CS112: Computer System Modeling Fundamentals
Homework 5

Due Friday, June 3, 4pm (in section)

Please refer to the course academic integrity policy for collaboration rules. Remember that you will be graded on both the correctness and the clarity of your solutions.

This assignment will be collected in the last discussion section on June 3, which will be a review session for the final. If you cannot make the section, you may submit it in class on June 2.

1. A very common prior used in Bayesian statistics to capture subjective beliefs is the *beta* prior. A beta distribution is parameterized by two values, a and b, which must both be at least 1. A beta random variable with parameters a and b has PDF

   $$f_\Theta(\theta) = \alpha\theta^{a-1}(1-\theta)^{b-1}$$

   for $\theta \in [0, 1]$, where α is the appropriate normalizing constant and depends on a and b. A beta random variable Θ has mean $a/(a+b)$.

   We would like to estimate the bias of a coin based on some observed data and our own subjective beliefs. Suppose that we assume a beta prior with parameters a and b on the coin's bias, and observe a sequence of coin flips with h heads and t tails.

   (a) Show that the *posterior* distribution of Θ is also a beta distribution with new parameters a' and b'. What are the values of a' and b' in terms of a, b, h, and t?

   Hint: You should not need to calculate the value of the normalizing constant for the posterior distribution exactly. Instead, use the fact that the integral from $-\infty$ to ∞ of any valid PDF must be 1, and look for terms that don't depend on θ.

   (b) What is the mean of θ with respect to the posterior? If you want your prior to have a strong influence on this mean, would you make a and b small or large? Why?

2. Construct two distinct directed acyclic graphs over four nodes (A, B, C, and D) such that each graph has exactly four edges, and the two graphs imply precisely the same set of independence and conditional independence assumptions when interpreted as Bayesian networks.

3. The *Markov Blanket* of a node X in a Bayesian network is the set containing the parents of X, the children of X, and the other parents of the children of X. Let M be the Markov Blanket of X. Prove that the set M d-separates X from any node $Y \notin M \cup \{X\}$. That is, prove that for any node $Y \notin M \cup \{X\}$, X and Y are d-separated given evidence set M.

4. An instruction for a RISC CPU is either R-type (R), Store/load (S), or Jump-type (J). If the current instruction is R-type, then the next instruction will be R, S, or J with probabilities 0.5, 0.4, and 0.1, respectively. If the current instruction is Store/load, then the probabilities for next

instruction to be R, S, or J are 0.3, 0.4, and 0.3, respectively. If the current instruction is Jump-type, then the next instruction will be R, S, or J with probabilities 0.2, 0.3, 0.5, respectively.

Suppose we would like to model this scenario using a Markov chain.

(a) Draw the state transition diagram.

(b) Find the transition probability matrix.

(c) In the long run, what proportion of instructions are of each type?

5. Consider a Markov chain with the following transition probability matrix.

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0.7 & 0 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 & 0 & 0 \\ 0.99 & 0 & 0.01 & 0 & 0 & 0 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.1 & 0.5 \end{pmatrix}$$

(a) Draw the state transition diagram of the Markov chain.

(b) For each of the six states, state whether it is recurrent or transient.

(c) What are the recurrent classes of this Markov chain?

(d) For each recurrent class, state whether it is periodic or aperiodic. If periodic, what is the period?

(e) Let X_t denote the state of the Markov chain at time t. What is $P(X_2 = 2 | X_0 = 4)$?

(f) In the long run, what proportion of the time will the Markov chain be in each state?

6. A workstation tries to transmit frames through the Ethernet. Suppose that whether or not collision occurs in the current transmission depends on the result of the last two transmissions the workstation had. In particular, if collisions occurred in the past two transmissions, then with probability 0.7 a collision will occur in the current transmission; if a collision occurred in the last transmission but not the transmission before the last one, then a collision will occur in the current transmission with probability 0.5; if a collision occurred in the transmission before the last one but not the last one, then one will occur in the current transmission with probability 0.4; if there have been no collisions in the past two transmissions, then a collision will occur in the current transmission with probability 0.2.

(a) Define appropriate states in order to model this scenario as a Markov Chain.

(b) Draw the state transition diagram.

(c) Find the transition probability matrix for the states defined in part (a).

(d) What fraction of frames suffer a collision in the long run?