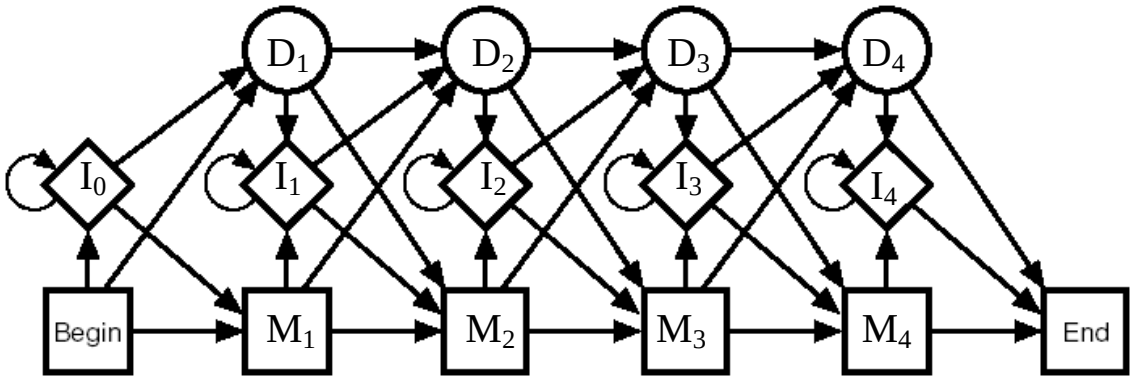
CENG 465
Fall 2017-2018

Due Date: November 22, 2017, 23:55

Assignment #2

Written Assignment

Consider the Hidden Markov Model (HMM) with the structure and parameters given below constructed for a small DNA sequence motif.



$\pi_{\text{Begin}} = 1.0$, π for all other states is 0.0

Transition probabilities:

	Begin	**M_1**	**M_2**	**M_3**	**M_4**	**I_0**	**I_1**	**I_2**	**I_3**	**I_4**	**D_1**	**D_2**	**D_3**	**D_4**	**End**
Begin	0	**0.8**	0	0	0	**0.1**	0	0	0	0	**0.1**	0	0	0	0
M_1	0	0	**0.6**	0	0	0	**0.3**	0	0	0	0	**0.1**	0	0	0
M_2	0	0	0	**0.7**	0	0	0	**0.1**	0	0	0	0	**0.2**	0	0
M_3	0	0	0	0	**0.7**	0	0	0	**0.2**	0	0	0	0	**0.1**	0
M_4	0	0	0	0	0	0	0	0	**0**	**0.4**	0	0	0	0	**0.6**
I_0	0	**0.6**	0	0	0	**0.2**	0	0	0	0	**0.2**	0	0	0	0
I_1	0	0	**0.4**	0	0	0	**0.4**	0	0	0	0	**0.2**	0	0	0
I_2	0	0	0	**0.5**	0	0	0	**0.4**	0	0	0	0	**0.1**	0	0
I_3	0	0	0	0	**0.7**	0	0	0	**0.2**	0	0	0	0	**0.1**	0
I_4	0	0	0	0	0	0	0	0	0	**0.3**	0	0	0	0	**0.7**
D_1	0	0	**0.3**	0	0	0	**0.2**	0	0	0	0	**0.5**	0	0	0
D_2	0	0	0	**0.4**	0	0	0	**0.5**	0	0	0	0	**0.1**	0	0
D_3	0	0	0	0	**0.4**	0	0	0	**0.4**	0	0	0	0	**0.2**	0
D_4	0	0	0	0	0	0	0	0	0	**0.4**	0	0	0	0	**0.6**
End	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Note: Each row sums up to 1, except the row for the End state.

Emission probabilities:

M_1: ω_A= 0.1, ω_C= 0.7, ω_G= 0.1, ω_T= 0.1

M_2: ω_A= 0.6, ω_C= 0.2, ω_G= 0.1, ω_T= 0.1

M_3: ω_A= 0.7, ω_C= 0.1, ω_G= 0.1, ω_T= 0.1

M_3: ω_A= 0.1, ω_C= 0.1, ω_G= 0.4, ω_T= 0.4

I_1, I_2, I_3, I_4: ω_A= 0.25, ω_C= 0.25, ω_G= 0.25, ω_T= 0.25

Note: Begin, End, and delete states do not emit any symbols.

Align the following sequence to the profile HMM using the Viterbi algorithm. In other words, find the sequence of states which is most likely to emit that sequence. Show the contents of the partial probability table you construct. What is the probability associated with the alignment path?

GCTAAACC

Note: Initialize the partial probability (or likelihood) table with $v_{Begin}(\text{“”})=1.0$. Do not use log of probabilites. You may indicate the probabilities in scientific notation, such as 0.234e-10. You may use some programming or tools like Excel to help in your calculations.

Submission:

Submit your solution via ODTU-Class before the deadline. Show both the sequence of states and the associated probability along with the partial probability table (i.e., the Viterbi table). Your solution may be a scanned copy of a handwritten solution or a document written on computer.