



Laborator PS

ALGORITMI DE COMPRESIE

Algoritmul Huffman neadaptiv

Prof. dr. ing. Dan STEFANOIU

As. Ing. Alexandru DUMITRASCU

2. Algoritmul Huffman neadaptiv


- Fazele de compresie si decompresie trec prin aceleasi etape ca in cazul alg. Shannon-Fano.
- Constructia arborelui binar pleaca de la frunze spre radacina → codurile simbolilor sunt construite de la bitul cel mai putin semnificativ la bitul cel mai semnificativ.

Pas 1: Se parcurge secvential D si se construiesc $A^0 \rightarrow N(s)$ pentru fiecare $s \in A^0$

Pas 2: Se ordoneaza simbolii alfabetului: $A^0 = \{s_1, s_2, \dots, s_N\}$ cu $N(s_1) \geq N(s_2) \geq \dots \geq N(s_N)$

Se renoteaza si reindexeaza elementele alfabetului astfel:

$$A^0 = \{t_{N+k-1} = s_k\}_{k \in \overline{1, N}}$$



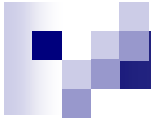
Simbolii $t_N, t_{N+1}, \dots, t_{2N-1}$ sunt utilizati pentru etichetarea frunzelor arborelui binar, iar alfabetul A^0 devine *primul alfabet curent* al algoritmului si va fi renotat prin $A^{0,N}$. Fiecare alfabet curent al algoritmului se va nota prin $A^{0,(N-k)}$, $k=0 \dots N-1$

$k=0 \Rightarrow$ frunzele arborelui sunt etichetate cu simbolii primului alfabet curent

Pas 3: Constructia arborelui binar se realizeaza simultan cu formarea noului alfabet curent.

Noul alfabet curent, $A^{0,(N-k-1)}$ cu $k=0 \dots N-2$, se construiește din vechiul alfabet $A^{0,(N-k)}$, aplicandu-se urmatoarea regula: se elimina 2 dintre simbolii cu cele mai mici contoare din vechiul alfabet curent, inlocuindu-le cu un simbol virtual (de ex. notat prin „ ”), care are contorul egal cu suma contoarelor celor doi simbolii.

Daca exista mai mult de doi simbolii slabi avand acelasi contor, se aleg ultimii doi din alfabet \Rightarrow cardinalul noului alfabet scade cu o unitate fata de cardinalul alfabetului precedent.



$$A^{0,(N-k-1)} = [A^{0,(N-k)} \setminus \{t_p, t_q\}] \cup \{t_{N-k-1}\}$$

$$t_{N-k-1} = \textit{simbol virtual}$$

$$N(t_{N-k-1}) = N(t_p) + N(t_q)$$

$$\# A^{0,(N-k-1)} = \# A^{0,(N-k)} - 1 = N - k - 1$$

Obs.:

- Simbolul virtual eticheteaza un nod intermediar al arborelui, avand ca descendenti cele doua noduri etichetate de simbolii eliminati.
- Fiecare nod intermediar din arbore are cate 2 descendenti, contorul fiului din stanga fiind cel putin egal cu cel al fiului din dreapta.

Pas 4: Se repeta Pas 3 de $N-1$ ori pentru fiecare alfabet curent $A^{0,(N-k)}$ pana cand se eticheteaza si radacina arborelui (t_1).

Pas 5: Se reindexeaza nodurile arborelui astfel incat fiii oricarui nod sa fie etichetati cu indici consecutivi, de la stanga la dreapta, incepand de la radacina.