

# ALGORITMI PER INTERNET E WEB: INDICIZZAZIONE DI TESTI

Appello 6 Giugno 2005

COGNOME

NOME

**Esercizio 1.** (*10 punti*) Dato il pattern  $P = ATATTAATATTA$ , costruire l'automa deterministico (sezione 1.2 delle dispense) per  $P$ . Discutere la complessità in tempo e spazio della costruzione dell'automa così ottenuto.

## ALGORITMI PER INTERNET E WEB: INDICIZZAZIONE DI TESTI

Appello 6 Giugno 2005

COGNOME

NOME

**Esercizio 2.** (10 punti) Data una stringa  $x$  e un insieme  $S$  di stringhe, si vuole calcolare la lunghezza della *sottosequenza comune più lunga* tra  $x$  e  $y$ , indicata con  $s(x, y)$ , per ogni stringa  $y \in S$  utilizzando il trie  $\delta$  costruito sulle stringhe in  $S$  (capitolo 5 delle dispense). Scrivere lo pseudocodice per realizzare efficientemente tale calcolo delle lunghezze  $s(x, y)$  mediante  $\delta$ . Quali sono i vantaggi rispetto al calcolo senza l'utilizzo del trie?

**ALGORITMI PER INTERNET E WEB: INDICIZZAZIONE DI TESTI**

Appello 6 Giugno 2005

**COGNOME**

**NOME**

**Esercizio 3.** *(10 punti)* Costruire l'automa di Aho-Corasick per l'insieme di stringhe ATATATAAT, ATA, TAT, ATAT, indicando le sue componenti  $\delta$  e  $\varphi$  ed evidenziandone gli stati finali.