

LANT – Descrierea soluției

1. Se citește textul și se memorează cuvintele distincte din text într-un tablou c . Fie n_c numărul de cuvinte distincte determinate. Fiecare cuvânt este numerotat de la 0 la $n_c - 1$ (indicii din tabloul c). Observați că numerotarea cuvintelor respectă ordinea primei apariții în text a cuvintelor.

2. Asociem problemei un graf orientat astfel:
- nodurile grafului sunt cuvintele distincte din text;
 - există arc de la nodul i la nodul j ($i < j$) dacă numărul minim de operații insert și delete necesare pentru a transforma cuvântul $c[i]$ în cuvântul $c[j]$ este $\leq k$.

Observați că graful asociat problemei nu conține circuite.

Pentru a determina arcele grafului trebuie să rezolvăm următoarea subproblemă:

să se determine numărul minim de operații delete și insert necesare pentru a transforma cuvântul x în cuvântul y .

Rezolvăm această subproblemă prin programare dinamică.

Fie $d[i][j]$ = numărul minim de operații insert și delete necesare pentru a transforma sufixul lui x care începe la poziția i în sufixul lui y care începe la poziția j .

Fie n =lungimea cuvântului x și m =lungimea cuvântului y .

$$d[n][j] = m - j, \text{ pentru orice } j = 0, m$$
$$d[i][m] = n - i, \text{ pentru orice } i = 0, n$$

```
d[i][j]=min {d[i+1][j+1], daca p[i]==q[j]; - move
              1+d[i][j+1] - insert
              1+d[i+1][j] - delete          }
```

Soluția este $d[0][0]$.

3. Numărul de lanțuri de k -similitudine este egal cu numărul de drumuri care încep cu nodul 0 și se termină într-un nod terminal al grafului (nod cu gradul exterior 0).

Să notăm: $nr[i] = \text{numărul de lanțuri de } k \text{ similitudine care încep cu cuvântul } i.$

Determinăm n_r folosind următoarea relație de recurență:

$nr[i] = 1$, dacă nodul i este terminal

$nr[i] = nr[i_1] + nr[i_2] + \dots + nr[i_k]$, unde i_1, i_2, \dots, i_k sunt noduri din graf cu proprietatea că există arc de la i la i_j , pentru $j=1, k$.