**CS423 Lab 5: String Alignment**

**Authors: Sara Perkins, Caleb Piekstra**

**Write-up (45 points):**

1. S and T
   1. -1
   2. AGC-TCT--A

TGCATCTCGT

* 1. 21
  2. GCGTCT

GCATCT

* 1. 5
  2. GCTCT

1. Yeast and Marty McFly
   1. 549
   2. In the one .zip file
   3. 660
   4. In the one .zip file
   5. 900
   6. In the one .zip file
2. Ran
   1. Less than—we are looking at a randomly generated sequence versus a naturally-selected mutated ortholog
   2. Less than—we are looking at a randomly generated sequence versus a naturally-selected mutated ortholog. Since it is an ortholog, they are fairly similar with a few insertions, deletions, or substitutions, whereas random has no similar heritage.
   3. Global—667, no
   4. Local—714, no
3. Blah
   1. Yes. Since the local alignment can skip bases at the beginning and/or end of the strings, it can skip a bunch of mismatches or insertions/deletions that the global alignment would have to account for, decreasing the global alignment score permanently. Also, the local alignment does not use negative values, simply replacing any negatives with 0 (a fresh start). In contrast, the global alignment saves negatives which can lower the score significantly compared to a 0.
   2. The two alignment values are equal when the strings align perfectly (all ordered bases in string one are identical to string two).
4. Larger for the long sequences because you have more chances to get higher scoring alignments.
5. Bigger for the short sequences because the long ones have a higher chance of getting lots of mismatches and gaps, lowering the score substantially.