

ALGORITMI PER INTERNET E WEB: INDICIZZAZIONE DI TESTI

Appello 18 Settembre 2003

COGNOME

NOME

Esercizio 1. (20 punti) Si vuole stabilire l'equivalenza tra automi per la ricerca di stringhe:

1. Per una stringa pattern x , sia δ_x l'automa deterministico, descritto nelle dispense, per trovare tutte le occorrenze di x in un testo. Dati due automi deterministici δ_x e δ_y (ma non le loro corrispondenti stringhe x e $y!$), descrivere un algoritmo che, presi in ingresso solamente δ_x e δ_y , decide se $x = y$ senza accedere ai caratteri del testo. Discuterne il costo in termini di complessità computazionale, in funzione del numero di stati degli automi.
2. Per una espressione regolare r , sia δ_r l'automa non-deterministico, descritto nelle dispense, per trovare tutte le occorrenze di r in un testo. Dati due automi non-deterministici δ_r e δ_s (ma non le loro corrispondenti espressioni regolari r e $s!$), descrivere un algoritmo che, presi in ingresso δ_r e δ_s , decide se r e s identificano le stesse occorrenze testuali. Discuterne il costo in termini di complessità computazionale, in funzione del numero di stati degli automi, nei due casi:
 - (a) Il testo è dato in ingresso insieme a δ_r e δ_s .
 - (b) Il testo è generico e non viene fornito in ingresso; ovvero, l'algoritmo decide se r e s identificano le stesse occorrenze usando solamente δ_r e δ_s .

ALGORITMI PER INTERNET E WEB: INDICIZZAZIONE DI TESTI

Appello 18 Settembre 2003

COGNOME

NOME

Esercizio 2. (10 punti) Siano date le sequenze $x = \text{TGCACA}$ e $y = \text{ATACAG}$.

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

	A	T	A	C	A	G
T						
G						
C						
A						
C						
A						

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