Project 4: Viterbi algorithm for protein family profile HMM alignment

Objective: Implement the Viterbi algorithm for protein family profile HMM alignment. You only need to implement the global version.

A. Input format

The format of the first input of the program, is the query profile HMM. You can download as an example a profile HMM file from http://pfam.xfam.org/family/PF00004/hmm. For more detailed instructions of the file format please see http://eddylab.org/software/hmmer3/3.1b2/Userguide.pdf (page 106).

The second input of the program is a FASTA file. It contains a single sequence. Please refer to HW1 for more detailed discussions on its format.

B. Output format

The output of the program is similar to that of HW1, except the following differences:

- 1: The score is the likelihood of the Viterbi path rather than the alignment score;
- 2: The sequence of the profile HMM is determined by the match-state characters that have the highest emission probability;
- 3: Use '+' to indicate a positive match if the emission probability of the matched character (in the target) in the matched state is higher than the background frequency of the matched character.

C: Test case

You are encouraged to check your result using HMMER3. You can download HMMER3 at http://hmmer.org/. Detailed manual and instructions can also be found from the website.

D: Submission
- Please submit your source code via Blackboard by April 30th 2021.
- Include a README file describing how to run your program.