

# ALGORITMI PER INTERNET E WEB: INDICIZZAZIONE DI TESTI

Appello 8 Luglio 2003

COGNOME

NOME

**Esercizio 1.** (*20 punti*) Sia dato un albero binario  $T$  con  $n$  nodi, i cui archi sono etichettati da singoli caratteri. Notare che  $T$  potrebbe essere un *trie* o meno, nel senso che l'unica proprietà richiesta è di avere gli archi etichettati con singoli caratteri, non necessariamente distinti. Un sotto-cammino di lunghezza  $k$  in  $T$  è una sequenza di archi  $e_1, e_2, \dots, e_k$  che collega un nodo in  $T$  a un suo discendente. Le etichette  $a_1, a_2, \dots, a_k$  corrispondenti a tale sotto-cammino formano una stringa di lunghezza  $k$  se concatenate.

Dato un tale albero  $T$  e una stringa pattern  $P$  di lunghezza  $m$ , si vogliono trovare tutti i sotto-cammini di lunghezza  $k = m$ , la cui concatenazione di etichette formi una stringa uguale a  $P$ . Si noti che tali sotto-cammini non partono necessariamente dalla radice.

1. Fornire un algoritmo che richiede tempo  $O(m + n)$ , utilizzando uno degli algoritmi lineari di *string matching* visti a lezione.
2. Discutere e motivare la complessità dell'algoritmo.

# ALGORITMI PER INTERNET E WEB: INDICIZZAZIONE DI TESTI

Appello 20 Giugno 2003

COGNOME

NOME

## Esercizio 2. (10 punti)

Siano date le stringhe  $x = \text{aiuola}$  e  $y = \text{paiolo}$ .

1. Calcolare la distanza di edit tra  $x$  e  $y$  riempiendo la tabella di programmazione dinamica:

	a	i	u	o	l	a
p						
a						
i						
o						
l						
o						

2. Scrivere *tutte* le sequenze minime di operazioni di edit per trasformare  $x$  in  $y$ .