Scribe Notes: Debbie Perouli

1 Ways to Describe a Stochastic Process

We will use the biased coin example to illustrate three ways with which we can describe a stochastic process, a Markov Chain (*MC*) in particular.

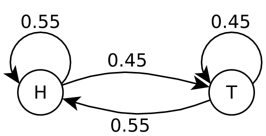
Probabilistic

Random variable X represents the outcome of the experiment where we toss a specific coin once. There is 0.55 probability for the result to be heads (H) and 0.45 to be tails (T).

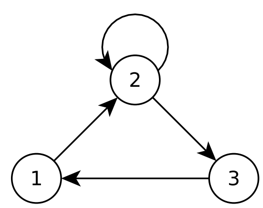
$$X_i = \begin{cases} H & \text{with probability } 0.55 \\ T & \text{with probability } 0.45 \end{cases}$$

Random Walk

Figure 1a shows a topological description of a sequence of coin tossings. Figure 1b illustrates a process which is relatively straight forward to define as a graph, but rather complicated to describe probabilistically.



(a) Biased Coin.



(b) It is not trivial to define the random variables which describe this system.

Figure 1: Stochastic Processes described as Random Walks.

Transition Matrix

A matrix description of the random walks 1a and 1b is P_a and P_b respectively. Variable x in P_b denotes a non-zero value. For example, $P_b(2,2)$ is the probability of state 2 to remain in this state and $P_b(2,3)$ is the probability of moving to state 3.

$$P_a = \begin{array}{c} \begin{matrix} H & \;\; T \end{matrix} \\ \begin{bmatrix} 0.55 & 0.45 \\ 0.55 & 0.45 \end{bmatrix} \end{array} \begin{matrix} H \\ T \end{matrix} \qquad P_b = \begin{bmatrix} 0 & x & 0 \\ 0 & x & x \\ x & 0 & 0 \end{bmatrix}$$

2 Strongly Connected Components

In the past lecture we mentioned different types of Markov states: absorbing, transient, recurrent, periodic. We will now introduce them in terms of strongly connected components of the random walk graph. We will use the symbol C_k for the strongly connected component k of the graph.

In a directed graph $G(V, E)$, where $i, j \in V$, the following property must hold: *if $i, j \in C_k$, then there is a path from node i to node j and there is a path from node j to node i as well.* The left column of Figure 2 highlights in dotted boxes the strongly connected components of two graphs.

In terms of random walks, if $i, j \in C_k$ then there exist $n_1, n_2 \in \mathbb{Z}^+$ such that:

$$\Pr(X_{n_1} = j \mid X_0 = i) > 0 \text{ and } \Pr(X_{n_2} = i \mid X_0 = j) > 0$$

In other words, the probability of visiting state j in n_1 steps after visiting state i is positive, and so is the probability of visiting state i in n_2 steps after state j.

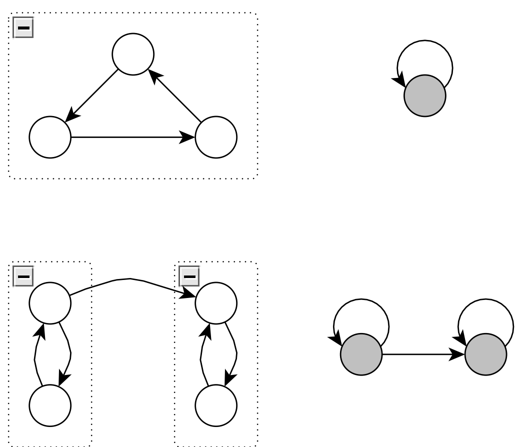


Figure 2: Strongly Connected Components of two Graphs (left) and the corresponding Component Graphs (right).

Component Graph

Let $G_C(V_C, E_C)$ be the component graph of graph $G(V, E)$, where:

$V_C = \{C_i : i = 1..|V_C|\}, E_C = \{(C_i, C_j) : u \in C_i, v \in C_j, (u, v) \in E\}$.

