NTIN 100 Základy přenosu a zpracování informace **zima 2021/2022**

2. domácí úlohy - Kódování

do 4. dubna 2022

Úloha 1. Pro každý z následujících kódů určete, zda je jednoznačně dekódovatalný. Pokud ano, nalezněte nekonečnou posloupnost, existuje-li, kterou lze dekódovat dvěma způsoby. Ukažte, že taková posloupnost neexistuje pro bezprefixové kódy.

a) $C_1 = \{0, 10, 11\}$, $C_2 = \{0, 01, 11\}$

b) $C_3 = \{0, 01, 10\}$, $C_4 = \{0, 01\}$

c) $C_5 = \{110, 11, 10\}$, $C_6 = \{110, 11, 100, 00, 10\}$

Úloha 2.

a) Ukažte, že pro každé pravděpodobnostní rozdělení $p_1, p_2, \ldots, p_n$ existuje bezprefixový kód s délkami slov $\ell_1, \ell_2, \ldots, \ell_n$, který má nejlepší možnou průměrnou délku kódového slova a splňuje následující tři vlastnosti:

1. Pokud $p_i \leq p_j$, pak $\ell_i \geq \ell_j$.
2. Dvě nejdelší slova mají stejnou délku.
3. Dvě nejmenší pravděpodobnosti mají přiřazena dvě nejdelší kódová slova, která se liší pouze v posledním bitu.

b) Ukažte, že Hammingův kód má nejmenší možnou průměrnou délku slova.

Úloha 3. Nalezněte náhodnou proměnnou X a nějaký kód $C : X \to \{0,1\}^*$ takový, že $E[|C(X)|] < H(X)$. Pro libovolnou náhodnou proměnnou X a její kód C nalezněte co nejlepší dolní odhad na $E[|C(X)|]$ v závislosti na $H(X)$.

Úloha 4. Mějme prvky $x_1 < \ldots < x_n$ a k nim příslušející pravděpodobnosti $p_1, p_2, \ldots, p_n$. Uvažujme statický, tedy neměnný, binární vyhledávací strom pro tyto prvky. Ukažte dolní mez ve formě entropie na průměrnou dobu vyhledání prvku v tomto stromu, kde prvek x_i je vyhledávaný s pravděpodobností p_i. Jak se odpověď změní pro dynamický vyhledávací strom, tedy strom, který se může po každém dotazu přeuspořádat.