

Examen Parcial de Recuparació de la Informació
Curs 2007-2008, Quadrimestre d'hivern

Temps: 2h

Exercici 1 (2 punts)

- Simuleu la cerca dels patrons *ATAA*, *AATA*, *TATA*, *TATAA* dins la seqüència *ATAAATATAA*, aplicant l'algorisme de SetHorspool amb una taula de dos símbols.

- Quin seria el valor esperat del salt de la finestra per a una seqüència aleatòria.

Exercici 2 (2 punts)

- Simuleu la cerca del patró *ARGA* dins la seqüència *AARARRTA* on *R* és un caràcter extés, aplicant l'algorisme BNDM.

- Si aplicant l'algorisme BNDM per al patró *GTGTA* obtenim $D_1 = (01010)$, $D_2 = (10100)$ i $D_3 = (00000)$, què es pot dir dels elements que formen el text?

Exercici 3 (2 punt) Apliqueu l'autòmata de Thompson per buscar l'expressió regular $A(A^*|G^*)G$ dins la seqüència $GAGGA$.

Exercici 4 (2 punts)

- Trobeu el millor alineament entre les 3 seqüències CA , A i AA .

- Busqueu el millor alineament local entre les seqüències $ATGTT$ i $ATTGTA$.

Exercici 6 (2 punts) Un jugador juga contra el casino segons el resultat de tirar un dau de 4 cares (un tetraedre) i guanya qui ha tret un resultat més alt. El jugador sempre juga amb un dau equilibrat pero el casino en té dos, un d'equilibrat i un altre on el 4 té la probabilitat de $2/5$ i la resta de valors 1,2,3 tenen la mateixa probabilitat, i en cada tirada el casino escull un del dos daus a l'atzar.

- Quina és la seqüència d'estats més probable si els resultats han estat 41342.

- Quina probabilitat té el casino de treure un 4?