Egzamin z Teorii informacji 5 lutego 2018.

Uzasadnienia nie były wymagane. Podajemy je dla informacji.

1. Przypuśćmy, że następujące kody kodują wartości zmiennych losowych X i Y,

$$\begin{array}{ll} X: & \{0, 100, 101, 110, 111\}, \\ Y: & \{000, 001, 01, 10, 11\} \end{array}$$

przy czym średnie długości tych kodów pokrywają się z entropiami odpowiednich zmiennych. Porównać $H(X)$ i $H(Y)$.

Odpowiedź: $H(X) < H(Y)$.

Uzasadnienie. Kod binarny, którego średnia długość pokrywa się z entropią istnieje tylko, gdy prawdopodobieństwa są całkowitymi potęgami $\frac{1}{2}$ i wtedy długości są logarytmami z odpowiednich prawdopodobieństw. To pozwala nam odtworzyć rozkłady X i Y. Dla X mamy 5 wartości o prawdopodobieństwach $p(x_1) = \frac{1}{2}$, $p(x_1) = \ldots = p(x_5) = \frac{1}{8}$; dla Y, $p(y_1) = p(y_2) = \frac{1}{8}$, $p(x_3) = p(x_4) = p(x_5) = \frac{1}{4}$, skąd $H(X) = \frac{1}{2} \cdot 1 + 4 \cdot \left(\frac{1}{8} \cdot 3\right) = 2$, $H(Y) = 2 \cdot \left(\frac{1}{8} \cdot 3\right) + 3 \cdot \left(\frac{1}{4} \cdot 2\right) = 2.25$.

2. Rzucamy dwa razy niezależnie sprawiedliwą kostką do gry. Zmienne losowe A_1, A_2 o wartościach w zbiorze $\{1, 2, \ldots, 6\}$ reprezentują wyniki rzutów. Określamy $A = A_1 + A_2$ i $B = A_1 \cdot A_2$. Uporządkować według wielkości liczby

$$I(A_1; A_2),\ I\left(A; B|(A_1, A_2)\right),\ I(A; B|A_1),\ H(A|A_1),\ H(B|A_2)$$

Odpowiedź: $I(A_1; A_2) = I\left(A; B|(A_1, A_2)\right) < I(A; B|A_1) = H(A|A_1) = H(B|A_2)$.

Uzasadnienie. $I(A_1; A_2) = 0$ z niezależności A_1 i A_2.

$I\left(A; B|(A_1, A_2)\right) = \underbrace{H(A|(A_1, A_2))}_{0} - \underbrace{H(A|B, (A_1, A_2))}_{0} = 0$, bo A jest funkcją (A_1, A_2).

$0 < I(A; B|A_1)$, bo A i B **nie** są niezależne względem A_1.

(Np. $0 < p(A = 2|A_1 = 1) = p(B = 1|A_1 = 1) = p(A = 2 \wedge B = 1|A_1)$.)

Z kolei $I(A; B|A_1) = H(A|A_1) - \underbrace{H(A|B, A_1)}_{0} = H(A|A_1)$, bo A jest funkcją B i A_1.

Analogicznie $I(A; B|A_1) = H(B|A_1)$. Ale z symetrii $H(B|A_1) = H(B|A_2)$, stąd ostatnia równość.

3. Niech, dla $k \geq 2$, C_k będzie kodem Hamminga o k bitach kontrolnych (*checking bits*); przypomnijmy, że kod ten składa się ze słów długości $2^k - 1$ otrzymanych przez przedłużenie wszystkich słów długości $2^k - k - 1$ konstrukcją Hamminga. Niech $\begin{pmatrix} P & Q \\ Q & P \end{pmatrix}$ będzie kanałem BSC. Wartość $Pr_E(\Delta, C_k)$ określamy, jak w Tw. Shannona. Przypuśćmy, że dla nieskończenie wielu k zachodzi $Pr_E(\Delta, C_k) = 0$. Co można wtedy powiedzieć o P i Q?

Odpowiedź: $P = 1, Q = 0$.

Uzasadnienie. Wiemy, że $Pr_E(\Delta, C) = 0$ implikuje $R(C) \leq C_\Gamma$. Ale ciąg $R(C_k) = \frac{2^k - k - 1}{2^k - 1}$ dąży do 1, zatem nasze przypuszczenie implikuje, że $C_\Gamma = 1$. Są dwa kanały BSC o tej własności: $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ i $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Jednak ten drugi możemy odrzucić, bo reguła najbliższego sąsiada dla żadnego nietrywialnego kodu nie da błędu 0. (Po przejściu przez taki kanał słowo wynikowe jest dokładnie **najdalszym** sąsiadem słowa wejściowego.)

4. Ile błędów wykrywają, a ile poprawiają następujące kody?

(a) słowa długości 100 (sto), w których 1 występuje dokładnie 5 razy,
Wykrywa 1, poprawia 0.

(b) słowa długości 100, w których liczba jedynek jest podzielna przez 5,
Wykrywa 1, poprawia 0.

(c) słowa długości 100 zbudowane z bloków tych samych liter długości 5, tzn. $\{0^5, 1^5\}^* \cap \{0,1\}^{100}$.
Wykrywa 4, poprawia 2.

Szacowania wynikają z charakteryzacji przez minimalną odległość między słowami kodowymi, która dla (4a) i (4b) wynosi 2, a dla (4c) 5.

5. Przypomnijmy, że słowo $w \in \{0,1\}^*$ jest losowe w sensie Kołmogorowa, jeśli $|w| \leq C_U(w)$. Rozważmy zdanie: W zbiorze $\mathcal{D}$ istnieje nieskończenie wiele liczb losowych. Dla których z poniższych zbiorów jest ono prawdziwe (proszę napisać: TAK lub NIE):

- palindromy, NIE
- słowa długości parzystej, TAK
- słowa długości n^2 kodujące macierz incydencji grafu skierowanego, TAK
- słowa długości n^2 kodujące macierz incydencji grafu symetrycznego, NIE
- słowa, stanowiące binarny zapis potęgi jakiejś liczby pierwszej. NIE

W przypadku odpowiedzi „NIE" wystarczy wskazać efektywny skrót, tzn. funkcję $w \mapsto \mathit{short}(w)$ i maszynę Turinga M (algorytm), która z $\mathit{short}(w)$ jednoznacznie odtwarza w, przy czym $|w| - |\mathit{short}(w)|$ asymptotycznie przekracza dowolną stałą. Wtedy $C_U(w) \leq |\mathit{short}(w)| + c_{MU} < |w|$.

Np. dla palindromów to może być „połowa" wraz z informacją o parzystości długości. W innych przypadkach podobnie. Odpowiedzi „TAK" wynikają z faktu, że słowa losowe istnieją dla każdej długości, a każde słowo długości n^2 jest macierzą incydencji jakiegoś grafu.