CENG465 Assignment 2

Mustafa Ozan Alpay

02/05/2021

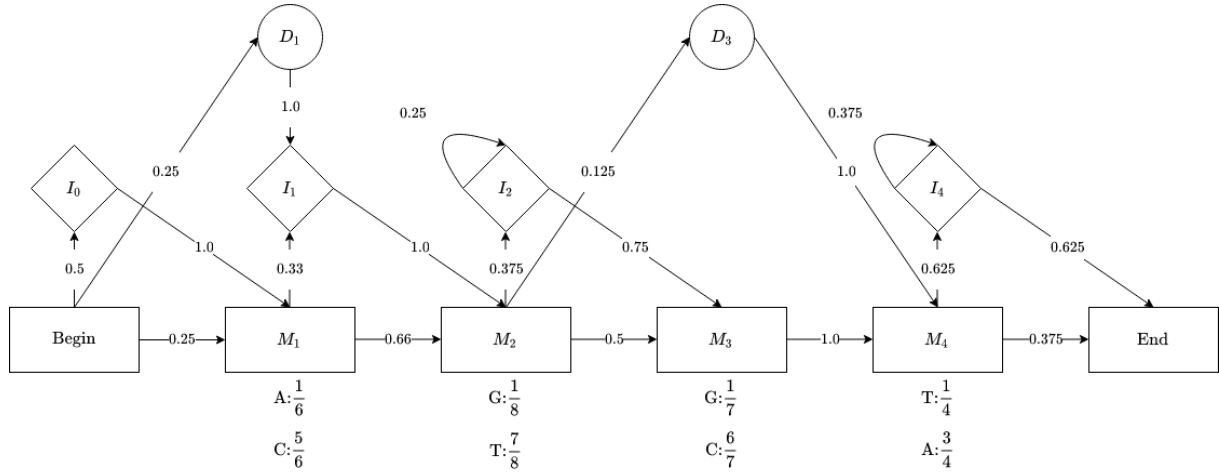## Part A

Figure 1: pHMM Structure

- Seq 1: $Begin \to D_1 \to I_1 \to M_2 \to M_3 \to M_4 \to I_4 \to End$

- Seq 2: $Begin \to I_0 \to M_1 \to I_1 \to M_2 \to M_3 \to M_4 \to I_4 \to I_4 \to End$

- Seq 3: $Begin \to M_1 \to I_1 \to M_2 \to I_2 \to I_2 \to M_3 \to M_4 \to End$

- Seq 4: $Begin \to I_0 \to M_1 \to M_2 \to D_3 \to M_4 \to I_4 \to I_4 \to I_4 \to End$

- Seq 5: $Begin \to D_1 \to I_1 \to M_2 \to M_3 \to M_4 \to End$

- Seq 6: $Begin \to I_0 \to M_1 \to M_2 \to I_2 \to M_3 \to M_4 \to I_4 \to End$

- Seq 7: $Begin \to I_0 \to M_1 \to M_2 \to I_2 \to M_3 \to M_4 \to End$

- Seq 8: $Begin \to M_1 \to M_2 \to M_3 \to M_4 \to I_4 \to End$

Part B

The best path for the alignment `AATGAC` is $Begin \to I_0 \to M_1 \to M_2 \to M_3 \to M_4 \to I_4 \to End$, and the probability associated with that is 0.000064 according to Table 1.

	" "	A	A	T	G	A	C
BEGIN	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
I0	0.000000	0.125000	0.000000	0.000000	0.000000	0.000000	0.000000
M1	0.000000	0.041667	0.020833	0.000000	0.000000	0.000000	0.000000
D1	0.250000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
I1	0.000000	0.062500	0.003472	0.001736	0.000000	0.000000	0.000000
M2	0.000000	0.000000	0.000000	0.012153	0.000217	0.000000	0.000000
I2	0.000000	0.000000	0.000000	0.000000	0.001139	0.000071	0.000004
M3	0.000000	0.000000	0.000000	0.000000	0.000868	0.000000	0.000046
D3	0.000000	0.000000	0.000000	0.001519	0.000027	0.000000	0.000000
M4	0.000000	0.000000	0.000000	0.000000	0.000000	0.000651	0.000000
I4	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000102
END	X	X	X	X	X	X	0.000064

Table 1: Partial Probability Table for AATGAC