BMI 651

Joshua Burkhart

February 29, 2016

HW5: Writeup

Question 1.a.:

Given that the probability of switching to a GC-rich region from an AT-rich region is 15% and the probability of inverse is 20%. Construct a transition matrix for the 2 state model. [APPENDIX 1A]

Question 1.b:

In AT-rich regions, the probability of each nucleotide has been found to be 30%, 10%, 20%, and 40% for A, C, G, and T respectively. In the GC-rich regions, is has been found to be 5%, 30%, 55%, and 10% respectively. Construct an emission matrix, for each state. [APPENDIX 1Bi, 1Bii]

Question 1.c:

Assume equal likelihood for starting in each state. Generate a 100 nucleotide long DNA sequence, by employing an HMM and the matrices defined in parts 1.a and 1.b. [APPENDX 1C]

Question 1.d:

Suppose the following DNA sequence: “A”, “A”, “G”, “C”, “G”, “T”, “G”, “G”, “G”, “G”, “C”, “C”, “C”, “C”, “G”, “G”, “C”, “G”, “A”, “C”, “A”, “T”, “G”, “G”, “G”, “G”, “T”, “G”, “T”, “C” What is the most probable states (i.e. either AT or GC rich states) that generated this sequence [APPENDIX 1D]