Problem set 3

1. Ehrenfest urn model: We have n identical balls and 2 urns labelled A and B. At each time step, we choose one of the n balls uniformly at random and move it to the other urn. (a) Let X_t denote the number of balls in urn A after the t^{th} time step. Briefly explain why X_t, $t \geq 0$ is a Markov chain. Draw an arrow diagram to describe the states and transition probabilities. (b) Specify its communicating classes and which states are transient and recurrent. Compute all of its invariant distributions.

2. Wright-Fisher model of population genetics: A gene has two alleles (comes in one of two types) denoted A and a. (For example, A could be the allele for brown eyes and a for blue eyes.) Each individual has two copies of each gene, but in this problem we are only interested in the total number of each allele in the population. The model assumes random mating and constant population size, and leads to the following description for the numbers of each allele. Let the total population size (of genes) be N and let N_t denote the number of A alleles in generation t. Generation $t+1$ is obtained as follows. Each gene in generation $t+1$ is sampled uniformly at random from the genes in generation t. A copy of the sampled gene is placed in generation $t+1$ (and the original is retained). This process is repeated N times to construct the population in generation $t+1$.
(a) Is N_t, $t \geq 0$, a Markov chain? If not, can you modify the state space to obtain a Markov chain?
(b) What are the communicating classes, transient states and recurrent states of your Markov chain?
(c) What are its invariant distributions?

3. Polya urn model: An urn has n white balls and m black balls to start with. At each time step, you draw a ball from the urn at random, then replace it and put in one more ball of the same colour. Let X_t denote the fraction of white balls at the t^{th} time step.
(a) Is X_t a Markov chain? If not, can you find a state space which makes

the process described above a Markov chain?
(b) Either compute the invariant distribution of this Markov chain or explain briefly why it doesn't have one.

4. We say that a random variable X has a Geometric distribution with parameter p, written $X \sim Geom(p)$ if

$$P(X = k) = p(1-p)^k, \quad k = 0, 1, 2, \ldots$$

(a) If $X \sim Geom(p)$, compute $P(X \geq n)$ and $P(X \geq m + n | X \geq n)$ for all $m, n \geq 0$.
(b) A wireless communication channel can be in one of two states, Good or Bad. When it enters the Good state, it remains in that state for a further $Geom(p)$ number of time slots before moving to the Bad state. Likewise, on entering the bad state, it remains in it for a further $Geom(q)$ number of time slots before moving to the Good state. Let X_t denote the channel state during the t^{th} time slot. Explain why X_t, $t \geq 0$ is a Markov chain. Draw an arrow diagram describing its states and transition probabilities and compute its invariant distribution.

5. Let $G = (V, E)$ be a connected undirected graph with finite node set V and edge set E. A node w is called a neigbour of a node v if there is an edge between them, i.e., if $(v, w) \in E$. The degree of a node v, denoted $d(v)$, is the number of neighbours v has.

Consider the following continuous-time random walk on the graph. If the walker enters some node v at time t, then it stays there for an exponentially distributed time with parameter $d(v)$, and then moves randomly to one of the neighbours of v.
(i) Each of the neigbours of v is equally likely to be chosen.
(ii) The choice of neighbour is independent of all past choices.
(iii) The choice neighbour is independent of the time spent in state v.

(a) Describe the above scenario using a Markov chain. This means saying what the possible states are, and what the transition rate is between every pair of states. It is fine to do this in words rather than in mathematical notation.
(b) Which of the properties (i),(ii) and (iii) are necessary for this to be a Markov chain?
(c) Show that the Markov chain is reversible and use this fact to compute its invariant probability distribution π.

6. You are given a directed graph $G = (V, E)$. Associated with each edge

$e \in E$ is a failure probability probability p_e. Edges fail independently of each other.
(a) What algorithm would you use to find the most reliable path from a given node s to another given node t? By most reliable, we mean the path with the smallest failure probability. A path fails if any edge on that path fails. (To answer this question, you can either just name one of the standard algorithms we have studied in class or come up with a new algorithm of your own.)
(b) Explain clearly how you would transform the stated problem to a form that lets you use the algorithm chosen in part (a).

7. [Hall's marriage theorem] Say there are n men and n women. Suppose that, for each k between 1 and n, *every* subset of k men know at least k different women between them. (This might not be evenly spread. For example, 1 man might know $k-1$ different women, while the other $k-1$ men each know only the same woman.) Hall's theorem says that is possible to arrange n marriages such that each man is married to a woman he knows.

(a) First, express this as an instance of a maximum flow problem. To do this, you must say what nodes there are, what arcs there are, and what the capacity of each arc is.
(b) Next, show that *every* cut has capacity at least n.
(c) Hence, argue why the theorem is true.