Stochastic Processes

Homework #6

Jan, 2023

Saba Hashemi - 97100581

Problem 1

Q. A little girl named Sarah loves making origami flowers. It's one of the three indoor activities that she loves to do almost equally. But when the weather is good, she prefers to go outside triple times as staying indoors. Each day Sarah makes origami flowers, she gives one flower to her mom. Assuming the weather condition stays the same as yesterday with a probability of 0.7:

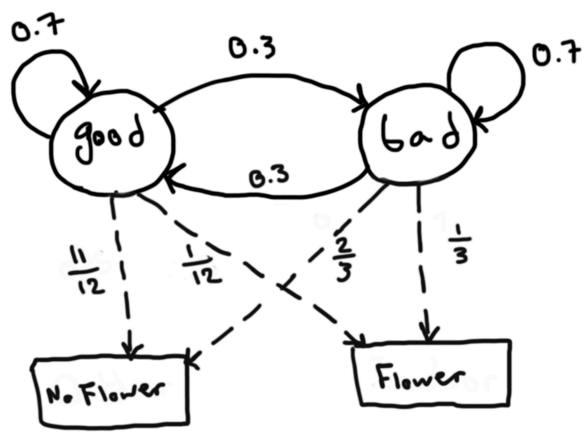
(a) Build a Markov chain representing the problem and calculate its re- quired probabilities.

(b) Calculate the expected number of flowers Sarah's mom has gathered in 10 years (yep she had continued her routine in 10 years!)

(c) This Saturday the weather was good. With what probability Sarah's mom will receive two flowers by Monday night (in 3 days)?

(d) If Sarah's mom received only one flower on Sunday (and none in Saturday and Monday), what had been the most probable weather conditions for Sunday and Monday?

A.



(a)

(b) We will define R.V. X as the total number of flowers in 10 years. Since 10 years is a big number and we do not have the initial probabilities, we can use stationary state of Markov Models to get an approximation for E[X].

$$E[N_{good}|X_0 = bad] = 0.3 * 1 + 0.7 * (1 + E[N_{good}|X_0 = bad])$$
$$E[N_{good}|X_0 = bad] = 1 + 0.7 * E[N_{good}|X_0 = bad]$$
$$E[N_{good}|X_0 = bad] = \frac{10}{3}$$

$$E[N_{good}|X_0 = good] = 0.7 * 1 + 0.3 * (1 + E[N_{good}|X_0 = bad])$$
$$E[N_{good}|X_0 = good] = 0.7 * 1 + 0.3 * (1 + \frac{10}{3})$$
$$= 2$$

$$\pi_{good} = \frac{1}{E[N_{good}|X_0 = good]} = \frac{1}{2}$$
$$\pi_{bad} = 1 - \frac{1}{2} = \frac{1}{2}$$

Problem 1 continued on next page...

So the long-run proportion of time that the weather has been good is $\frac{1}{2}$.

$$\begin{aligned}\mathrm{E}[X] &= (\frac{1}{2} * \frac{1}{12} + \frac{1}{2} * \frac{1}{3}) * 10 * 365 \\ &= 760\end{aligned}$$

(c) Lets show the total number of flowers in three days by X. We have:

$$P(X=2) = P(FFN) + P(FNF) + P(NFF)$$

We will use forward algorithm to calculate each of the probabilities. $P(FFN)$:

