

Cvičení

20. dubna 2020

Problém 1 (Podle S. Rotman: Introduction to Coding and Information Theory). Mějme informační zdroj nad abecedou $\Sigma = \{a, b, c, d, e\}$ s pravděpodobnostmi jednotlivých písmen postupně $(1/5, 1/5, 3/10, 1/10, 1/5)$. Které kódovací schéma nám dá v průměru kratší kódové slovo?

a) $a \rightarrow 0, b \rightarrow 10, c \rightarrow 110, d \rightarrow 1110, e \rightarrow 11110$ nebo

b) $a \rightarrow 000, b \rightarrow 001, c \rightarrow 010, d \rightarrow 011, e \rightarrow 100$?

Problém 2. Na přednášce jsme našli Huffmanovo kódování C pro zdroj příkazů Alíkovi. Jaká bude průměrný počet bitů potřebný na odeslání jednoho znaku pokud použijeme C , ale pravděpodobnosti jednotlivých příkazů Alíkovi se změní všechny na $1/8$? Jaké kódování by bylo v tomto případě lepší zvolit?

Problém 3. Dokažte, že entropie informačního zdroje nad n -prvkovou abecedou Σ , který vysílá každý znak náhodně nezávisle s rovnoměrným rozdělením nad Σ , je $\log_2 n$.

Problém 4. Mějme informační zdroj, který vysílá s pravděpodobností p znak 0 a s pravděpodobností $1 - p$ znak 1. Dokažte, že entropie tohoto zdroje je maximální pro $p = 1/2$.

Problém 5. Dokažte, že v Huffmanovu kódování není žádné kódové slovo předponou jiného kódového slova (tj. neexistuje $x_1x_2 \dots x_n \in C$ takové, že by existovalo $y_1y_2 \dots y_m$, že $m > n$ a $x_i = y_i$ pro $i \leq n$).