

Algorithmen und Datenstrukturen in der Bioinformatik

Fünftes Übungsblatt WS 09/10

Abgabe Montag 23.11.2008 12:00

Name: _____ Übungsgruppe: A B C

Matrikelnummer: _____ Ich kann Aufgabe _____ nicht vorrechnen.

Exercise 11: Affine gap costs

Calculate the global alignment with affine gap costs for the following sequences.

$$S_1 = \text{ACCCT} \text{ and } S_2 = \text{GACT}$$

Use 3 as match score, -5 as mismatch score, 3 as gap open cost and 1 as gap extension cost. Write down the matrices for M , I_x and I_y .

		A	C	C	C	T
	M					
	I_x					
	I_y					
G	M					
	I_x					
	I_y					
A	M					
	I_x					
	I_y					
C	M					
	I_x					
	I_y					
T	M					
	I_x					
	I_y					

Exercise 12: Bounded gap costs

Extend the algorithm for affine gap costs to handle *bounded* gap costs. Bounded gap costs are defined by a gap opening cost d and a gap extension cost e , just like affine gap costs. However,

they also have a parameter l which is the bound on the gap costs. Gaps with costs larger than l instead have cost of just l . For example, assume that $d = 3$ and $e = 1$. This means that a gap of length 3 has total cost 5, and a gap of length 4 has cost 6. Suppose now that $l = 6$. Then, any gap of length > 4 has the same cost as a gap of length 4, i.e. 6.

Write down the necessary modifications to the pseudo-code of global alignment with linear gap costs to allow bounded gap costs.

Exercise 13: Banded alignment

Calculate the banded alignment for the sequences $S_1 = ACCCT$ and $S_2 = GACCT$ with a match score of 1, mismatch costs and gap costs of -1 and $k = 1$ using the **KBand** algorithm. Repeat the calculation with $k = 0$. What difference do you observe? How do you explain it?
