nakładający słowo na siebie koryguje błędy kanału.) Mamy oczywiście $-\log\left(\frac{1}{2} + \frac{|cycle(x,x)|}{2(n-1)}\right) \leq \log \frac{2(n-1)}{|cycle(x,x)|}$, a ponadto, dla wszystkich y, $\log \frac{2(n-1)}{|cycle(x,y)|} = \log 2(n-1) - \log|cycle(x,y)| \leq \log 2(n-1)$ (bo $|cycle(x,y)| \leq n-1$). Oznaczmy na chwilę $wait = -\frac{1}{2} \cdot \log\left(\frac{1}{2} + \frac{|cycle(x,x)|}{2(n-1)}\right)$. Z równości (3) otrzymujemy

$$\begin{aligned} H(B|x) &\leq wait + \sum_y \frac{|cycle(x,y)|}{2(n-1)} \cdot \log 2(n-1) \\ &= wait + \frac{1}{2} \cdot \log 2(n-1) \end{aligned}$$

(w ostatniej równości znów korzystamy z (5)). Z drugiej strony mamy $-\log\left(\frac{1}{2} + \frac{|cycle(x,x)|}{2(n-1)}\right) \leq -\log\frac{1}{2} = 1$, tak więc $wait \leq \frac{1}{2}$. Stąd otrzymujemy $(\forall x)\ H(B|x) \leq 1 + \frac{1}{2}\log(n-1)$ i w konsekwencji tezę zadania.

Rozwiązanie 2 — „wysokopoziomowe". Idąc za pierwszym, intuicyjnym opisem kanału wprowadzimy zmienną losową $turn$ o wartościach w zbiorze $\{0, 1, \ldots, n-1\}$ i rozkładzie $p(turn = 0) = \frac{1}{2}$, $p(turn = i) = \frac{1}{2(n-1)}$, dla $i = 1, \ldots, n-1$. Zauważmy, że

$$\begin{aligned} H(turn) &= \frac{1}{2} + (n-1) \cdot \frac{1}{2(n-1)} \log 2(n-1) && (6) \\ &= 1 + \frac{1}{2}\log(n-1). && (7) \end{aligned}$$

Ważną obserwacją jest, że zmienna A jest funkcją B i $turn$. (Jest tak dlatego, że obrót o daną liczbę pozycji jest bijekcją na słowach; z tej własności korzystaliśmy także w Rozwiązaniu 1.)
Oszacujemy teraz $H(A|B)$. Mamy

$$\begin{aligned} H(A|B) &\leq H(A, turn|B) && (8) \\ &= H(turn|B) && (9) \\ &\leq H(turn). && (10) \end{aligned}$$

Nierówność (8) otrzymujemy z warunkowej reguły łańcuchowej $H(A, turn|B) = H(turn|A,B) + H(A|B)$. Równość (9) otrzymujemy z warunkowej reguły łańcuchowej w jej symetrycznym wariancie $H(A, turn|B) = H(A|turn, B) + H(turn|B)$, gdyż pierwszy składnik jest równy 0, ponieważ, jak zauważyliśmy, A jest funkcją zmiennych $turn$ i B. Ostatnia nierówność (10) jest ogólną własnością entropii warunkowej.
Biorąc A o rozkładzie jednostajnym (czyli $H(A) = n$) otrzymujemy tezę zadania.

T.Gogacz, D.Malinowski, D.Niwiński