3.26

Considerem el problema d'imprimir de manera polida una frase amb una impressora. El text d'entrada és una seqüència de n mots amb longitud $l_1, l_2, \ldots, l_n$, on cada longitud ve donada en caràcters. Cada línia pot contenir com a màxim M caràcters. Realitzem la impressió de manera que si una línia conté els mots de i fins a j, on $i \leq j$, i deixem exactament un espai entre mots, aleshores el nombre de caràcters en blanc al final de cada línia és $M - j + i - \sum_{k=i}^{j} l_k$, que ha de ser no-negatiu. Volem minimitzar la suma, sobre totes les línies excepte la darrera, dels quadrat d'aquestes magnituds. Dissenyeu un algorisme per imprimir de la manera indicada, un paràgraf amb n mots. Analitzeu les complexitats espacials i temporals del vostre algorisme.

Una solució. Anomenem $B(i,j) = M - j + i - \sum_{k=i}^{j} l_k$ al nombre de caràcters en blanc al final si col·loquem els mots i a j en una línia.

Si $B(i,j) < 0$, els mots no hi caben a una línia, si no el valor ens dona el nombre de espais blancs addicionals.

Suposem que tenim una distribució de cost òptim del paràgraf, mots 1 a n. En cas que el paràgraf càpiga a una línia el cost és zero.

Si no, la primera línia conté els mots 1 fins al i ($1 \leq i \leq n$) i les línies posteriors contenen els mots $i+1$ fins a n, distribuïdes amb cost òptim.

Llavors hem identificat el subproblemes d'interès.

Sigui $C(i)$ = el cost òptim d'un imprimir els mots i a n. Volem obtenir $C(1)$. D'acord amb l'estructura de suboptimalitat tenim la recurrència

$$C(j) = \begin{cases} 0 & \text{si } B(i,n) \geq 0 \\ \min_{i \leq j \leq n, B(i,j) \geq 0} \{C(j+1) + B(i,j)^2\} & \text{altrament} \end{cases}$$

Podem evitar calcular i emmagatzemar els valors $B(i,j)$ precalculant les sumes prefixades de les longituds, $P(i) = \sum_{1 \leq k \leq i} l_i$ en temps $O(n)$. Llavors,

$$B(i,j) = M - i + j - (P(j) - P(i-1)).$$

Tenint en compte que com a molt podem ficar $M/2$ mots a una línia, el rang del min a l'equació es $O(M)$. Això ens dona un cost total de $O(M)$ per subproblema i un cost total de $O(Mn)$. En un cas pràctic podem assumir que la mida de la línia es constant, i per tant l'algorisme té cost $O(n)$.