The right part of Figure 2 shows the component graphs of the two examples on the left. Note that we characterize components in the same way we characterize states of the original graph, e.g. there can be absorbing or transient components.

An important property of G_C is that it is a dag (directed acyclic graph). To see this, suppose there is a cycle between two components as shown in Figure 3a. According to the definition of a component, though, C_i and C_j should form a single component. Therefore, there cannot be a cycle.

(a) The Component Graph cannot have a Cycle.

(b) Invalid Markov Chain graph.

Figure 3: Two Special Cases of Component Graphs.

A consequence of this property is that G_C and G have a set of starting vertices and a set of terminal ones as illustrated in Figure 4. The terminal nodes of a Markov Chain are required to have self-loops (if one reaches the terminal state, one never leaves it). A graph like Figure 3b is not a valid one for a Markov Chain, since there needs to be a transition out of each state.

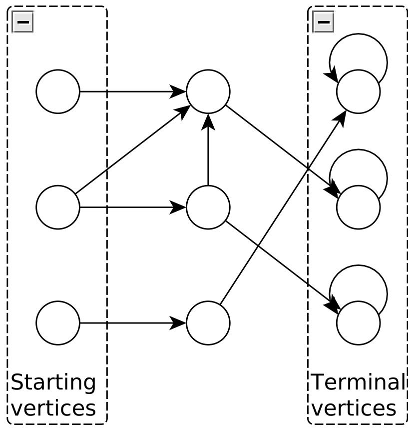


Figure 4: A Markov Chain Graph consists of starting and terminal vertices/components. The terminal nodes need to have self-loops.

3 States

Recurrent

$\Pr(X_n \text{ revisits } i \mid X_0 = i) = 1$

The terminal components of a Markov Chain belong to some type of recurrent state. The possible types are: absorbing, periodic, aperiodic.

Absorbing

$\Pr(X_{n+1} = i \mid X_n = i) = 1$. Equivalently:

$$P_{ij} = \begin{cases} 1 & \text{if i = j} \\ 0 & \text{otherwise} \end{cases}$$

Periodic

The graph can be partitioned in classes of states K_i as shown in Figure 5. Each of the classes contains no edges inside it, so the transition matrix has a diagonal of zeros. The cycles have a period p. Figure 6 shows a simple example. Note that it still may not be possible to reach the state in p steps. Say that for a particular state, n can take the values 9, 12, 15, ... Although the period is 3, there are values of n (like 3 and 6) for which one cannot revisit the state.

Aperiodic

It is a periodic state with period 1. A sufficient condition for a state to be aperiodic is to have a self-loop.

Transient

$\Pr(X_n = i \mid X_0 = i) < 1$

The non-terminal components of a graph are transient. More formally, we introduce the event:

$E_i^t = \Pr(X_t = i, X_1, ...X_{t-1} \neq i \mid X_0 = i)$. Then, $\sum_{i=1}^{\infty} E_i^t < 1$

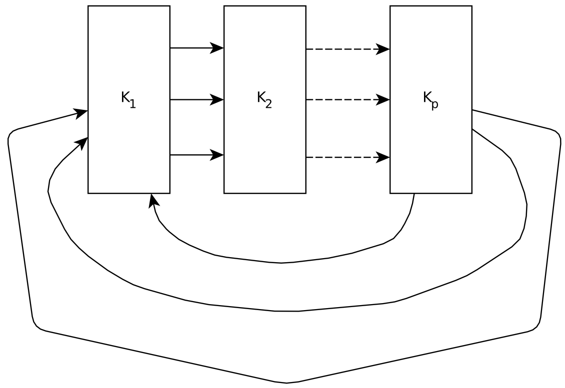


Figure 5: Classes of states which form cycles of period p. There are no edges inside each class K_i.

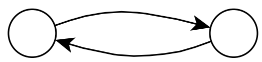


Figure 6: Example of a Markov Chain with period 2.

