

ALGORITMI PER INTERNET E WEB: INDICIZZAZIONE DI TESTI

Appello 20 Giugno 2003

COGNOME

NOME

Esercizio 1. (*12 punti*) Utilizzando il paradigma SHIFT-AND, mostrare i passi di calcolo per la ricerca del pattern $P = ABABC$ nel testo $T = ABAABABCAB$ senza utilizzare confronti diretti tra i caratteri di entrambe le stringhe. La descrizione dei passi di calcolo consiste nel riportare, analogamente a quanto illustrato nelle dispense del corso, le variazioni della parola di memoria D durante la scansione del testo T , e i prefissi del pattern P via via associati ad esse.

ALGORITMI PER INTERNET E WEB: INDICIZZAZIONE DI TESTI

Appello 20 Giugno 2003

COGNOME

NOME

Esercizio 2. (18 punti) La lunghezza del prefisso comune più lungo tra due stringhe x e y è il massimo numero di caratteri iniziali in comune, e viene indicato con $lcp(x, y)$. Per esempio, $lcp(x, y) = 2$ per $x = \underline{v}ariable$ e $y = \underline{v}alore$. Descrivere a parole e in pseudocodice un algoritmo che, preso in ingresso un trie δ per un insieme di stringhe S , fornisce in uscita tutti i possibili valori $lcp(x, y)$ per ogni coppia di stringhe $x, y \in S$ con $x \neq y$. Discutere la complessità al caso pessimo in tempo e in spazio dell'algoritmo proposto.