Discussion 06D

Markov Chains Intro II

Note 22

Recall that a Markov chain is defined with the following: the state space $\mathscr{X}$, the transition matrix P, and the initial distribution π_0. This implicitly defines a sequence of random variables X_n with distribution π_n, which denote the state of the Markov chain at timestep n. This sequence of random variables also obey the Markov property: the transition probabilities only depend on the current state, and not any prior states.

A before B: Suppose we want to compute the probability of reaching state A before reaching state B. To compute this quantity, let $\alpha(i) = \mathbb{P}[A \text{ before } B \mid \text{at } i]$. Then, we have:

$$\begin{aligned}
\alpha(A) &= 1 \\
\alpha(B) &= 0 \\
\alpha(i) &= \sum_j P(i,j)\alpha(j)
\end{aligned}$$

Here, we use the law of total probability when computing $\alpha(i)$; we consider all possible transitions *out of* state i. These are called the **first step equations (FSE)**.

Hitting time: Suppose we want to compute the expected number of steps until you reach state A. To compute this quantity, let $\beta(i) = \mathbb{E}[\text{steps until } A \mid \text{at } i]$. Then, the first step equations become:

$$\begin{aligned}
\beta(A) &= 0 \\
\beta(i) &= 1 + \sum_j P(i,j)\beta(j)
\end{aligned}$$

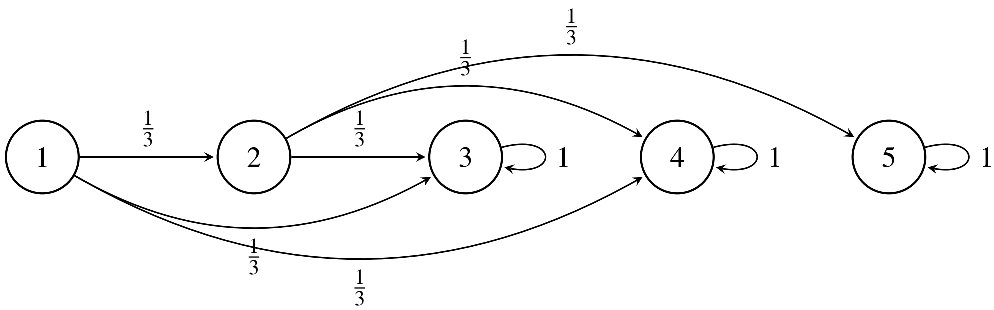
Here, we use the law of total expectation when computing $\beta(i)$; we consider all possible transitions *out of* state i.

1 Skipping Stones

Note 22

We consider a simple Markov chain model for skipping stones on a river, but with a twist: instead of trying to make the stone travel as far as possible, you want the stone to hit a target. Let the set of states be $\mathscr{X} = \{1,2,3,4,5\}$. State 3 represents the target, while states 4 and 5 indicate that you have overshot your target. Assume that from states 1 and 2, the stone is equally likely to skip forward one, two, or three steps forward. If the stone starts from state 1, compute the probability of reaching our target before overshooting, i.e. the probability of $\{3\}$ before $\{4,5\}$.

Solution: Here is the Markov Chain we are working with:



Let $\alpha(i)$ denote the probability of reaching the target before overshooting, starting at state i. Then:

$$
\begin{aligned}
\alpha(5) &= 0 \\
\alpha(4) &= 0 \\
\alpha(3) &= 1 \\
\alpha(2) &= \frac{1}{3}\alpha(3) + \frac{1}{3}\alpha(4) + \frac{1}{3}\alpha(5) = \frac{1}{3} \\
\alpha(1) &= \frac{1}{3}\alpha(2) + \frac{1}{3}\alpha(3) + \frac{1}{3}\alpha(4) = \frac{1}{9} + \frac{1}{3}
\end{aligned}
$$

Therefore, $\alpha(1) = 1/9 + 1/3 = 4/9$.

2 Three Tails

Note 22

You flip a fair coin until you see three tails in a row.

