Úvod do kódovania – Úloha č. 2

Termn odovzdania 22. apríl 2025, vo formáte pdf cez *Google Classroom*

(R) bude znamenať odkaz na knižku S. Romana, (G) P. Garretta a (H) R. Hilla; uvedené bude spravidla číslo príkladu/problému (kapitola.sekcia.problém), alebo číslo strany, kde sa o danej veci píše. Neočakávajú sa úplne kompletné a perfektné riešenia. Aj čiastkové riešenia s drobnými opomenutiami, logickými medzerami, či neporiadnym zápisom si môžu vyžadovať veľa práce a námahy a môžu dostať plný počet bodov. T.j. nemusíte týmito domácimi úlohami stráviť všetok svoj voľný čas počas nadchádzajúceho semestra. Na druhej strane, očakáva sa preukázanie výraznejších snáh, aby domáce úlohy splnili svoj účel – naučiť sa, resp. samostatne objaviť niečo netriviálne z preberaného materiálu.

Vždy je tu možnosť absolvovania konzultácií. Tiež môže pomôcť preskúmanie viacerých príkladov daného fenoménu pri hľadaní dôkazu všeobecného tvrdenia. Akceptované budú aj riešenia založené na počítačových simuláciách, pokiaľ budú primerane zdôvodnené a bude to v danom kontexte dávať zmysel.

Domáca úloha bude obsahovať príklady s celkovým ohodnotením prevyšujúcim 50 bodov, čo je maximum, ktoré sa v rámci jednej úlohy dá získať. To znamená, že si môžete zvoliť, ktorým príkladom sa budete venovať a ktoré nakoniec odovzdáte. Keďže sa dá očakávať, že nie všetky riešenia budú za plný počet bodov, má zmysel odovzdať príklady, ktorých celkové hodnotenie prevyšuje 50 bodov.

1. (10 bodov) Nech $\mathcal{A}$ je r-árna abeceda a $n \geq 1$. Nájdite jednoznačne rozkódovateľný kód $\mathcal{C}$ obsahujúci n slov z $\mathcal{A}^*$, ktorý má najmenšiu celkovú dĺžku spomedzi všetkých jednoznačne rozkódovateľných kódov obsahujúcich n slov z $\mathcal{A}^*$. Svoju odpoveď zdôvodnite.

2. (10 bodov)

(a) Nájdite podmienku pre pravdepodobnostnú distribúciu $(p_1, p_2, \dots, p_n)$, ktorá zaručí jednoznačnosť priebehu vytvárania Huffmanovho kódu (t.j. jeho jednoznačnosť až na zámenu premenovania $0 \to 1$ a $1 \to 0$)

(b) Nájdite formulu pre počet redukcií v Huffmanovom algoritme pre kód veľkosti n a základ r.

(c) Dokážte alebo vyvráťte: Nech $n = 2^k$ a $\mathcal{C}$ je binárny Huffmanov kód pre rovnomernú pravdepodobnostnú distribúciu $(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$ s dĺžkami kódových slov $\ell_1, \ell_2, \dots, \ell_n$. Potom platí

$$\sum_{i=1}^{n} \frac{1}{r^{\ell_i}} = 1.$$

3. (10 bodov) Nech (p_1, p_2, p_3, p_4) a (q_1, q_2, q_3, q_4) sú dve pravdepodobnostné distribúcie a $W(p_1, p_2, p_3, p_4)$ a $W(q_1, q_2, q_3, q_4)$ označujú priemerné dĺžky optimálnych jednoznačne rozkódovateľných kódov pre tieto distribúcie. Nájdite pravidlo alebo popíšte algoritmus, ktorý rozhodne ktorá z hodnôt bude väčšia/menšia bez toho, aby sa skonštruovali optimálne kódy. Závisí odpoveď od veľkosti Σ? Bude nerovnosť medzi $W(p_1, p_2, p_3, p_4)$ a $W(q_1, q_2, q_3, q_4)$ rovnaká ako medzi ich entropiami $H(p_1, p_2, p_3, p_4)$ a $H(q_1, q_2, q_3, q_4)$?

4. (15 bodov)

a) Aký je maximálny počet kódových slov v okamžite rozkódovateľnom r-árnom kóde s maximálnou dĺžkou slova k?

b) Dokážte/vyvráťte, že Huffmanov kód $\mathcal{C}$ vždy obsahuje slová všetkých dĺžok $1, 2, \dots, \ell$, kde ℓ je maximálna dĺžka slova v $\mathcal{C}$.

c) Nech $\mathcal{C}$ je Huffmanov kód zložený z n slov nad r-árnou abecedou $\mathcal{A}$, ℓ je dĺžka najdlhšieho slova v $\mathcal{C}$ a α_i označuje počet slov dĺžky i v $\mathcal{C}$ pre $1 \leq i \leq \ell$. Sú čísla α_i jednoznačne určené? Svoju odpoveď dokážte. Dajú sa vysloviť nejaké všeobecné tvrdenia o číslach α_i?

