

CMSC423 Project 1  
Handed out: 2/21/07  
Due: 3/14/07

For this project you will have to create a program that takes as input a sequence and searches for all good alignments of this sequence inside a database, using the Smith-Waterman dynamic programming algorithm described in class.

**Inputs:**

- one query sequence in FASTA format
- multiple sequence database in FASTA format
- similarity matrix (format?)
- minimum % identity
- gap extension penalty
- gap creation penalty

**Outputs:**

- precise alignment for all hits over %identity

**Sample inputs:**

See syllabus page: <http://www.cbcb.umd.edu/confcour/CMSC423-syllabus.shtml>

**Output format:**

See syllabus page: <http://www.cbcb.umd.edu/confcour/CMSC423-syllabus.shtml>

**Submission:**

Email me and Behjat your code and a README file that describes how to compile, install, and run your code.

**Grading!** We will grade all aspects of the code, including how “pretty” it looks. Specifically pay attention to the following aspects:

1. Please make sure that your code works as advertised in the README file you provided. If your code doesn't work as indicated in the README file you will automatically lose 50% of the grade for this assignment.
2. Please provide copious comments and format your code so that it is easy to read. Part of your grade will be based on the formatting of the code.
3. Fastest programs get additional credit:
  - a. 20 points for fastest
  - b. 12 points for second fastest
  - c. 5 points for third fastest

Please contact me as soon as possible if you have any questions regarding this assignment, or if you “get stuck” and might not be able to complete the assignment on time. Once the assignment is due I will no longer accept any excuses.

**Good Luck!**

