

Examen Parcial de Recuparació de la Informació
Curs 2009-2010, Quadrimestre d'hivern

Temps: 2h

Exercici 1 (1.5 punts) Simuleu la cerca del patró $AATR$, on R és A o G , dins la seqüència $ATCTAATGCRACC$, aplicant l'algorisme BNDM.

Exercici 2 (0.5 punts) Perquè la gràfica de l'algorisme de força bruta mostrava que el temps de cerca és quasi bé independent de la llargada del patró?

Exercici 3 (1.5 punts) Simuleu la cerca dels patrons *ATAA*, *ATT*, *TATA* dins la seqüència *ATATAATTAT* aplicant l'algorisme SBOM.

Exercici 4 (0.5 punts) En quines situacions l'algorisme SBOM és el més indicat i perquè?

Exercici 4 (2 punts) Un jugador juga contra el casino segons el resultat de tirar un dau de 4 cares (un tetraedre) i perd qui ha tret un resultat més baix. El jugador sempre juga amb un dau equilibrat pero el casino en té dos, un d'equilibrat i un altre on cada valor i ($i = 1, 2, 3, 4$) té la probabilitat de $i/10$, 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 41.

- En cada tirada quina és la probabilitat de que perdi el casino?

Exercici 5 (0.5 punts cada apartat)

- Has creat una seqüència de valors amb el model

$M = 6$

$N = 2$

A:

0.5 0.5

0.5 0.5

B:

0.5 0.1 0.1 0.1 0.1 0.1

0.1 0.1 0.1 0.1 0.1 0.5

pi:

0.5 0.5

Quin creus que seria el model que *esthm* t'estimaria i perquè?

- Si hem de buscar molts patrons en una seqüència, quins són els paràmtres a tenir en compte per escollir el millor algorisme?

- Perquè no és possible buscar el millor aliniament entre més de 10 seqüències llargues?

- Perquè s'utilitzen suffix-trees per buscar MUMs?