Omogene – descrierea soluției

Prof. Dan Pracsiu, Liceul Teoretic Emil Racoviță Vaslui

Rezolvăm mai întâi problema liniară. Se consideră un vector `t` de lungime n care memorează doar valori din mulțimea {1,2,3}. Să se determine numărul secvențelor omogene (adică în care numărul valorilor de 0 este egal cu numărul valorilor de 1 și egal cu numărul valorilor de 2).

Construim sirul sumelor partiale:
s0[i]= numărul valorilor de 0 din t[1..i]
s1[i]= numărul valorilor de 1 din t[1..i]
s2[i]= numărul valorilor de 2 din t[1..i]
Apoi construim șirurile
d01[0] = d12[0] = d20[0] = 0 și
d01[i] = s0[i] – s1[i], i=1..n
d12[i] = s1[i] – s2[i] , i=1..n
d20[i] = s2[i] – s0[i], i=1..n
Vom determina apoi câte perechi de triplete { d01[i],d12[i], d20[i] } egale sunt în acest șir de lungime n+1 (i=0..n). Dacă două triplete { d01[i],d12[i], d20[i] } și { d01[k],d12[k], d20[k] }(i < j) sunt identice, atunci secvența t[i+1], t[i+2],…t[j] este o secvență omogenă.

Pentru a afla numărul perechilor de triplete egale, se poate sorta șirul tripletelor. În șirul sortat, tripletele identice sunt alăturate. Dacă identificăm X triplete identice, atunci numărul perechilor de triplete identice este X(X-1)/2 (oricare două).

Deci complexitatea determinării numărului de secvențe omogene din vectorul t este dată se sortarea tripletelor, deci O(n log n).

Revenind la problema bidimensională: Din restricții, L<=C și L*C<=65536. De unde rezultă că
L <= sqrt(65536) = 256.
Deci numărul de linii din matrice nu poate fi mai mare de 256.
Construim două matrice de sume parțiale:
z[i][j]=numărul de valori de 0 aflate pe coloana j, începând de la poziția (i,j) în sus.
u[i][j]= numărul de valori de 1 aflate pe coloana j, începând de la poziția (i,j) în sus.

Pentru orice două linii k și p (unde k<=p), dorim să aflăm câte submatrice omogene de înălțime H=p-k+1 există între cele două linii. Vom construi deci vectorul t0[j]=z[p][j]-z[k-1][j], j=1..C, deci în care t[j] memorează numărul valorilor de 0 aflate pe coloana j între liniile k și p. Similar, se construiește t1[j]=u[p][j]-u[k-1][j], j=1..C, deci în care t1[j] memorează numărul valorilor de 1 aflate pe coloana j între liniile k și p. Atunci t2[j]=H-t0[j]-t1[j], pentru j=1..C.
Apoi problema se rezolvă ca în cazul unidimensional.

Complexitatea va fi deci O(L * L * C * log C)