(a) What is the average number of timesteps until you get TTT?

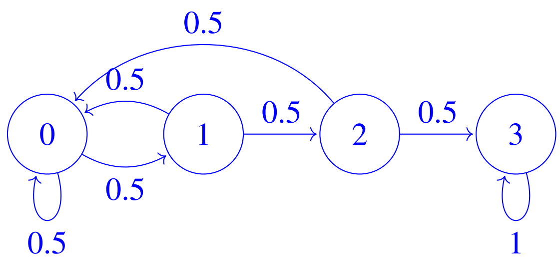
(b) What is the average number of heads that you'll see until you get TTT? *Hint*: Modifying your equations from part (a) slightly to solve the original question.

Solution:

(a) **Solution 1**: We can model this problem as a Markov chain with the following states:

- 0: Currently, we've seen 0 consecutive tails.

- 1: Currently, we've seen 1 consecutive tail.
- 2: Currently, we've seen 2 consecutive tails.
- 3: Currently, we've seen 3 consecutive tails. This concludes the game.



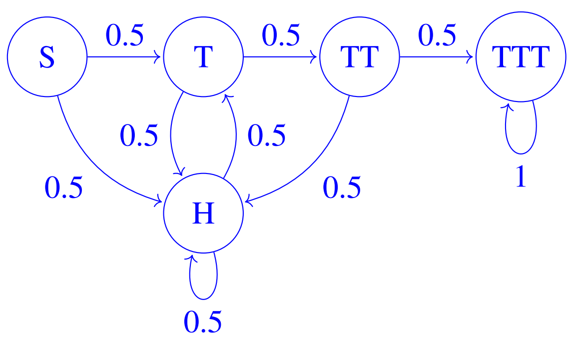
Here are the hitting time equations for the number of **timesteps**, defining $\beta(S)$ as the expected number of timesteps to reach state 3, given that we're currently at state S. Note that we definitively start at state 0.

$$\beta(0) = 1 + 0.5\beta(0) + 0.5\beta(1) \tag{1}$$
$$\beta(1) = 1 + 0.5\beta(0) + 0.5\beta(2) \tag{2}$$
$$\beta(2) = 1 + 0.5\beta(0) + 0.5\beta(3) \tag{3}$$
$$\beta(3) = 0 \tag{4}$$

Solving yields $\beta(0) = 14$, $\beta(1) = 12$, $\beta(2) = 8$, and $\beta(3) = 0$. On average, it'll take us 14 timesteps (14 coin flips) before getting TTT.

Solution 2: Alternatively, we can model this problem as a Markov chain with the following states:

- S: Start state, which we are only in before flipping any coins.
- H: We see a head, which means no streak of tails currently exists.
- T: We've seen exactly one tail in a row so far.
- TT: We've seen exactly two tails in a row so far.
- TTT: We've accomplished our goal of seeing three tails in a row and stop flipping.



We can write the first step equations and solve for $\beta(S)$. The equations are as follows:

$$\beta(S) = 1 + 0.5\beta(T) + 0.5\beta(H) \tag{5}$$
$$\beta(H) = 1 + 0.5\beta(H) + 0.5\beta(T) \tag{6}$$
$$\beta(T) = 1 + 0.5\beta(TT) + 0.5\beta(H) \tag{7}$$
$$\beta(TT) = 1 + 0.5\beta(H) + 0.5\beta(TTT) \tag{8}$$
$$\beta(TTT) = 0 \tag{9}$$

From (6), we see that
$$0.5\beta(H) = 1 + 0.5\beta(T)$$
and can substitute that into (7) to get
$$0.5\beta(T) = 0.5\beta(TT) + 2.$$

Substituting this into (8), we can deduce that $\beta(TT) = 8$. This allows us to conclude that $\beta(T) = 10$, $\beta(H) = 12$, and $\beta(S) = 14$.

