

## Homework #1 (deadline: 3/6/2017)

1. Write a program that given a fasta format DNA sequence file,

```
>seq1
acgtaccggt...
>seq2
CGATCCGGTT...
```

computes the **local** pairwise alignments between every pair of sequences in the file, and outputs the optimal alignments with the alignment scores. Your scoring function is specified as  $w(\text{match}) = 3$ ,  $w(\text{mismatch}) = -1$  and  $w(\text{indel}) = -3$ . Your test examples are in HW1File1.

2. Write a program to compute *dissimilarity dependent* **global** pairwise alignment (banded DP), and output one optimal alignment with the alignment score. Your scoring function is  $d(\text{match}) = 0$ ,  $d(\text{mismatch}) = 1$  and  $d(\text{indel}) = 1$ . Note that this is a distance scoring function, so the optimal alignment produces the minimum distance. Your test examples are in HW1File1.

Each student should submit the homework online (1) with a MS Word report with the idea of implementation, summary and analysis of results (alignment scores and running time), and two example alignments (one from HW1File1 and the other from HW1File2), (2) comparisons of these two programs, and (3) your source code and your binary executable file (for C/C++ users).