$$\begin{aligned}
\alpha_1(G) &= P(F, q_1 = G) = \pi_G * P(F|G) = 1 \times \frac{1}{12} = 0.0833 \\
\alpha_1(B) &= P(F, q_1 = B) = \pi_B * P(F|B) = 0 \\
\alpha_2(G) &= P(FF, q_2 = G) = (\alpha_1(G) \times a_{GG} + \alpha_1(B) \times a_{BG}) * P(F|G) \\
&= (\frac{1}{12} \times \frac{7}{10} + 0) \times \frac{1}{12} = 0.0048 \\
\alpha_2(B) &= P(FF, q_2 = B) = (\alpha_1(G) \times a_{GB} + \alpha_1(B) \times a_{BB}) * P(F|B) \\
&= (\frac{1}{12} \times \frac{3}{10} + 0) \times \frac{1}{3} = 0.0083 \\
\alpha_3(G) &= P(FFN, q_3 = G) = (\alpha_2(G) \times a_{GG} + \alpha_2(B) \times a_{BG}) * P(N|G) \\
&= (0.0048 \times \frac{7}{10} + 0.0083 \times \frac{3}{10}) \times \frac{11}{12} = 0.0053 \\
\alpha_3(B) &= P(FFN, q_3 = B) = (\alpha_2(G) \times a_{GB} + \alpha_2(B) \times a_{BB}) * P(N|B) \\
&= (0.0048 \times \frac{3}{10} + 0.0083 \times \frac{7}{10}) \times \frac{2}{3} = 0.0048 \\
P(FFN) &= \alpha_3(G) + \alpha_3(B) = 0.0053 + 0.0048 = 0.01027
\end{aligned}$$

I have used a python code to calculate probabilities for the other two sequences:

$$\begin{aligned}P(FNF) &= 0.01277 \\ P(NFF) &= 0.03214\end{aligned}$$

So:

$$\begin{aligned}P(X=2) &= P(FFN) + P(FNF) + P(NFF) \\ &= 0.01027 + 0.01277 + 0.03214 \\ &= 0.05518\end{aligned}$$

(d) We want $\arg\max_{Q_1Q_2Q_3} P(q_1 = Q_1, q_2 = Q_2, q_3 = Q_3|NFN)$.
We have:

$$\begin{aligned}P(q_1 = Q_1, q_2 = Q_2, q_3 = Q_3|NFN) &= P(Q_1Q_2Q_3|NFN) \\ &= \frac{P(Q_1Q_2Q_3, NFN)}{P(NFN)} \\ &= \frac{P(NFN|Q_1Q_2Q_3)P(Q_1Q_2Q_3)}{P(NFN)}\end{aligned}$$

Since $P(NFN)$ is constant with respect to states, we only need to find the Q_is that maximize $P(NFN|Q_1Q_2Q_3)P(Q_1Q_2Q_3)$. We can simply calculate this value for all 2^2 options (Since day 1 is certainly good as it is stated in the question).

We want $\arg\max_{Q_1 Q_2 Q_3} P(q_1 = Q_1, q_2 = Q_2, q_3 = Q_3|NFN)$.

$$\begin{aligned}
P(NFN|GGG)P(GGG) &= P(N|G) \times P(F|G) \times P(N|G) \times \pi_G \times a_{GG} \times a_{GG} \\
&= 0.03431 \\
P(NFN|GGB)P(GGB) &= 0.01069 \\
P(NFN|GBG)P(GBG) &= 0.02520 \\
P(NFN|GBB)P(GBB) &= 0.04277
\end{aligned}$$

So the answer is good-bad-bad.

Problem 2

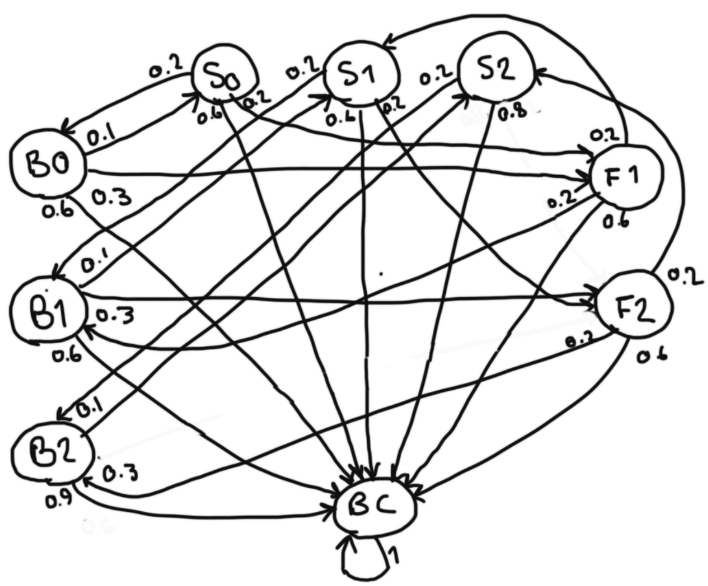
Q. A traveler travels from town to town and never stops! The probability of traveling into each town based on the previous location is presented in the following table. Assuming the traveler is now in Bangs, what is the probability of visiting Fries twice in four next travels without visiting Bluff and Cool? (Highly recommended! Do the computations with a computer!)

