

Homework 3

Due: 4/27/2010

1. In class we discussed a dynamic programming algorithm for folding an RNA molecule. This algorithm is global in nature, in that the entire molecule gets folded. There are biological contexts in which it makes sense to explore a local formulation of the folding algorithm - find substrings of the original sequence that fold well. One example are transcription terminators - stretches of DNA that form a hairpin structure which blocks the transcription process.

Describe how you would modify the RNA folding algorithm to find high quality local alignments. You should also provide a definition of what you call a "high quality" alignment.

2. Write out pseudo-code that performs progressive multiple alignment. Specifically describe how you would align one sequence to an existing multiple alignment, as well as how you would align two multiple alignments to each other.

3. Describe an algorithm that solves the multiple longest increasing subsequence problem: given k sequences of integers, find the largest subset of integers that occur in increasing order in all the sequences. Discuss how your algorithm might change if some integers are repeated in some sequences.

For example, for the 3 sequences:

5 7 **1** **3** 2 **8** 4 6

2 5 **1** 2 **3** 9 **8** 10

1 4 9 2 5 7 **3** 6 **8**

a possible answer might be 1,3,8 (bolded above).