Zadanie egzaminacyjne nr 1

Na pewnej spośród n pozycji znajduje się pierścień. Pozycję pierścienia opisuje zmienna losowa $\mathbf{X}$ o wartościach w zbiorze $\{1, 2, \ldots, n\}$ i rozkładzie jednostajnym.

Zakładamy, że n jest podzielne przez 6 i dlatego zbiór pozycji możemy podzielić na

- dwie **połówki**:$\{1, \ldots, \frac{n}{2}\}$, $\{\frac{n}{2} + 1, \ldots, n\}$, lub
- trzy równe **części trzecie** $\{1, \ldots, \frac{n}{3}\}$, $\{\frac{n}{3} + 1, \ldots, \frac{2n}{3}\}$, $\{\frac{2n}{3} + 1, \ldots, n\}$.

Hobbit poszukuje pierścienia i może spytać o radę dwóch trolli, z których pierwszy wskazuje, w której połówce odcinka, a drugi w której jednej trzeciej części odcinka znajduje się pierścień.

Problem w tym, że pierwszy troll **co trzeci raz** kłamie tzn. mówi prawdę jedynie z prawdopodobieństwem $\frac{2}{3}$, przy czym decyzja o kłamstwie jest niezależna od położenia pierścienia. Natomiast drugi troll kłamie podobnie, ale **co drugi raz** i w takim wypadku daje jedną z dwóch złych odpowiedzi z jednostajnym prawdopodobieństwem. Odpowiedzi pierwszego trolla opisuje zmienna $\mathbf{T_1}$ (jej dwiema możliwymi wartościami są więc połówki odcinka), a odpowiedzi drugiego trolla opisuje analogicznie zmienna $\mathbf{T_2}$.

Którego z trolli bardziej opłaca się zapytać; innymi słowy, która z informacji wzajemnych $\mathbf{I(X; T_i)}$ jest większa ?

Jak zmieni się odpowiedź, jeśli opuścimy założenie, że zła odpowiedź drugiego trolla wybierana jest z jednostajnym prawdopodobieństwem?

Uwaga. Można skorzystać z szacowania

$$1.584962500 < \log_2 3 < 1.584962501$$

Komentarz do zadania 1

Jak większość autorów poprawnie wykazała, drugi troll – który daje dokładniejszą informację, choć częściej kłamie – okazuje się lepszym informatorem (choć nieznacznie).

Natomiast pytanie dodatkowe

> Jak zmieni się odpowiedź, jeśli opuścimy założenie, że zła odpowiedź drugiego trolla wybierana jest z jednostajnym prawdopodobieństwem?

dało pole do różnych interpretacji.

Intuicyjnie, jeśli troll (odtąd zajmujemy się tylko drugim) różnicuje prawdopodobieństwo swoich kłamliwych odpowiedzi, może przekazać dodatkową informację i stać się jeszcze lepszym informatorem.

Istotnie, entropia warunkowa $H(T_2 \mid X)$ może się wtedy jedynie zmniejszyć. Nie rozstrzyga to jednak jeszcze pytania o zmianę $I(T_2; X)$, ponieważ entropia $H(T_2)$ może się zmniejszyć także.

Dla analizy problemu przyjmijmy, że troll wybiera kłamliwą odpowiedź rzucając monetą (zapewne fałszywą) z pradopodobieństwami wyników p i $q = 1 - p$. Nie determinuje to jednak jeszcze do końca sytuacji, dopóki nie wskażemy *który* wynik rzutu prowadzi do *której* z dwóch złych odpowiedzi.

Na przykład, jeśli „zawiniemy" odcinek w cykl i określimy strategię trolla: z pradopodobieństwem p podaj odcinek późniejszy, a z pradopodobieństwem q odcinek wcześniejszy, to rozkład zmiennej T_2 się nie zmieni i wobec uwagi powyżej widzimy od razu, że informacja wzajemna między X i T_2 się zwiększyła.

Jeśli jednak wybierzemy inną strategię: z pradopodobieństwem p podaj zły odcinek bardziej na lewo, a z pradopodobieństwem q bardziej na prawo, to rozkład i w konsekwencji entropia T_2 będą inne (dla $p \neq q$).

Z uwagi na symetrię, nietrudno jest zobaczyć, że w modelu z jedną monetą są to jedyne dwie możliwości (z punktu widzenia rozkładu T_2).

