# Egzamin z Teorii informacji 9 lutego 2021

### Zadanie egzaminacyjne nr 1

Jest troje drzwi: 1,2,3. Za jednymi z nich została ukryta nagroda, jej położenie opisuje zmienna losowa $X \in \{1, 2, 3\}$ z jednostajnym rozkładem prawdopodobieństwa.

Uczestnik turnieju wybiera drzwi $Y \in \{1, 2, 3\}$; zakładamy, że zmienna $Y$ ma rozkład jednostajny oraz $I(X; Y) = 0$.

Z kolei prowadzący turniej losowo[1]wybiera drzwi $Z$ spośród tych, które nie są równe $X$ ani $Y$.

(Możemy myśleć, że otwiera te drzwi, pokazując, że nie ma za nimi nagrody.)

Należy obliczyć $I(X; Y|Z)$.

Jak zmieni się odpowiedź, gdy prowadzący może wybrać[2] $Z = Y$ (o ile $Y \neq X$) ?

Jak zmieni się relacja między wartościami $I(X; Y|Z)$ w obu rozważanych wyżej przypadkach, jeśli nie zakładamy, że zmienna $Y$ ma rozkład jednostajny, choć nadal zakładamy, że $X$ ma rozkład jednostajny oraz, że $I(X; Y) = 0$ ?

**Solution.**

Let us start with the second case, when we allow $Z = Y$. We know that $I(X; Y \mid Z) = I(X; Y) - R(X; Y; Z)$, where $I(X; Y) = 0$. On the other hand

$$\begin{aligned} R(X; Y; Z) &= I(X; Z) - I(X; Z \mid Y) \\ &= 0, \end{aligned}$$

since by assumption the host opens the doors randomly, independently of $Y$. Therefore $I(X; Y \mid Z) = 0$ in this case. Intuitively, the player gains no information seeing the open doors.

Note that we have not used the assumption of $Y$ being uniform, so the argument applies for any $Y$ independent of $X$.

Now let us compute $R(X; Y; Z)$ in the first case. It is convenient to present it as $I(X; Z) - I(X; Z \mid Y)$. We have

$$I(X; Z) = \underbrace{H(X)}_{\log 3} - \underbrace{H(X|Z)}_{1}$$

where the equality $H(X|Z) = \log 2 = 1$ follows from symmetry. We also have

$$I(X; Z \mid Y) = H(Z \mid Y) - H(Z \mid X; Y).$$

Now again $H(Z \mid Y) = \log 2 = 1$ follows from symmetry[3]. To compute $H(Z \mid X; Y)$, we consider two cases (recall that $\Pr(Z \neq X, Y) = 1$). For $x = y \neq z$, $\Pr(Z = z \mid X = Y = x) = \frac{1}{2}$, whereas for $x, y, z$ all different, $\Pr(Z = z \mid X = Y = x) = 1$. Since $\Pr(X = Y) = \frac{1}{3}$, we obtain

$$H(Z \mid X; Y) = \sum_{x,y} H(Z \mid X = x, Y = y) = \underbrace{H(\frac{1}{2})}_{1} \cdot \frac{1}{3} + \underbrace{H(1)}_{0} \cdot \frac{2}{3} = \frac{1}{3},$$

[1] A więc, jeśli jest dwoje takich drzwi, wybiera jedne z nich z prawdopodobieństwem $\frac{1}{2}$, a gdy $X \neq Y$, wybiera jedyne możliwe drzwi różne od $X$ i $Y$.

[2] A więc w każdym przypadku wybiera z prawdopodobieństwem $\frac{1}{2}$ jedne z dwojga drzwi różnych od $X$.

