Universität Bielefeld
Technische Fakultät

AG Genominformatik
Dipl.-Inform. Klaus-Bernd Schürmann

Lecture: Spezielle Algorithmen der Sequenzanalyse
Summer semester 2006

Exercises

Exercise 3, Discussion: 04/26/2006.

1. **Computing optimal alignments.**
   Given the two strings u = GAG and v = CTGGAGT. Calcualte the dynamic programming matrices for the similarity function according to the score function δ, visualize the maximizing paths, and give the optimal alignments
   (a) for the global alignment model and
   (b) the end-gap free alignment model.

δ	A	C	G	T	-
A	4	-5	-2	-5	-3
C	-5	3	-5	-1	-3
G	-2	-5	4	-5	-3
T	-5	-1	-5	3	-3
-	-3	-3	-3	-3	

2. **Number of co-optimal alignments.**
   Since the traceback paths in a dynamic programming table correspond one-to-one with the optimal alignments, the number of distinct co-optimal global alignments can be obtained by computing the number of distinct traceback paths. Give an algorithm to compute this number in $O(n\,m)$ time.
   Hint: Use dynamic programming.

3. **Affine gap cost.**
   The recurrences relations we developed for the affine gap model follow the logic of paying `gapinit` + `gapext` when a gap is "initiated" and then paying `gapext` for each additional space used in that gap. An alternative logic is to pay `gapinit` + `gapext` at the point when the gap is "completed". Write recurrence relations for the affine gap model following that logic. The recurrences should compute the alignment in $O(n\,m)$ time.
   Recurrences of this type are part of Gotoh's algorithm.