W jeszcze ogólniejszym scenariuszu rozważonym w jednej z prac troll – pozostając przy naszej metaforze – wykorzystuje trzy różne monety (ze swoimi prawdopodobieństwami) w zależności od tego, w której trzeciej części odcinka znajduje się pierścień. Jak wynika z rachunków, w obu przypadkach informacja $I(T_2; X)$ rośnie, potwierdzając naszą intuicję.

Pozostał do rozważenia najogólniejszy wariant, w którym przy każdym położeniu pierścienia moneta może być inna. Oczywiście dla $n = 3$ ta sytuacja pokrywa się z przypadkiem opisanym przed chwilą; nazwijmy ten przypadek **(*)**, bo z niego za chwilę skorzystamy. Ogólnie, jeśli przyjmiemy oznaczenie

$$\delta(X) \quad = \quad \left\lceil \frac{3X}{n} \right\rceil,$$

to założenie o drugim trollu sprowadza się do jednej równości:

$$\Pr(T_2 = \delta(X)) \quad = \quad \frac{1}{2}.$$

Jednak z nierówności teorio-informacyjnej *data processing inequality* wiemy, że

$$I(\delta(X); T_2) \quad \leq \quad I(X; T_2),$$

dlatego wystarczy oszacować z dołu $I(\delta(X); T_2)$, a to sprowadza się do przypadku **(*)**.

Zadanie egzaminacyjne nr 2

W poniższym zadaniu ustalamy uniwersalną (jednotaśmową) maszynę Turinga U i stwierdzenie *słowo* $w \in \{0,1\}^*$ *jest słowem losowym* oznacza, że

$$C_U(w) \quad \geq \quad |w|.$$

Rozstrzygnąć, czy następujące zdania są prawdziwe.

1. Dla prawie wszystkich n i dla każdych słów $x, y, z \in \{0,1\}^n$, jeśli z jest słowem losowym i $x \oplus y = z$, to przynajmniej jedno ze słów x, y jest losowe[1].
2. Dla prawie wszystkich n i dla każdego słowa $x \in \{0,1\}^n$, jeśli słowo x jest losowe, to każde jego przesunięcie cykliczne też jest losowe[2].
3. Dla każdego $k \in \mathbb{N}$, istnieje słowo losowe zawierające 0^k jako podsłowo.

Zadanie 3

Niech S_k oznacza zbiór wszystkich permutacji zbioru $[k] = \{1, 2, \ldots, k\}$, przy czym $k \geq 2$. Dla $n \geq 2$ rozważamy kanał $\Gamma_{k,n}$, którego alfabet wejściowy i wyjściowy stanowią słowa nad alfabetem $[k]$ o długości n, a którego zasada działania jest następująca. Dla wejścia $w \in [k]^n$, wybieramy losowo (z jednostajnym prawdopodobieństwem) permutację $f \in S_k$ i stosujemy ją do słowa w litera po literze (tzn. dla wejścia $w = w_1 \ldots w_n$ wyjściem jest słowo $f(w_1) \ldots f(w_n)$).

Rozważmy również kanał $\overline{\Gamma}_{k,n}$ nad tym samym alfabetem zdefiniowany analogicznie, ale w miejsce S_k wstawiamy $[k]^{[k]}$, tzn. zbiór wszystkich funkcji z $[k]$ w $[k]$.

Porównaj pojemności kanałów $\Gamma_{k,2}$ i $\overline{\Gamma}_{k,2}$.

[1] Gdzie $\oplus$ oznacza XOR po współrzędnych, np. $110 \oplus 011 = 101$.

[2] Przesunięciem cyklicznym słowa $w_1 \ldots w_n$ jest każde słowo $w_{i+1} \ldots w_n\, w_1 \ldots w_i$, gdzie $i = 0, 1, \ldots, n-1$.

Zaprojektuj kod $C \subseteq [k]^n$ pozwalający przesyłać przez kanał $\Gamma_{k,n}$ wiadomości z zerowym prawdopodobieństwem błędu; im większy kod, tym lepiej. W tym poleceniu wolno założyć, że n jest dużo większe od k.

Czy jest możliwe zaprojektowanie kodu o podobnej własności dla kanału $\overline{\Gamma}_{k,n}$?

Bonus. Porównaj pojemności kanałów $\Gamma_{k,n}$ i $\overline{\Gamma}_{k,n}$ dla $n \geq 2$.

dn