5. (15 bodov)

a) Predpokladajme, že $\mathcal{C}$ je optimálny kód pre zdroj $\mathcal{S}$ s pravdepodobnostnou distribúciou $(p_1, p_2, \dots, p_n)$ nad binárnou abecedou $\Sigma = \{0, 1\}$ získaný pomocou Huffmanovho algoritmu. Dá sa kód $\mathcal{C}$ použiť ako základ pre vytvorenie optimálneho kódu nad ternárnou abecedou? Skúste zodpovedať otázku a v prípade kladnej odpovede navrhnite algoritmus (+ zdôvodnite jeho korektnosť). Ak rozhodnete, že z optimálneho binárneho kódu $\mathcal{C}$ sa nedá získať optimálny ternárny kód žiadnym rozumným spôsobom, zdôvodnite svoju odpoveď.

b) Predpokladajme, že $\mathcal{C}$ je optimálny kód pre zdroj $\mathcal{S}$ s pravdepodobnostnou distribúciou $(p_1, p_2, \ldots, p_n)$ nad r-árnou abecedou Σ získaný pomocou Huffmanovho algoritmu. Dá sa kód $\mathcal{C}$ použiť ako základ pre vytvorenie optimálneho kódu nad abecedou Σ' veľkosti $r' > r$? Skúste zodpovedať otázku a v prípade kladnej odpovede navrhnite algoritmus (+ zdôvodnite jeho korektnosť). Ak rozhodnete, že z optimálneho r-árneho kódu $\mathcal{C}$ sa nedá získať optimálny r'-árny kód žiadnym rozumným spôsobom, zdôvodnite svoju odpoveď.

6. (15 bodov)

a) Predpokladajme, že $\mathcal{C}$ je optimálny kód pre zdroj $\mathcal{S}$ s pravdepodobnostnou distribúciou $(p_1, p_2, \ldots, p_n)$ nad binárnou abecedou $\Sigma = \{0, 1\}$ získaný pomocou Huffmanovho algoritmu. Existuje spôsob ako použiť kód $\mathcal{C}$ ako základ pre vytvorenie optimálneho kódu pre rozšírený zdroj $\mathcal{S}^2$? Skúste zodpovedať otázku a v prípade kladnej odpovede navrhnite algoritmus (+ zdôvodnite jeho korektnosť). Ak rozhodnete, že z optimálneho kódu $\mathcal{C}$ pre zdroj $\mathcal{S}$ sa nedá získať optimálny kód pre zdroj $\mathcal{S}^2$ žiadnym rozumným spôsobom, zdôvodnite svoju odpoveď.

b) Predpokladajme, že $\mathcal{C}$ je optimálny kód pre zdroj $\mathcal{S}$ s pravdepodobnostnou distribúciou $(p_1, p_2, \ldots, p_n)$ nad binárnou abecedou $\Sigma = \{0, 1\}$ získaný pomocou Huffmanovho algoritmu. Existuje spôsob ako použiť kód $\mathcal{C}$ ako základ pre vytvorenie optimálneho kódu pre n-krát rozšírený zdroj $\mathcal{S}^n$, $n \geq 2$? Skúste zodpovedať otázku a v prípade kladnej odpovede navrhnite algoritmus (+ zdôvodnite jeho korektnosť). Ak rozhodnete, že z optimálneho kódu $\mathcal{C}$ pre zdroj $\mathcal{S}$ sa nedá získať optimálny kód pre zdroj $\mathcal{S}^n$ žiadnym rozumným spôsobom, zdôvodnite svoju odpoveď.

7. (10 bodov) Úloha č. 2.3.7 z Romanovej knihy (str. 66).

8. (10 bodov) Úloha č. 3.1.6 z Romanovej knihy (str. 79).

9. (10 bodov) Úloha č. 3.2.4 z Romanovej knihy (str. 86).

10. (10 bodov) Úloha č. 3.2.7 z Romanovej knihy (str. 87).

11. (10 bodov) Úloha č. 3.2.8 z Romanovej knihy (str. 87).

12. (15 bodov) Úloha č. 3.2.16 z Romanovej knihy (str. 88).