A note on Huffman's algorithm

Christian Wulff-Nilsen [*]

1 Optimal substructure

Here is the alternative proof of optimal substructure for Huffman's algorithm presented during the Greedy Algorithms lecture. At the oral exam, you are free to choose between this proof and that in the course book CLRS.

We start with a lemma.

Lemma 1. *Let T be a parse tree for an alphabet C (with associated frequencies f_x for each $x \in C$) and let W be the set of vertices of T except the root. Then $B(T) = \sum_{x \in W} f_x$.*

Proof. Let $e = (u, v)$ be an edge of T where v is a child of u. The bit associated with e (0 if v is the left child of u and 1 otherwise) occurs f_v times in the compressed string since f_v is the sum of frequencies of leaves of the subtree of T rooted at v. Summing f_v over all edges (u, v) of T gives $\sum_{x \in W} f_x$ which is the total length $B(T)$ of the compressed string. □

Showing optimal substructure for a greedy algorithm amounts to showing that, if the greedy choice x belongs to an optimal solution for a problem, this optimal solution consists of x and an optimal solution to the subproblem generated.

To show optimal substructure for Huffman, let T_1 be an optimal tree for alphabet C and assume it contains the greedy choice, i.e., assume that symbols x and y of minimal frequencies are siblings in T_1. Let z be their parent (whose frequency f_z equals $f_x + f_y$) and let $C' = (C \setminus \{x, y\}) \cup \{z\}$. Let T_1' be T_1 with leaves x and y removed; T_1' is a tree for alphabet C'. Furthermore, let T_2' be an optimal tree for C' and let T_2 be a tree for C obtained from T_2' by adding x and y as children of z.

Showing optimal substructure amounts to proving that T_1' is an optimal tree for C', i.e., that $B(T_1') = B(T_2')$. By the lemma above,

$$B(T_1') = B(T_1) - f_x - f_y,$$
$$B(T_2') = B(T_2) - f_x - f_y.$$

Since T_1' is some tree for C', we have $B(T_1') \geq B(T_2')$. It remains to show $B(T_1') \leq B(T_2')$.

Since T_2 is some tree for C, we have $B(T_1) \leq B(T_2)$. Hence,

$$B(T_1') = B(T_1) - f_x - f_y \leq B(T_2) - f_x - f_y = B(T_2'),$$

as desired.

[*] Department of Computer Science, University of Copenhagen, koolooz@di.ku.dk, http://www.diku.dk/~koolooz/.