Luca Rosso Matr. 887 433

Mattone Andrea Matr. 882 956

**ALGORITMI E STRUTTURE DATI 2019-20 LABORATORIO**

**Relazione Esercizio 2**

L’applicazione Esercizio2.java calcola la edit distance minore tra due parole utilizzando i metodi edit\_distance oppure edit\_distance\_dyn (versione in programmazione dinamica).

Attraverso uno Scanner ci salviamo in due ArrayList di stringhe la correctme.txt e la dictionary.txt

Per due stringhe inizializziamo una matrice quadrata con dimensione la massima lunghezza delle due stringhe con ogni campo impostato a “-1”.

Per sapere la distanza della parola w dalla parola corrente del file correctme utilizziamo un vettore di interi "distanze".

Questo vettore memorizza la edit distance tra le due parole.

Il vettore distanze verrà aggiornato un numero di volte pari al numero di parole del file correctme.

Ad esempio per una parola w il vettore distanze sarà {1, 2, 1, 4, ...} dove gli elementi di indice j

nel vettore distanze sono le edit distance tra w e la parola contenuta in dictionary in indice j.

Ci salviamo in una variabile minValue la minima edit distance contenuta in questo vettore, cosi da sapere

qual è la edit distance minore tra tutte.

Successivamente viene fatto uno scorrimento delle parole nell'ArrayList e vengono stampate quelle con edit distance uguale a minValue.

La funzione edit\_distance\_dyn è così implementata:

Abbiamo deciso di implementare la memorizzazione attraverso una matrice quadrata chiamata "matrix" di lato la lunghezza massima tra le due stringhe prese in considerazione.

In posizione [i,j] della matrice sarà contenuta la edit\_distance tra la prima stringa di lunghezza s1.length()-i e la seconda stringa di lunghezza s2.length()-j. (Quando una delle due sarà uguale a zero verrà ritornata la restante lunghezza della stringa non nulla in quanto corrisponderà al numero di eliminazioni/inserimenti da fare, come da definizione).

L’applicazione utilizzando la funzione edit\_distance\_dyn impiega circa 1 secondo, per ogni parola w contenuta in correctme.txt, a trovare la lista di parole in dictionary.txt con edit distance minima da w. Questo indipendentemente dal fatto che ne trovi una sola (nel caso in cui ne trovi una uguale a w) come per esempio nel caso di “made”, oppure se ne trovi svariate, come nel caso di “perpeteva” in cui troviamo ad esempio “erpete, permetteva, perpendeva, ecc.”

Abbiamo effettuato un test sugli stessi dataset utilizzando la funzione edit\_distance(s1,s2) ma i tempi sono molto grandi, per trovare la edit distance della prima parola (sui dataset correctme.txt e dictionary.txt) ci mette diversi minuti tanto da rendere fallimentare l’esperimento.