

Tehtävä 2: Aineistotehtävä - Vastaus

1. Luettele pituusjärjestyksessä kaikki sekvenssin $\langle C, A, G, A, T, A, G, A, G \rangle$ etuosat (5p).

Vastaus: $\langle \rangle, \langle C \rangle, \langle C, A \rangle, \langle C, A, G \rangle, \langle C, A, G, A \rangle, \langle C, A, G, A, T \rangle, \langle C, A, G, A, T, A \rangle, \langle C, A, G, A, T, A, G \rangle, \langle C, A, G, A, T, A, G, A \rangle$ ja $\langle C, A, G, A, T, A, G, A, G \rangle$.

2. Luettele kaikki sekvenssin $\langle A, G, C, G, A \rangle$ neljän pituiset alisekvenssit (5p).

Vastaus: $\langle A, G, C, G \rangle, \langle A, G, C, A \rangle, \langle A, G, G, A \rangle, \langle A, C, G, A \rangle$ ja $\langle G, C, G, A \rangle$.

3. Tarkastellaan sekvenssejä $X = \langle A, C, A \rangle$ ja $Y = \langle G, T, C \rangle$.

- (a) Laske sekvenssien X ja Y pisimmän yhteisen alisekvenssin pituus aineistossa esitetyllä menetelmällä. Anna vastauksessasi Taulukon 1 kaltainen taulukko. (3p)

Vastaus:

Taulukko 1: Tehtävän 3 (a) vastaus.

		j	0	1	2	3
i		y_j	G	T	C	
0	x_i	0	0	0	0	
1	A	0	0	0	0	
2	C	0	0	0	1	
3	A	0	0	0	1	

- (b) Mikä on sekvenssien X ja Y pisin yhteinen alisekvenssi? (1p)

Vastaus: $\langle C \rangle$.

4. Tarkastellaan sekvenssejä $X = \langle T, A, C, A, G, C, T, C \rangle$ ja $Y = \langle A, T, C, G, G, A, C, T \rangle$.

- (a) Laske sekvenssien X ja Y pisimmän yhteisen alisekvenssin pituus aineistossa esitetyllä menetelmällä. Anna vastauksessasi Taulukon 1 kaltainen taulukko. (6p)

Vastaus:

Taulukko 2: Tehtävän 4 (a) vastaus.

		j	0	1	2	3	4	5	6	7	8
i	x_i	y_j	A	T	C	G	G	A	C	T	
		0	0	0	0	0	0	0	0	0	0
1	T	0	0	1	1	1	1	1	1	1	1
2	A	0	1	1	1	1	1	2	2	2	2
3	C	0	1	1	2	2	2	2	3	3	3
4	A	0	1	1	2	2	2	3	3	3	3
5	G	0	1	1	2	3	3	3	3	3	3
6	C	0	1	1	2	3	3	3	4	4	4
7	T	0	1	2	2	3	3	3	4	5	5
8	C	0	1	2	3	3	3	3	4	5	5

- (b) Luettele kaikki sekvenssien X ja Y pisimmät yhteiset alisekvenssit. (5p)

Vastaus: Pisimpiä yhteisiä alisekvenssejä on neljä: $\langle A C A C T \rangle$, $\langle A C G C T \rangle$, $\langle T C A C T \rangle$ ja $\langle T C G C T \rangle$.

Viitteet

Cormen, T.H., Leiserson, C.E. ja Rivest, R.L, *Introduction to Algorithms*, 1990, McGraw-Hill.

Charras, C. ja Lecroq, T., *Longest common subsequences*, <http://www-igm.univ-mlv.fr/~lecroq/seqcomp/node4.html> (luettu maaliskuussa 2008).