Homework Set #7

Due: 11/26/03

1 Problem: Properties of utility functions

Show that the decisions made under the maximum expected utility rule are invariant to scaling and offsetting of the utility function, $U_{new}(s) = \alpha U_{old}(s) + \beta$. What are the ranges of α and β for which the above statement holds?

2 Problem: Decision making in static situations

Consider the alarm and burglary example we discussed in the class. Answer the following questions:

1. Find Markov blankets for every node (random variable) in the network.
2. Compute the value of information of node $Burglary$ given the evidence $Call = No$.
3. Now compute the value of information of node $Call$ given the evidence $Burglary = Yes$.
4. Comment on results in (2) and (3).
5. What is the optimal action assuming we received the following evidence: $Newscast = Quake$.
6. What is the optimal action assuming we received the following evidence: $Newscast = No\_Quake$.
7. Comment on results in (5) and (6).

3 Problem: Sequential decision making

Consider the CS440 class confusion problem we discussed in the recitation session on Nov. 5. Assume that the camera detectors are absolutely accurate and that we know the true state of class confusion. However, unlike the recitation session problem when we could do nothing about the class confusion state, we now have the ability to make the lectures more interesting (or not). If we make a lecture more interesting, the probability that the class will be confused in the next lecture is given by the following conditional probability table:

$$P(C_t|C_{t-1}, a_{t-1} = VERY\_INTERESTING) = \begin{bmatrix} 0.3 & 0.1 \\ 0.7 & 0.9 \end{bmatrix}$$

If the lecture is not very interesting, the confusion model remains the same as in the recitation session.

If the class is confused, the professor gets an immediate reward of 10. If the class is not confused, the reward is 20. However, there is also another "reward" being issued, related to the number of student questions. The less confused the students are, the more they want to know, hence they ask more questions. The professor has to spend more time answering those questions, thus he makes his lectures longer and the students are not happy. Hence, his other "reward" (actually, penalty) for class not being confused, if the lecture is overtime is -10. If the lecture is on time, the penalty is 0.

1. Denote by O_t the binary variable that indicates whether the lecture is overtime or not, $O_t \in \{Yes, No\}$. Assume that whether the lecture at time t is overtime does not depend on the lecture being overtime at $t-1$, but it does (as described in the problem) depend on the state of the class confusion at time t. We are told that, if the class is confused, the probability of overtime is 0.1. If the class is not confused, that probability increases to 0.4.

   Draw the Markov decision network that describes this problem. Include action nodes, random variables, and reward nodes. Define all network parameters. Show at least two time slices of this network.

2. Convert the network you constructed in (1) to a basic MDP network we discussed in the class. You may need to do some node grouping in the process of conversion. Define all the parameters of this new network. (Note: if you need to group two utility nodes, assume the individual utilities are additive.) Show two time slices of this network.

3. Write the Bellman equations for this problem. Solve them using the value iteration algorithm. Show how the utility estimates change as the iterations progress. How many iterations are necessary for 1% relative accuracy of the utility values? Assume initial utilities of 0 for all states.

4. Let us assume we know the class is confused at time $t = 5$ and the lecture is overtime. What is the best course of action at this point?

5. Assume we know the class is confused at time $t = 10$, but we forgot our watches and do not know if it is overtime or not. What is the best decision in this case?

4 Problem: Learning

Consider Problem 2 in Assignment 6 (Asia network).

1. Use the 1000 samples you drew from the network to learn all of the network parameters using the maximum likelihood learning model. Show sufficient statistics you used to learn the network parameters.

2. Now assume you have no evidence about $Visit\_to\_Asia$ and $Positive\_X\_ray$. Starting with the parameters you computed in (1) run three iterations of the EM algorithm to reestimate the network parameters. For all three iterations show the following: $P(Visit\_to\_Asia, Positive\_X\_ray|Evidence)$, cumulative statistics for $Visit\_to\_Asia$ and $Positive\_X\_ray$ nodes, and current network parameters.

3. Comment on these results and how they differ from each other and the true network parameters.

5 Problem: Decision trees

Build decision trees for nodes $Has\_tuberculosis$, Has_lung_cancer, and $Has\_bronchitis$ from the 1000 samples in problem 4. Use the information gain as the attribute selection criterion. Compare the decision tree results with those obtained by the Asia Bayesian network.