

Kódování

1. Existuje binární prefixové kódování číslic 0, 1, 2, ..., 9 takové, že číslice 9 je kódována nulou a ostatní číslice mají kód délky nejvýše pět?

Existenci kódu ověříme tak, že zjistíme, zda je splněna Kraftova nerovnost

$$n^{-d_1} + n^{-d_2} + \dots + n^{-d_{|K|}} \leq 1,$$

kde n je počet znaků kódové abecedy a d_i jsou délky kódových značek.

2. Vytvořte tříznakový prefixový kód, který bude kódovat abecedu {A, B, C, D, E, F}. Dvě písmena budou kódována značkou délky 1, čtyři písmena značkou délky 2. Alternativně vytvořte kód, kde budou dvě písmena kódována značkou délky 2, dvě písmena značkou délky 2 a dvě písmena značkou délky 3.

Než se pustíme do vytváření kódu, ověříme Kraftovu nerovnost. Zjistili bychom, že pro dvě písmena s kódovou značkou délky 1 a čtyři písmena s kódovou značkou délky 2 kód neexistuje. Uděláme tedy kód pro druhou alternativu, pro kterou je Kraftova nerovnost splněna.

Vezmeme např. první dvě písmena vstupní abecedy a přiřadíme jim kódy – první a druhý znak z kódové abecedy. Kódy dalších dvou písmen musí začínat třetím znakem kódové abecedy, aby kódy prvních dvou znaků nebyly jejich prefixem. Na druhou pozici (jsou to značky délky 2) dáme opět první, respektive druhý znak kódové abecedy. Kódy zbylých dvou písmen (kódové značky délky 3) musí mít na první a druhé pozici třetí znak kódové abecedy. Na třetí pozici dáme opět první, respektive druhý znak kódové abecedy.

3. Je dána abeceda {A, B, C, D, E, F} s pravděpodobnostmi výskytu znaků {2/8, 1/8, 1/8, 1/8, 2/8, 1/8}. Kódová abeceda je {0, 1, 2, 3}. Najděte kód s minimální střední délkou značky.

Obecný postup pro vytvoření kódu s minimální střední délkou značky je následující:

1. Seřadíme znaky vstupní abecedy podle pravděpodobnosti.
2. Rozdělíme znaky do skupin po $n-1$, kde n je počet znaků kódové abecedy.
3. Sloučíme poslední skupinu. Pravděpodobnost výskytu některého znaku ze skupiny je rovna součtu pravděpodobností znaků ve skupině. Sloučenou skupinu zařadíme podle pravděpodobnosti jako nový prvek mezi zbývajícím znakem.
4. Dále slučujeme (už bez ohledu na skupiny) po n znacích.

Slučování provádíme zdola nahoru. Mezi sloučené prvky pak shora dolů distribuujeme znaky kódové abecedy a vytváříme kódové značky.

