Bioinformatics (Akiyama) Exercise #08

Student ID Name

1) Consider about the profile HMM model for DNA sequence motif. (see next page)

Character emission probabilities are given as :

M1 : P(A)=0.5, P(T)=0.25, P(G)=0.25, P(C)=0.0

M2 : P(A)=1.0, P(T)=P(G)=P(C)=0.0

M3 : P(A)=0.5, P(T)=0. 5, P(G)= P(C)=0.0

INS : P(A)=P(T)=P(G)=P(C)=0.25

Now we observed an output character sequence as :

start →T→ A→ A→ T→end

(a) Calculate the total probability of state transitions and symbol emissions, by assuming the hidden state transition path below.

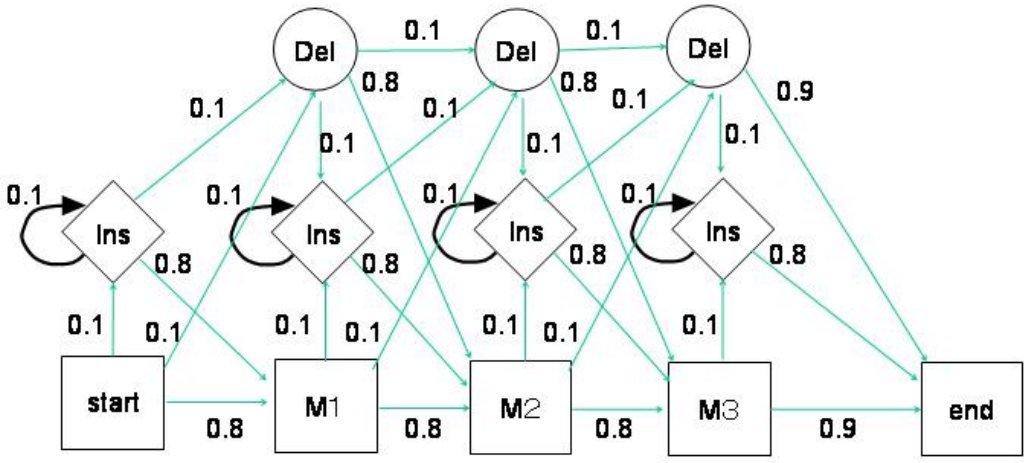
start → M1 → M2 → M3 → Ins →end

Ans.

(b) Calculate the total probability of state transitions and symbol emissions, by assuming the hidden state transition path below.

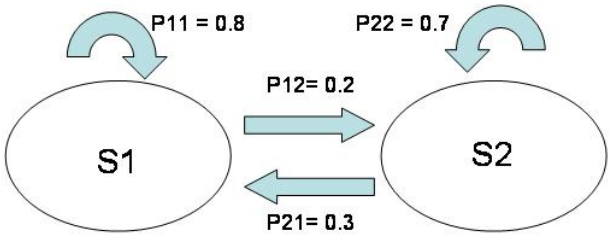
start → M1 → Ins → M2 → M3 →end

Ans.



*Correction : transition probability from last Ins state to 'end' is not 0.8 but 0.9.

2) Calculate the stationary distribution for the following first-order Markov chain model. The system has two states named S1 and S2. Calculate P1（∞）and P2（∞）, where P1（∞）+P2（∞）= 1.



$$\begin{pmatrix} P1(t+1) \\ P2(t+1) \end{pmatrix} = \begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix} \begin{pmatrix} P1(t) \\ P2(t) \end{pmatrix}$$

Ans. P1(∞)= P2(∞)=