(b) **Solution 1**: Now, we can modify the first step equations from (a) and solve for $\beta_H(S)$ defined as the expected number of heads we see before we reach state 3, given that we're currently at state S. We will provide some motivation for how to set up the equations. In the previous equations, we had $\beta(0) = 1 + 0.5\beta(0) + 0.5\beta(1)$ which represents that we take a timestep, and move into either state 0 or state 1 with equal probability by flipping a coin. We can think of the same equation in an alternate viewpoint: consider instead flipping the coin to decide whether to move into state 0 or state 1, and in either case, we took a timestep, so we add 1 to our counter. Thus, the same equation can be rewritten as $\beta(0) = 0.5(1 + \beta(0)) + 0.5(1 + \beta(1))$.

Now consider how we may edit these equations to account for the number of **heads**, instead of the number of **timesteps**. We can modify the same equation to be $\beta_H(0) = 0.5(1 + \beta_H(0)) + 0.5(\beta_H(1))$ we only add 1 to our counter if we move into state 0, and we do not add anything to our counter if we move into state 1. We can set up the whole system of equations as follows:

$$\beta_H(0) = 0.5(1 + \beta_H(0)) + 0.5(\beta_H(1)) \tag{10}$$
$$\beta_H(1) = 0.5(1 + \beta_H(0)) + 0.5(\beta_H(2)) \tag{11}$$
$$\beta_H(2) = 0.5(1 + \beta_H(0)) + 0.5(\beta_H(3)) \tag{12}$$
$$\beta_H(3) = 0 \tag{13}$$

Solving these equations yields $\beta_H(0) = 7$, $\beta_H(1) = 6$, $\beta_H(2) = 4$. This is the same answer we got before, resulting in an expected number of 7 heads before we see three tails in a row. On average, we expect to see 7 heads before flipping three tails in a row.

Solution 2:

We can write the first step equations and solve for $\beta(S)$, only counting heads that we see since we are not looking for the total number of flips. The equations are as follows:

$$\beta(S) = 0.5\beta(T) + 0.5\beta(H) \tag{14}$$
$$\beta(H) = 1 + 0.5\beta(H) + 0.5\beta(T) \tag{15}$$
$$\beta(T) = 0.5\beta(TT) + 0.5\beta(H) \tag{16}$$
$$\beta(TT) = 0.5\beta(H) + 0.5\beta(TTT) \tag{17}$$
$$\beta(TTT) = 0 \tag{18}$$

From (15), we see that

$$0.5\beta(H) = 1 + 0.5\beta(T)$$

and can substitute that into (16) to get

$$0.5\beta(T) = 0.5\beta(TT) + 1.$$

Substituting this into (17), we can deduce that $\beta(TT) = 4$. This allows us to conclude that $\beta(T) = 6$, $\beta(H) = 8$, and $\beta(S) = 7$.

Note on Symmetry: You may have noticed that the expected number of heads (7) we see before we see three tails in a row is half the expected number of timesteps (14). We can't directly say that this follows as a consequence of heads and tails being symmetric, as not every coin flip in our experiment ends up having the same distribution when conditioned on the fact that we end on TTT. For example, if we end on TTT, we know that the last three flips must all be tails, and have no chance of being heads. However, symmetry still ends up being applicable. What if we were to revisit the previous set of equations, but instead calculate the expected number of **tails** we see before we see three tails in a row? We can set up the equations as follows:

$$\beta_T(0) = 0.5(\beta_T(0)) + 0.5(1 + \beta_T(1)) \tag{19}$$
$$\beta_T(1) = 0.5(\beta_T(0)) + 0.5(1 + \beta_T(2)) \tag{20}$$
$$\beta_T(2) = 0.5(\beta_T(0)) + 0.5(1 + \beta_T(3)) \tag{21}$$
$$\beta_T(3) = 0 \tag{22}$$

The $(+1)$ term ends up being applied to the latter term, as we are counting tails instead of heads. Now, we notice that these two sets of equations are actually the exact same, when we distribute out the $0.5 * 1$ term from every equation. Thus, the expected number of tails we see before we see three tails in a row is the same as the expected number of heads we see before we see three tails in a row, which is 7.