Matematyka dyskretna, WMat 2024

Lista 1: Prawdopodobieństwo i entropia

1. Mieszkańcy wyspy mówią prawdę z prawdopodobieństwem 1/3, w pozostałych przypadkach kłamią. Pytasz jednego z nich o drogę, po czym dla pewności pytasz również stojącą obok osobę, czy wskazówki, które padły, były prawdziwe. Pada odpowiedź „tak”. Jaka jest szansa, że wskazówki były istotnie prawdziwe?

2. Obserwujemy strumień bitów, w którym z prawdopodobieństwem 1/3 każdy kolejny bajt to x = 10101010, y = 01111101 lub z = 11100111, przykładowa sekwencja xyxxzyzxyyxzz... Zaczynamy obserwować strumień asynchronicznie, tj. nie mamy informacji, ile bitów wystąpiło przed zaczęciem ich sczytywania.

(a) Jakie jest prawdopodobieństwo, że w pierwszym obserowanym bicie wypadnie 1, a jakie, że 0?

(b) Podaj algorytm, który sczytując kolejne bity pozwala rozpoznać dalszą sekwencję bajtów x, y lub z. Ile najmniej początkowych bitów trzeba zaobserwować, aby jednoznacznie rozpoznać pierwszy bajt?

(c) Przelicz podpunkt (a), ale z modyfikacją, że kolejno w sekwencji występują x', y,z z prawdopodobieństwami 1/3, gdzie x' to podwojony x, x' = xx.

3. Zaproponuj metodę kompresji strumienia bitów, którego przykładowy fragment jest postaci 00000000000000000000000000000000000000000101000000000000010000000000000001000000000000010000000000000000000000000001000000000000000010000000000000000000010000000000000000000000000000...

4. Udowodnij, że możemy alternatywnie zdefiniować zbiór przeliczalny jako zbiór, którego każdemu elementowi możemy nadać jednoznaczną etykietę będącą skończonym słowem złożonym z symboli dowolnego alfabetu.

5. Niech $0 \leq p \leq 1$ będzie prawdopodobieństwem i $q = 1 - p$, $a = \ln(p/q)$. Jakim wzorem wyraża się p jako funkcja a? Naszkicuj jej wykres. Co kiedy $\tilde{a} = \log_2(p/q)$?

6. Niech X oraz Y będą zmiennymi losowymi o wartościach bitowych. Pokaż, że

$$\log \frac{\mathrm{P}(X = \mathtt{1}|Y)}{\mathrm{P}(X = \mathtt{0}|Y)} = \log \frac{\mathrm{P}(Y|X = \mathtt{1})}{\mathrm{P}(Y|X = \mathtt{0})} + \log \frac{\mathrm{P}(X = \mathtt{1})}{\mathrm{P}(X = \mathtt{0})}$$

7. Rozpatrzmy 4 losowe bity (każdy ma prawd 1/2 na 0 i 1/2 na 1) i zmienną X, która zawiera wszystkie bity z tych 4 po pierwszej 1, np. 0110 daje $X = \mathtt{10}$. Zmienna X może być pusta, np. w przypadku 0000. Jaka jest jej entropia?

8. Rozpatrzmy zestaw 52 kart, z których jedną losowo wyrzuciliśmy, po czym zastąpiliśmy inną losowo wybraną z drugiego zestawu. Jaką entropię ma zestaw po tej operacji?

9. Niech X będzie zmienną losową zawierającą 2 bity. Niech każdy bit będzie losowany niezależnie i z prawdopodobieństwem p, że wypadnie `1`. Oblicz entropię tej zmiennej.

10. Udowodnij własność dekompozycji entropii. Dla dowolnego rozkładu $\boldsymbol{p} = [p_1, p_2, \ldots, p_n]$ zachodzi, że:

$$H(\boldsymbol{p}) = H(p_1, 1 - p_1) + (1 - p_1)H\left(\frac{p_2}{1 - p_1}, \ldots, \frac{p_n}{1 - p_1}\right).$$

Notacja: $H(p_1, p_2, \ldots)$ to entropia rozkładu $p_1, p_2, \ldots$.

11. Niech w strumieniu wartości `1` i `0` występują z prawdopodobieństwami p i $1 - p$. Zaczynamy obserwować strumień i liczymy, ile `1` zaobserwowaliśmy do pierwszego `0`, np. `11110` daje 4. Oblicz entropię tej wartości.

12. Jaka jest minimalna wartość entropii dla zmiennej o skończonej ilości wartości? Jakie rozkłady $\boldsymbol{p}$ go realizują?

13. Jaka jest relacja między entropią $H(X)$ a $H(f(X))$, gdzie f jest dowolną funkcją X?