

Programmēšanas vingrinājumi

Lietiškie algoritmi, 2019.g. rudens

Terminš: 2019-10-22

1.uzdevums. Diskrēts gadījumielums. M.Bendiks reizēm pērk Rūjienas saldējumu ar rozīnēm (<https://bit.ly/2001xdn>). Pieņemsim, ka rozīņu skaits vienā saldējumā pakļaujas Puasona sadalījumam ar vidējo vērtību $\lambda = 5$. Aprēķinu vienkāršošanai var pieņemt, ka rozīņu skaits nevienā saldējumā nepārsniedz 20 (t.i. gadījumielums $X \in \{0, 1, \dots, 20\}$); saldējumi, kuros $X > 20$, pārdošanā nenonāca.

Atrasto rozīņu skaitus M.Bendiks vēlas nosūtīt, izmantojot Hafmana kodus.

1. Cik garš ir visīsākais Hafmana kods; kuram rozīņu skaitam tas atbilst?
2. Cik garš Hafmana kods atbilst vērtībai $X = 0$?
3. Cik garš Hafmana kods atbilst vērtībai $X = 20$?
4. Kāda ir saldējumā atrodamā rozīņu skaita entropija jeb caurmēra ziņojumā esošās informācijas saspiežamības robeža? (*Noapaļot rezultātu līdz 3 cipariem aiz komata.*)

2.uzdevums. Aritmētiskais kods. Aritmētiskais kods definēts, sākot ar 16.lappusi G.Bleloka konspektā: <https://www.cs.cmu.edu/~guyb/realworld/compression.pdf>. Dots apriori varbūtību sadalījums sekojošiem ziņojuma alfabēta burtiem: $\{(A, 6/10), (B, 2/10), (C, 1/10), (\$, 1/10)\}$. Izveidot aritmētisko kodu sekojošai 8 ziņojumu virknei: **ABABACA\$**. Zināms, ka ziņojums beidzas ar "end-of-text" simbolu (dolāru), kam piedēvēta varbūtība $1/10$ (un nekur ziņojuma iekšienē šāda simbola nav). Kodēšana beidzas ar intervālu $[\ell_8, \ell_8 + s_8)$. Aritmētiskais kods ir īsākais skaitļa c binārais pieraksts ar k (binārajiem) cipariem aiz komata, kuram $[c, c+2^{-k})$ ir intervāls, kurš pilnībā ietilpst intervālā $[\ell_8, \ell_8 + s_8)$.

3.uzdevums. Move to Front kods: Pieņemsim, ka visi 26 mazie latīņu alfabēta burti sākotnēji sakārtoti masīvā alfabētiskā secībā ($A[0] = a, \dots, A[25] = z$). Izveidot "Move to Front" kodējumu sekojošam vārdam: **interchangeableness**

4.uzdevums. Markova ķēdes arhivācija: Dota Markova ķēde no 6 stāvokļiem - sk. zīmējumu <https://bit.ly/2LDJqo8>. Izveidot ar šo Markova ķēdi simbolu virknīti tieši miljons ciparu garumā (katru ciparu attēlot kā ASCII simbolu "1", "2", "3", "4", "5", "6"). Arhivēt šo simbolu virkni ar WinZIP (vai Linux "gz") arhivatoru, lai iegūtu LZ78-veida arhīva garuma novērtējumu.

Saskaitīt, cik šajā virknē ir katra veida ciparu un izrēķināt entropiju (un teorētiski īsāko sasiestās virknes garumu). Salīdzināt, cik tuvu ir ZIP faila izmērs teorētiskajam minimumam.

5.uzdevums. Diskrētais kosinusu pārveidojums:

$$X_k = \sum_{n=0}^{N-1} x_n \cos \left[\frac{\pi}{N} \left(n + \frac{1}{2} \right) k \right] \quad k = 0, \dots, N-1.$$

Diskrēto kosinusu pārveidojumu uzdod augšminētā formula, kur $N = 100$. Iedomāsimies, ka x_n ir virkne, kas definēta ar vienādojumu $x_n = \sin x$, kur $n = 0, \dots, 99$. Atrast, kuri ir trīs lielākie no DCT koeficientiem X_k , un kurām k vērtībām tie atbilst.