

Prova pratica del 03/10/03

La "Distanza di Levenshtein" (DL) è un parametro che indica quanto due stringhe siano simili. In particolare, date due stringhe s e t , esso rappresenta il numero di caratteri che devono essere inseriti, cancellati o modificati in s per renderla identica a t .

L'algoritmo per il calcolo della distanza di Levenshtein è noto dal 1965; i passi risolutivi sono:

1. si prendano in ingresso due stringhe s e t
2. si calcolino $n=\text{lunghezza}(s)$ e $m=\text{lunghezza}(t)$
 - 2.1. se $n=0$ si ha $DL=m$, stampare DL e terminare
 - 2.2. se $m=0$ si ha $DL=n$, stampare DL e terminare
3. inizializzare una matrice d di $n+1$ colonne e $m+1$ righe:
 - 3.1. la prima riga della matrice deve contenere i valori da 0 a n
 - 3.2. la prima colonna della matrice deve contenere i valori da 0 a m
4. si prendano in considerazione ciascun carattere di s ($s[i]$ con i che varia da 0 a $n-1$) e ciascun carattere di t ($t[j]$ con j che varia da 0 a $m-1$)
 - 4.1. se $t[j]=s[i]$ allora costo=0 altrimenti costo=1
 - 4.2. si imposti il valore della cella della matrice d di coordinate $(j+1,i+1)$ pari al minimo valore di:
 - valore della cella di coordinate $(j+1,i) + 1$
 - valore della cella di coordinate $(j,i+1) + 1$
 - valore della cella di coordinate $(j,i) + \text{costo}$
5. DL è uguale al valore della cella della matrice d di coordinate (m,n) ; stampare DL e terminare.

Scrivere un programma in C++ che implementi l'algoritmo descritto. Si ipotizzi una lunghezza massima ragionevole per le stringhe in ingresso.

Esempio di utilizzo del programma (input in neretto):

```
Inserire stringa s: ROMBO  
Inserire stringa t: TROMBA
```

La distanza di Levenshtein è 2