

Examen Parcial de Recuperació de la Informació
Curs 2004-2005, Quadrimestre d'hivern

Temps: 2h45m

Exercici 1 (2 Punts)

Donat el text T i la finestra que conté al patró:

T:

	...	G	G	A	C	A	...
--	-----	---	---	---	---	---	-----

A	C	A	G	A	C	A
---	---	---	---	---	---	---

- Simuleu l'execució de l'algorisme BNDM i digueu en quina posició es situarà la finestra en el pas següent.

- Construïu l'autòmat factor oracle i digueu en quina posició es situarà la finestra en el pas següent.

Exercici 2 (2 Punts) Donada la seqüència ACAGACA\$, construïu el seu Suffix Tree aplicant l'algorisme quadràtic de construcció.

Exercici 3 (2 Punts) Donada la taula de l'algorisme semi-global

	ACTAGATA TACA										
	0	0	0	0	0	0	0	0	0	0	0
	/ / / / / / / / / / / /										
G	1	1	1	1	1	0	-1	1	1	1	1
	/ / / / / / / / / / / /										
A	2	1	-2	2	1	1	0	-1	1	-2	1
	/ / / / / / / / / / / /										
T	3	2	2	2	2	2	1	0	-1	1	-2
	/ / / / / / / / / / / /										
A	4	3	3	3	2	-3	2	1	0	-1	1

- Com podem saber si hi ha subseqüències de ACTAGATATACA a una distància d'edició d de GATA.

- Com es busquen aquestes subseqüències?

- indiqueu quines subseqüències de ACTAGATATACA estan a una distància d'edició u de GATA.

Exercici 4 (1 Punt) Ompliu les cel·les simulant l'algorisme de programació dinàmica:

		A	
		C	
	0	?	
-A	?	?	

Exercici 5 (1.5 Punts) A partir de les similaritats entre totes les seqüències,

	S1	S2	S3	S4	S5
S1		20	12	5	1
S2			16	9	3
S3				4	5
S4					8

construiu l'arbre de l'aliniament múltiple.

Exercici 6 (1.5 Punts) Indica què faries en cadascún dels següents casos:

- Com buscaries de forma eficient les subseqüències comunes a tres seqüències conegudes?

- Si en el text hi tens símbols d'un alfabet extés respecte de l'alfabet del patró, quin algorisme es pot adequar a aquesta situació i perquè ?

- Quin és el cost de buscar k patrons utilitzant l'algorisme SBOM ?