

Algorithms in Bioinformatics I, WS2002/3

Assignment sheet # 6

Daniel Huson

November 17, 2002

1 The number of mixed cycle inequalities for pairwise alignments (2 points)

Given two sequences x and y of lengths n and m , respectively. Consider the ILP formulation of the maximal trace problem for the alignment of two sequences. Write down a formula for the number of different mixed-cycle inequalities that we need to state in the ILP. Please explain your result!

2 Implementation of the fill stage of the Nussinov algorithm (4 points)

Implement the fill stage of the Nussinov algorithm. In other words, write a program that reads as input a RNA sequence, fills the γ matrix and then reports the maximum number of base pairs attainable in a secondary structure.

To do this, modify Program8.java, which is available from:

www-ab.informatik.uni-tuebingen.de/teaching/ws02/abi1/programs/program08.zip. This download also contains data files to run the algorithms on. Please run the program on all supplied data sets. The file `example.output` shows the result expected for input file `example.fa`.

The program has one command-line option, the name of the fasta file containing the RNA sequence.

3 Implementation of the traceback stage of the Nussinov algorithm (4 points)

Modify Program8.java so that it performs the recursive traceback discussed in the lecture. The program should print out each base pair i, j as (i, j) on a separate line (with $i < j$).

Additional useful activities

Draw the secondary structures that your program computes.

Due by 10am, Monday, 25 Nov 2002