[3] If you are skeptical, compute directly $\Pr(z|y)$, using the fact that $\Pr(Z = z \wedge Y = y) = \Pr(Z = z \wedge Y = y \wedge X = y) + \Pr(Z = z \wedge Y = y \wedge X \neq y)$..

and hence

$$\begin{aligned}
R(X;Y;Z) &= I(X;Z) - I(X;Z \mid Y) \\
&= (\log 3 - 1) - (1 - \frac{1}{3}) \\
&= \log 3 - \frac{5}{3} \\
I(X;Y \mid Z) &= \underbrace{I(X;Y)}_{0} - R(X;Y;Z) \\
&= \frac{5}{3} - \log 3.
\end{aligned}$$

To complete the solution, we should show that, for any $Y$ independent of $X$, we have $I(X;Y \mid Z) > 0$.

**Zadanie egzaminacyjne nr 2**

Przypomnijmy, że bezprefiksową złożoność Kołmogorowa słowa $y \in \{0,1\}^*$ określamy jako

$$K(y) \quad = \quad \min\{|x| : U(x) = y\},$$

gdzie $U$ jest pewną ustaloną bezprefiksową uniwersalną maszyną Turinga.

Dla liczby naturalnej $n$, $K(n)$ określamy jako bezprefiksową złożoność Kołmogorowa jej zapisu binarnego (bez dodatkowych zer).

Dowieść, że jeśli $p$ i $q$ są liczbami pierwszymi, to

$$K(p+q) \quad \leq \quad K(p \cdot q) + c,$$

gdzie stała $c$ zależy tylko od maszyny $U$.

**Rozwiązanie.** Konstruujemy maszynę Turinga $M$, która dla wejścia $x$, najpierw symuluje maszynę $U$ na $x$. Jeśli $U$ zatrzymuje się z wynikiem $U(x)$, to $M$ traktuje ten wynik jak zapis binarny pewnej liczby naturalnej (ewentualnie ignorując początkowe zera), rozkłada tę liczbę na czynniki pierwsze i dodaje je do siebie. Wynikiem obliczenia $M(x)$ jest więc zapis binarny sumy wszystkich dzielników pierwszych (być może z krotnościami) liczby reprezentowanej przez $U(x)$.

Bezprefiksowość maszyny $M$ wynika wprost z bezprefiksowości $U$. Dalej, jeśli $U(x)$ jest zapisem liczby $p \cdot q$, dla $p$ i $q$ pierwszych (nie zakładamy, że są różne), to $M(x)$ jest zapisem liczby $p+q$; w skrócie $M(x) = p+q$. Jeśli ponadto $|x| = K(p \cdot q)$, to otrzymujemy

$$K_M(p+q) \quad \leq \quad K(p \cdot q),$$

gdzie

$$K_M(p+q) \quad = \quad \min\{|z| : M(z) = p+q\}.$$

Stąd, z uniwersalności $U$,

$$\begin{aligned}
K(p+q) &\leq K_M(p+q) + |\langle M \rangle| \\
&\leq K(p \cdot q) + |\langle M \rangle|.
\end{aligned}$$

**Zadanie egzaminacyjne nr 3**

Porównać przepustowości kanałów

$$\begin{pmatrix} p & 1-p \\ q & 1-q \end{pmatrix} \tag{1}$$

$$\begin{pmatrix} \frac{p}{2} & \frac{p}{2} & 1-p \\ \\ \frac{q}{2} & \frac{q}{2} & 1-q \end{pmatrix} \tag{2}$$

**Wskazówka**. Nie trzeba tych przepustowości obliczać.

**Rozwiązanie** . Przepustowości są równe. Załóżmy, że $(A, B)$ i $(A, B')$ są parami wejścia-wyjścia dla pierwszego i drugiego kanału, odpowiednio. Wykażemy, że

$$H(A|B) \quad = \quad H(A|B').$$

Załóżmy, że alfabetem wejściowym dla obu kanałów jest $\{0, 1\}$, a alfabety wyjściowe to odpowiednio $\{a, b\}$ i $\{a, a', b\}$. Z postaci kanałów mamy