States	Bangs	Speed	Fries	Bluff	Cool
Bangs	0	0.1	0.3	0.4	0.2
Speed	0.2	0	0.2	0.1	0.5
Fries	0.2	0.2	0	0.2	0.4
Bluff	0.2	0.4	0.2	0	0.2
Cool	0.3	0.5	0.1	0.1	0

A.
We can view Bluff and Cool as absorbing states, so that if we enter them, we never leave them. Also, for making computations easier, we can merge Bluff and Cool into one state named BC:

States	Bangs	Speed	Fries	BC
Bangs	0	0.1	0.3	0.6
Speed	0.2	0	0.2	0.6
Fries	0.2	0.2	0	0.6
BC	0	0	0	1

Now we can create a new Markov Chain as follows:



B0: Be in Bangs and have never visited Fries
B1: Be in Bangs and have visited Fries only once
B1: Be in Bangs and have visited Fries exactly twice
S0: Be in Speed and have never visited Fries
S1: Be in Speed and have visited Fries only once

Problem 2 continued on next page...

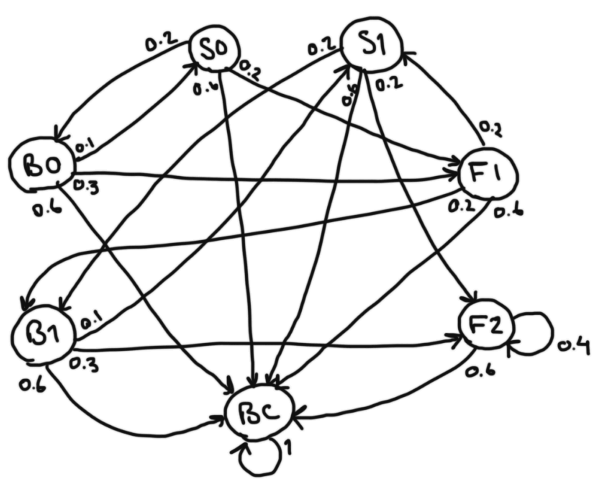
S2: Be in Speed and have visited Fries exactly twice
F1: Be in Fries and have visited Fries only once
F2: Be in Fries and have visited Fries exactly twice
BC: Have visited Bluff or Cool, or have visited Fries more than twice

States	B0	B1	B2	S0	S1	S2	F1	F2	BC
0:B0	0	0	0	0.1	0	0	0.3	0	0.6
1:B1	0	0	0	0	0.1	0	0	0.3	0.6
2:B2	0	0	0	0	0	0.1	0	0	0.9
3:S0	0.2	0	0	0	0	0	0.2	0	0.6
4:S1	0	0.2	0	0	0	0	0	0.2	0.6
5:S2	0	0	0.2	0	0	0	0	0	0.8
6:F1	0	0.2	0	0	0.2	0	0	0	0.6
7:F2	0	0	0.2	0	0	0.2	0	0	0.6
8:BC	0	0	0	0	0	0	0	0	1

If we name the transition matrix as P, the probability of visiting Fries exactly twice and never visiting Bluff or Cool in four next travels by starting from Bangs will be equal to:

$$\begin{aligned} P^4_{B0F2} + P^4_{B0S2} + P^4_{B0B2} &= P^4_{07} + P^4_{05} + P^4_{02} \\ &= 0.00680 + 0.006 + 0.006 \\ &= 0.0188 \end{aligned}$$

If we want to calculate the probability of visiting Fries ”atleast” twice, we can done it using following Markov Model and calculate P^4_{B0F2}.



