CS/CNS/EE 154: Artificial Intelligence
Problem Set 3

Handed out: 14 Nov 2010
Due: 24 Nov 2010

---

1 Bayesian Networks [20 points]

As discussed in class, conditional independence properties entailed by a Bayesian Network can be read directly from the graph, using the framework of *d-separation*, where the "d" stands for "directed." Given the Bayesian network in figure 1, which of the following conditional independence statements hold:

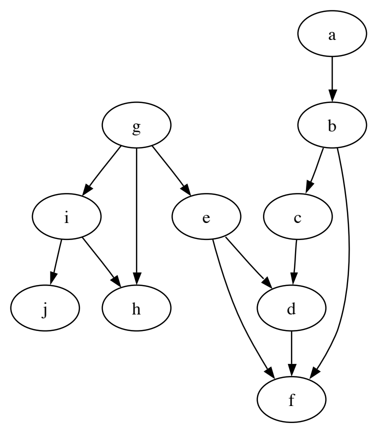

Figure 1: Bayes' net for Problem 1

1. $a \perp f$
2. $a \perp g$
3. $b \perp i \mid f$
4. $d \perp j \mid g, h$
5. $i \perp b \mid h$
6. $j \perp d$
7. $i \perp c \mid h, f$

2 Markov Decision Processes (MDPs) [40 points]

Let us consider the following problem related to probabilistic planning in the PacMan domain. In particular, consider the scenario where one of the PacMen, P1, has not consumed a power pellet, whereas P2 has eaten a power pellet and is chasing after P1.

1. **[20 points]** Suppose the maze is a simple (infinite) chain of nodes, each node labeled with a number: P1 starts at 0, P2 starts at -2. P1 always tries to move away from P2, but only succeeds with probability p, and with probability $1 - p$ it gets stuck (i.e., with probability p, it moves 1 step to the right, from node i to $i + 1$, and gets 1 unit of reward; with probability $1 - p$, it doesn't change its location and gets 0 units of reward). P2 always chases after P1 and never gets stuck. If P2 catches P1, P1 incurs -10 reward in the timestep in which it got caught (and 0 reward in all subsequent time steps). Both P1 and P2 move simultaneously. Write down the state space with the transition probabilities. For a discount factor γ, what is the expected long term future reward as a function p and γ? Calculate its value for $p = .9$ and $\gamma = .95$. *Hint: You may want to consider the relative positions of P1 and P2 instead of their absolute positions when choosing your state representation.*

2. **[20 points]** Now, suppose the maze is a "T", i.e, an (infinitely large) tree, where only one node, the starting node of P1, has degree 3. In the first round, P1 has the choice of either moving "right" and being chased (the same as above); or moving "down" and not being chased. If Pacman moves "down", it will also get stuck with probability $1 - p$ like above, but only incur reward 1/2 for each step moved (which happens with probability p. In all subsequent actions, P1 continues to (attempt to) move in the same direction as in the first round (i.e., once it decides to move right, it has to continue to move right etc.) What is the expected long term future reward in this case, as a function of p and γ? Calculate its value for $p = .9$ and $\gamma = .95$. For these values of p and γ, which initial action should P1 take? For a value of $\gamma = .95$, give an explicit rule on how P1 should choose its initial action as a function of p. Compute the critical (decision-relevant) values of p (you may have to do this numerically).

3 Information Gathering [40 points]

1. **[20 points]** Consider the following dataset (imagine it's part of the Avian Asker problem, where S stands for the species and Y_1, Y_2, Y_3 and Y_4 are the answers to four different questions). Each row in the table is a bird. For a bird that is picked uniformly at random, you try to identify its species by asking questions.

   Calculate the information gain $I(S; Y_i)$ for each i. Which question is the most informative one? Draw the decision tree constructed using the adaptive greedy algorithm which always picks the most informative question conditioned on what it has already observed. Stop expanding the tree after the correct species has been uniquely determined. In practice, do you think it makes sense to always expand the tree until the true species has been uniquely determined, or can you think of a reason to stop earlier?

S	Y_1	Y_2	Y_3	Y_4
1	0	1	0	0
1	1	1	0	1
2	0	0	0	1
1	1	1	0	0
2	0	0	0	0
3	0	1	1	0
3	1	0	1	1
3	1	1	1	0
3	0	1	0	1
4	1	0	1	1
4	0	0	1	1
4	1	0	1	1

Table 1: Dataset

2. **[20 points]** Now, suppose that after you stop asking questions, you have to choose an action a. You can either choose to predict a species $a \in \{1, 2, 3, 4\}$ or to say "don't know" ($a = 0$). If you correctly predict the species, you get 1 unit of reward. If you incorrectly predict the species, you get -5 units of reward. If you say "don't know", you get -1 unit reward. What is the expected value of (perfect) information for each single question? What is the value of information for the set of questions $\{Y_1, Y_4\}$? How about $\{Y_2, Y_3\}$? What does this imply about submodularity of value of information?