$$\Pr(A = i \wedge B' = a) = \Pr(A = i \wedge B' = a') = \frac{1}{2} \Pr(A = i \wedge B = a)$$

i oczywiście $\Pr(A = i \wedge B = b) = \Pr(A = i \wedge B' = b)$, dla $i = 0, 1$. Wynika stąda, że

$$\Pr(B' = a) = \Pr(B' = a') = \frac{1}{2} \Pr(B = a), \quad \Pr(B' = b) = \Pr(B = b)$$

i dalej

$$\begin{aligned} \Pr(A = i \mid B' = a) = \Pr(A = i \mid B' = a') = \Pr(A = i \mid B = a), \\ \Pr(A = i \mid B' = b) = \Pr(A = i \mid B = b), \end{aligned}$$

dla $i = 0, 1$. W konsekwencji

$$\begin{aligned} H(A \mid B') &= H(A \mid B' = a) \cdot \Pr(B' = a) + H(A \mid B' = a') \cdot \Pr(B' = a') + H(A \mid B' = b) \\ &= H(A \mid B = a) \cdot \Pr(B = a) + H(A \mid B = b) \cdot \Pr(B = b) \\ &= H(A \mid B). \end{aligned}$$

A zatem $I(A; B) = I(A; B')$ i wobec dowolności wyboru $A$, przepustowości obu kanałów są sobie równe.

**Uwaga** . Możliwe jest jest rozwiązanie bardziej „wysokopoziomowe”, odwołujące się do kaskadowego złożenia kanałów (patrz praca pana Jakuba Jasiulewicza).

**Zadanie egzaminacyjne nr 4**

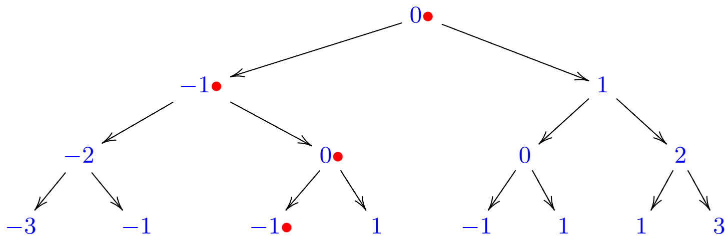
Pionek wędruje po pełnym drzewie binarnym głębokości $N$ od korzenia w kierunku liści, w każdym wierzchołku (z wyjątkiem liści) wybierając ruch w lewo lub w prawo z prawdopodobieństwem $\frac{1}{2}$.

Niech zmienna $T_k$, dla $k = 0, 1, \ldots, N$, opisuje położenie pionka po $k$ ruchach. (Jej wartościami są wierzchołki drzewa na $k$-tym poziomie.)

Zakładamy ponadto, że pionek płaci za każdy ruch w lewo 1 zł, a za każdy ruch w prawo zyskuje 1 zł. Początkowy kapitał pionka jest $S_0 = 0$, a zmienna $S_k$ opisuje jego kapitał po $k$ ruchach.

Dla $k < N$, która spośród zmiennych $S_k$, $S_{k+1}$, $T_{k+1}$, daje najwięcej, a która najmniej informacji o $T_k$ ?

**Przykład.** $N = 3$, na rysunku na niebiesko zaznaczona jest wartość kapitału osiągana w danym wierzchołku.




Źródłem losowości są jedynie wybory kierunku, w lewo lub w prawo, z prawdopodobieństwem $\frac{1}{2}$. Możemy więc przyjąć, że przestrzeń zdarzeń elementarnych, na których określone są nasze zmienne losowe, stanowią wszystkie ścieżki od korzenia do liścia; prawdopodobieństwo jednej ścieżki jest $\left(\frac{1}{2}\right)^N$. Jeśli przyjąć reprezentację wierzchołków drzewa binarnego przez słowa binarne (0 – lewo, 1 – prawo), to dla ścieżki zaznaczonej na rysunku (przez $\bullet$), mamy

$$\begin{array}{lcl@{\qquad\qquad}lcr} T_0 &=& \epsilon & S_0 &=& 0 \\ T_1 &=& 0 & S_1 &=& -1 \\ T_2 &=& 01 & S_2 &=& 0 \\ T_3 &=& 010 & S_3 &=& -1 \end{array}$$

Jak widać na rysunku, ta sama wartość kapitału może być przyjmowana na danym poziomie dla różnych ścieżek, np.

$$\Pr(S_2 = 0) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

Oczywiście, rozważane przez nas zmienne nie są niezależne; zależności między nimi są właśnie przedmiotem zadania.