Problem 3

Q. Suppose we have a random DNA-like sequence generator! DNA is only composed of A, C, G, and T letters. Our generator simply randomizes the next letter based on the current one. The generation probabilities are presented in the following table. Calculate the probability of observing the following patterns in two states: Starting from letter T and in a random location in long time generation:

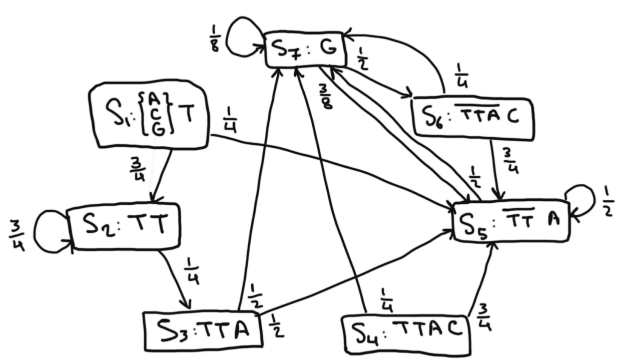
(a) TTAC

(b) GGCGG

States	A	C	G	T
A	$\frac{1}{2}$	0	$\frac{1}{2}$	0
C	$\frac{3}{4}$	0	$\frac{1}{4}$	0
G	$\frac{3}{8}$	$\frac{1}{2}$	$\frac{1}{8}$	0
T	$\frac{1}{4}$	0	0	$\frac{3}{4}$

A.

(a) Since the probability of going to C from A is zero, the TTAC pattern cannot be observed in neither of two conditions. We can also create a Markov Chain for this sequence:



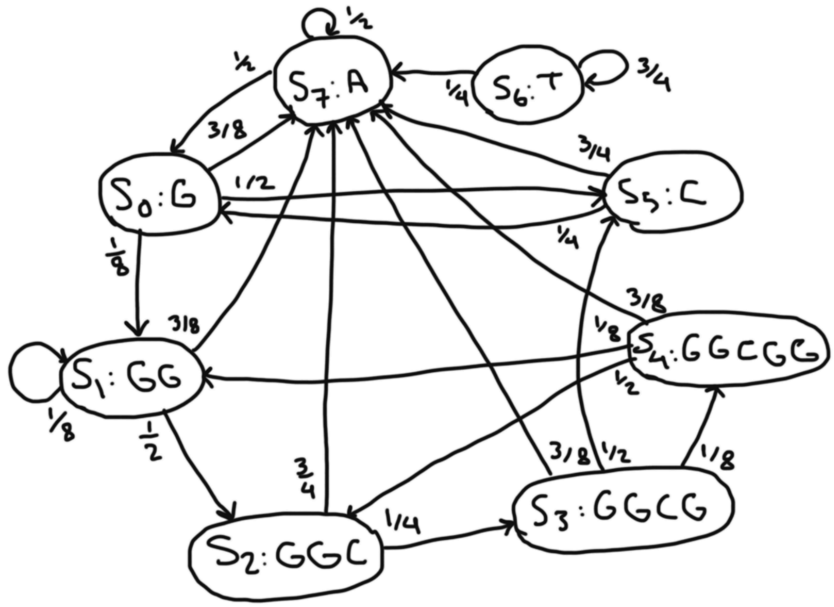
($\overline{XYZ}$ means any pattern except XYZ.)

S_4 is the goal state and as we can see, there is no way to reach this goal.