3.1 Quiz

We were given a random walk and were asked to identify the components and their type. The solution is shown in Figure 7. For instance, the only absorbing component of the graph is a terminal component of size 1! (factorial).

3.2 Stochastic Matrix Permutations

See the discussion in Lecture 8 (last lecture).

4 Λ-Step Transition Probabilities

$\Pr(X_n = j \mid X_0 = i) =$

$\sum_r \Pr(X_n = j \mid X_{n-1} = r) \Pr(X_{n-1} = r \mid X_0 = r) =$

(apply the Markov property that history doesn't matter)

$\sum_r P_{rj} \Pr(X_{n-1} = r \mid X_0 = r) =$

(repeat the argument)

$\sum_{r,s} P_{rj} P_{sj} \Pr(X_{n-2} = s \mid X_0 = r) =$

$\sum_s (P^2)_{sj} \Pr(X_{n-2} = r \mid X_0 = r) \Rightarrow$

$\Pr(X_n = j \mid X_0 = i) = (P^n)_{ij}$

For the starting state, we have:

$X_i^{(0)} = \Pr(X_0 = i)$

$\mathbf{x}^{(o)T} P^n = \sum_i X_i^{(0)} P_{ij}^n$

$(P^n)^T \mathbf{x}^{(0)} = \sum_i X_i^{(0)} P_{ij}^n$

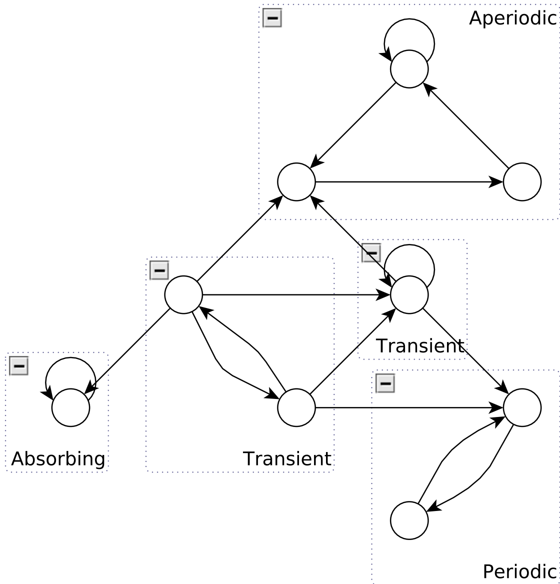
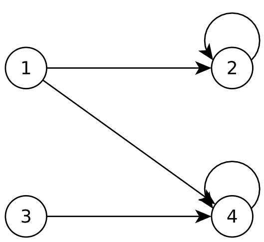


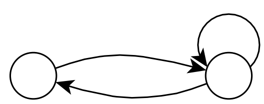
Figure 7: The components and their type.

5 Does $\mathbf{x}_0$ matter, if n is "big enough"?

It depends. Sometimes yes, sometimes no. Figure 8 shows two examples. If one starts at node 3, then one ends up at 4. However, starting at node 1 does not guarantee arriving at 4. The graph in Figure 8b has a single recurrent state that is aperiodic. For this reason, the initial state does not matter.



(a) $\mathbf{x}_0$ matters.



(b) $\mathbf{x}_0$ does not matter.

Figure 8: Examples showing the role of $\mathbf{x}_0$.

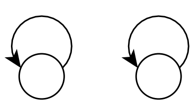


Figure 9: Two classes that do not communicate.

6 When does P^n converge to "something" simple?

P^n is the matrix for n state transitions. An example of two classes that don't communicate is given in Figure 9. In this case $P = I$.

We want: $P^n \to \mathbf{e}\boldsymbol{\pi}^T$ (rank 1 matrix)

$P^T\boldsymbol{\pi} = \boldsymbol{\pi}$ (eigenvalue-vector problem)

P^T has a unique eigenvalue $\iff$ it converges to something simple