MO640 – Biologia Computacional
Segundo Semestre de 2014
Primeira Lista de Exercícios

1. Given two sequences, which value is larger: their local similarity or their global similarity? Why? How does their semi-global similarity compare with the other two values?

2. Show all optimal global alignments between sequences x = ACTGTGCT and y = ATGGTCT, using $match = +3$, $mismatch = -2$, and $gap = -5$.

3. *Longest Common Subsequence (LCS) Problem*: given two sequence x and y, find the longest subsequence present in both of them. A subsequence is a sequence that appears in the same relative order, but not necessarily contiguous. For example, in the string abcdefg, "abc", "abg", "bdf", "aeg" are all subsequences. Use the Needleman-Wunsch algorithm to solve the LCS problem. Justify your answer.

4. An alignment of circular strings is defined as an alignment of linear strings forming by cutting (linearizing) these circular strings an arbitrary position. Devise an efficient algorithm to find an optimal global alignment of circular strings.

5. A local alignment between two different strings x and y finds a pair of substrings, one in x and the other in y, with maximum similarity. Suppose that we want to find a pair of (nonoverlapping) substrings within string v with maximum similarity (*Optimal Inexact Repeat problem*). Computing an optimal local alignment between v and v does not solve the problem, since the resulting alignment may correspond to overlapping substrings. Devise an algorithm for the Optimal Inexact Repeat problem.

6. A string x is called a supersequence of a string y if y is a subsequence of x. For example, ABLUE is a supersequence for BLUE and ABLE. Given strings x and y, devise an efficient algorithm to find the shortest supersequence for both x and y.