(b) We can create a Markov chain as follows:

	S_0	S_1	S_2	S_3	S_4	S_5	S_6	S_7
S_0	0	1/8	0	0	0	1/2	0	3/8
S_1	0	1/8	1/2	0	0	0	0	3/8
S_2	0	0	0	1/4	0	0	0	3/4
S_3	0	0	0	0	1/8	1/2	0	3/8
S_4	0	1/8	1/2	0	0	0	0	3/8
S_5	1/4	0	0	0	0	0	0	3/4
S_6	0	0	0	0	0	0	3/4	1/4
S_7	1/2	0	0	0	0	0	0	1/2

Problem 3 continued on next page...



If we calculate P to the power of a large number, we will get following probabilities:

$$\pi = [0.28645833, 0.04101562, 0.02083333, 0.00520833, 0.00065104, 0.14583333, 0., 0.5]$$

So, the probability of seeing the pattern starting from a random state in the long run will be $\pi_0 = 0.28645833$.

Problem 4

Q. Prove that a random walk on an undirected, connected graph is aperiodic if and only if the graph is not bipartite.

A.
We will first prove if the random walk is aperiodic, then the graph is not bipartite.
Proof by contradiction: Assume the graph is bipartite. So we can separate the graph nodes into two sets, A and B, in which no edges exist inside nodes of one set. Suppose we are in state i in set A at time 0; we will always be in a node in set B at odd times. Thus, we can only be in state i again after an even number of times. This is true for all states; thus, the random walk on this graph is not aperiodic.

Now we should prove that if the graph is not bipartite, it is aperiodic.
We know that if a graph is not bipartite, it has an odd cycle. Let's show its length with C. For a state i, we will show the length of the path from i to one of the nodes in the cycle by n_i. (The graph is connected, and thus, this path exists.) Now consider $N_i = 2n_i + C_i$. We will show that for each $M \geq N_i$, we can return to node i after M time steps. To do this, if M is even, we can simply go to one of the i's neighbors and come back $\frac{M}{2}$ times. If M is odd, we can go to the cycle, go around it, and come back in $2n_i + C_i$ time steps; then we can go to one of the i's neighbors and come back $\frac{M-(2n_i+C)}{2}$ times. So i is aperiodic. This is true for all states; thus, the graph is aperiodic.

Problem 5

Q. Show that every Markov chain with $M < \infty$ states contains at least one recurrent set of states.

A. Assume all the states are transient. We know that a transient state is visited only a finite number of times since for a transient state i, $\sum_{n=1}^{\infty} P_{ii}^n < \infty$. So for each transient state i there is a time t_i after which we never visit state i. Since $M < \infty$, we can consider the maximum of t_i for all states and name it T. So after time T, none of the states will be revisited, but this is impossible since we need to be in one state. Thus, there must be a class of recurrent states in a finite-state Markov chain.

Problem 6

Q.

(a) Show that an ergodic Markov chain with M states must contain a cycle with $\tau < M$ states.

(b) Let X be a fixed state on this cycle of length τ. Let $T(m)$ be the set of states accessible from X in m steps. Show that:

$$\forall m \geq 1; T(m) \subseteq T(m+\tau)$$

A.

(a) Since an ergodic Markov chain is recurrent, there exist atleast one cycle. Consider the shortest length cycle. If it has $\tau < M$ states, then we are done. Suppose it has $\tau = M$ states. If there wouldn't be any edge between states other than the cycle edges, then this Markov chain is periodic with period M since we can only get back to one state after M steps. This contradicts with the Markov chain being ergodic.

(b) Consider a path of length m to a state A in $T(m)$. We can simply go across the cycle and come back to X in τ steps and then go to A in m steps. So A is in $T(m+\tau)$. Since we have shown that $A \in T(m) => A \in T(m+\tau)$, so $T(m) \subseteq T(m+\tau)$ for all $m \geq 1$.