BINF 730 - Biological Sequence Analysis

Final Exam

Due December 17, 2018

Please answer questions fully and show all work.

1. Please read the article <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2766791/>. What are the three types of hidden Markov models described? How are they useful for sequence analysis?
2. Please read section 1.2 at the web site <http://www.cs.cmu.edu/~epxing/Class/10701-11f/HW/hw3_solution.pdf> entitle “HMM for DNA Sequence”. Please work through the steps (reproducing calculations if possibl). How might a similar model be used to analyze DNA sequences?
3. If you had sequences from several different organisms and wanted to find how the organisms were related what would you do? Be sure to mention specific methods we discussed in class.
4. Suppose we had identified a protein motif and we wanted to find out if other sequences has this motif. Please describe how you would do this. Be sure to mention specific methods we discussed in class.
5. Suppose on Mars a diverse variety of life was discovered with a new genetic code that differed from that on Earth. What challenges do you think there would be for understanding the genome, genes, and comparing different species? How might one solve these challenges?