

Laboratorio Algoritmi e Strutture

Dati Relazione esercizio 2

Descrizione dell'esercizio:

L'obiettivo dell'esercizio è determinare la minima distanza di edit tra due stringhe, ovvero il numero minimo di operazioni necessarie per trasformare una stringa di partenza s_2 in una stringa di destinazione s_1 , tramite un algoritmo denominato "edit_distance".

Le operazioni disponibili sono:

- cancellazione di un carattere
- inserimento di un carattere
- non effettuare nessun cambiamento tra i caratteri

Si vuole implementare l'algoritmo in due versioni: statica e dinamica. La versione dinamica viene usata per garantire una migliore efficienza e la conseguente riduzione del tempo computazionale utilizzato dall'algoritmo.

Analisi dei risultati:

Dopo aver implementato l'algoritmo nella sua versione dinamica, si è deciso di testarlo, come da consegna, su un file denominato "correctme.txt", il quale è un file che presenta appositamente errori di battitura. Per testare l'algoritmo si fa anche uso di un altro file chiamato "dictionary.txt", con il quale si analizzano e confrontano le parole con distanza di edit minima rispetto alle parole del file precedentemente nominato.

L'approccio utilizzato per effettuare questo testing è il seguente:

- per ogni parola letta nel file "correctme.txt" si legge l'intero dizionario
- viene chiamato all'uso l'algoritmo edit_distance
- si salva la distanza di edit minima
- si rilegge il dizionario, così da poter stampare ogni parola con la distanza di edit pari a quella minima

Successivamente si trova un'immagine d'esempio durante il testing:

```
Suggested words:
che
Word analyzed:
non
Suggested words:
non
Word analyzed:
avevano
Suggested words:
avevano
Word analyzed:
capito
Suggested words:
capito
Word analyzed:
la
Suggested words:
la
Word analyzed:
wita.
Suggested words:
aita
cita
dita
gita
iota
irta
ista
iuta
lita
pita
sita
ta
vita
witz
zita
dimu@dimu-ubuntu-boot:~/Scrivania/Progetto-ASD/ex2$
```

Come mostrato nell'immagine, l'algoritmo è in grado di restituire tutte le parole con distanza di edit minima rispetto a quella in analisi. Per ogni parola si possono ottenere molteplici suggerimenti in base al confronto che esse hanno avuto con le parole del dizionario.

Questo confronto con le parole del dizionario porta a un rallentamento del tempo di esecuzione dovuto al numero di parole scritte al suo interno, dunque più parole ci sono nel dizionario, più confronti sono da fare con le parole in analisi e più tempo servirà all'algoritmo per portare a termine il suo compito. Inoltre, si denota anche che se una parola risulta "sbagliata" contestualmente,

potrebbe comunque essere una parole esistente, portando a non ottenere la parola del dizionario ricercata per correggerla.