**Rozwiązanie.** Przypadek $k = 0$ jest trywialny ($H(T_0) = 0$); dalej zakładamy, że $k \geq 1$. Z definicji $T_k$ jest funkcją $T_{k+1}$, więc $I(T_k; T_{k+1}) = H(T_k)$ osiąga maksymalną wartość, jaką może osiągnąć $I(T_k; X)$, dla jakiejkolwiek zmiennej losowej $X$.

Podobnie $S_k$ jest funkcją $T_k$, więc $I(T_k; S_k) = H(S_k)$. Mamy oczywiście $H(S_1) = 1 = H(T_1)$, ale, począwszy od $k \geq 2$, $S_k$ ma mniej wartości niż $T_k$, więc $H(S_k) < k = H(T_k)$, zatem

$$I(T_k; S_k) \quad < \quad I(T_k; T_{k+1}).$$

Pozostaje porównać $H(S_k)$ i $I(T_k; S_{k+1})$. Można to zrobić na różne sposoby. Z definicji wynika, że dla ustalonego wierzchołka $t$ na poziomie $k$, rozkład $S_{k+1}$ jest jednostajny dwupunktowy, zatem

$$\begin{array}{rcl} I(T_k; S_{k+1} \mid S_k) &=& \underbrace{H(S_{k+1}|S_k)}_{1} - \underbrace{H(S_{k+1}|T_k, S_k)}_{1} \\ &=& 0, \end{array}$$

podczas gdy

$$\begin{array}{rcl} I(T_k; S_k \mid S_{k+1}) &=& H(S_k|S_{k+1}) - \underbrace{H(S_k|T_k, S_{k+1})}_{0} \\ &>& 0, \end{array}$$

gdzie $H(S_k|S_{k+1}) > 0$, bo $S_k$ nie jest funkcją $S_{k+1}$ (nawet dla $k = 1$). Zatem

$$I(T_k; S_{k+1}) = \underbrace{I(T_k; S_{k+1} \mid S_k)}_{0} + R(T_k; S_k; S_{k+1}) < \underbrace{I(T_k; S_k \mid S_{k+1})}_{>0} + R(T_k; S_k; S_{k+1}) = I(T_k; S_k).$$

Alternatywnie, możemy zauważyć[4], że powyższa uwaga o rozkładzie warunkowym implikuje, że

$$I(T_k; S_{k+1}) = H(S_{k+1}) - \underbrace{H(S_{k+1} \mid T_k)}_{1},$$

ale także

$$\begin{aligned} I(S_k; S_{k+1}) &= H(S_{k+1}) - \underbrace{H(S_{k+1} \mid S_k)}_{1} \\ &= H(S_k) - H(S_k \mid S_{k+1}), \end{aligned}$$

gdzie $H(S_k \mid S_{k+1}) > 0$, co również daje żądaną nierówność. Jeszcze innym sposobem jest obliczenie rozważanych wartości i ich porównanie (z wykorzystaniem własności symbolu Newtona).

W konkluzji, $I(T_1; S_2) < I(T_1; S_1) = I(T_1; T_2)$, oraz

$$I(T_k; S_{k+1}) < I(T_k; S_k) < I(T_k; T_{k+1}),$$

dla $k \geq 2$.

*dn*

[4] To chyba najelegantsze rozwiązanie podał pan Jakub Jasiulewicz.