

Huffmanové kódovanie

Projekt z premetu KKO 2016/2017

František Matečný, xmatec00

Adaptívne Huffmanové kódovanie konvertuje vstupné znaky na bitový reťazec rôznej dĺžky. Znak, ktorý sa v stupnom toku nachádza najčastejšie, sú kódované menším počtom bitov, čo spôsobuje kompresiu výstupných dát.

Na začiatku je vytvorený koreň stromu, ktorý obsahuje ESC kód. Následne sa postupne načítavajú vstupné znaky a ukladajú sa do stromu tak, že v mieste, kde bol list s ESC kódom, sa vytvorí nový podstrom, a do jeho ľavého syna sa vloží ESC kód a kód 0, a do pravého syna načítaný znak a kód 1.

Kódovanie

Funkcia *AHEDEncoding* na začiatku inicializuje strom t. j. vytvorí koreň stromu. Následne sa postupne načítavajú jednotlivé znaky zo vstupného súboru.

- Ak sa znak nenachádza v strome do výstupného súboru sa zapíše ESC kód a načítaný vstupný symbol. V strome sa vytvorí podstrom, kde ľavý syn obsahuje ESC kód a pravý syn obsahuje načítaný symbol. Potom sa strom aktualizuje.
- Ak sa znak nachádza v strome, načíta sa kód znaku postupným prechádzaním cez uzly od listu s daným znakom až ku korenu. Načítaný kód sa uloží v reverznom poradí, keďže pri dekódovaní sa prehľadáva strom od korena k listom. Potom sa strom aktualizuje.

Dekódovanie

Funkcia *AHEDDecoding* načítavá kódovaný súbor po 1 bajte a následne po podľa jednotlivých bitov postupuje v strome a hľadá znaky odpovedajúce vstupnému kódu.

- Ak sa na vstupe nachádza ESC kód, nasledujúcich 8bitov prdsavuje nekódovaný znak, ktorý sa zapíše na výstup a uloží sa do stromu.
- Ak sa na vstupe nenachádza ESC kód, vstupné bity udávajú kód znaku v strome. Ten sa vyhľadá, zvýši sa je početnosť a zapíše sa na výstup. Koniec súboru je detekovaný pomocou kódu ESC a EOF.