**B551 Final Practice Problems: Fall 2012**

The final exam will be similar in breadth, length, and difficulty to these problems. All topics covered in class will be fair game for the final exam.

**I. Traveling on a roadmap**

You are given a roadmap of some country in the form of a connected non-directed graph in which nodes represent cities and edges represent roads between cities. (A connected graph is one in which every two nodes are connected by a path made of one or several consecutive edges.) Each edge (*i*,*j*) is labeled by the length *l*(*i*,*j*) of the road between cities *i* and *j*.

Two friends live in two different cities, *a* and *b*, of the map. They want to meet in a city of the map (any one). To do this, they move in successive turns. On every turn, the two friends start moving at the same time. Each friend moves to a neighboring city on the map; he/she cannot stay in the same city. The amount of time needed to move from city *i* to neighboring city *j* is equal to the length *l*(*i*,*j*) of the road between cities *i* and *j*. So, the two friends may not reach their respective new cities at the same time. The friend that arrives first to his/her new city must wait until the other arrives to his/her new city (each one calls the other on his/her cell phone when he/she arrives to a new city) before the next turn can begin. The two friends want to meet as quickly as possible. Note that the goal for the two friends is to meet in a city, not anywhere on a road.

1. Formulate this problem as a state-space search problem:

a) What is the state space? N^2 -1 states

b) What are the initial and the goal states?

c) What is the successor function? Cartesian product of list of adjacent cities

d) What is the step cost function? Max(L(i1,j1),L(i2,j2)) or Sum of the paths taken by both

2. Let D(*i*,*j*) be the straight-line distance between any two cities *i* and *j* in the map. Which, if any, of the following heuristic functions are admissible? Why?

For the given cost (Max), if the weights of each edge were 3 and 3 , b would say 4 but the real cost is 3.

a) D(*i*,*j*)

b) D(*i*,*j*)2

c) D(*i*,*j*)/2

3. Is the following statement true or false: “There are connected maps for which no solution exists”? If you answer ‘true’, give an example of such a map. If you answer is ‘false’, prove it.

Two nodes, connected. Both have to keep moving.

**II. Modified Tic-Tac-Toe**

Consider the game of 2×2 tic-tac-toe where each player has the additional option of passing, i.e., of marking no square on the 2×2 board. The two players, MAX and MIN, take turns, with MAX going first.

1. Draw the full game tree down to depth 2 (recall that the root of the tree is at depth 0). Do not show the nodes that are rotations or reflections of siblings already shown (your tree should have five leaves). [Draw your tree nicely since you will have to use it again for your answers to questions 2 and 3.]

2. Let the evaluation function of MAX be the number of MAX’s marks on the board minus the number of MIN’s marks. Give the values of the evaluation function for all leaves of the tree constructed in Question 1 and the values backed-up by the Minimax algorithm for all internal nodes. [Show these values on the tree drawn in Question 1.]

3. Circle all nodes [in the tree that you drew in Question 1] that would not be evaluated by the Alpha-Beta algorithm during a left-to-right depth-first exploration of your tree.

4. Suppose we wanted to solve the game to find the optimal move of MAX (i.e., by constructing a game tree with no depth limit). Explain why Alpha-Beta pruning with an appropriate node ordering can do it, while Minimax can’t.

**III. Automotive Diagnosis**

Consider the following simple network for car diagnosis:

Each variable is Boolean, and a value of True indicates that the aspect of the car is working properly.

1. How many independent probability values would be listed in the joint probability table for these six variables, if no independence assumptions were made? 2^6 -1
2. How many independent probability values are listed in the conditional probability tables of this BN? 12
3. Given what other variables can you say that Battery is independent of Moves? Spark Plugs,start Given what other variables can you say that Battery is NOT independent of Moves? Radio gas
4. Given what other variables can you say that Gas is independent of Radio? battery spark plugs Given what other variables can you say that Gas is NOT independent of Radio? Start move
5. List at least two algorithms that can perform inference on this Bayesian network.

Variable elimination, monte carlo ,enumeration

**IV. Support Vector Machines**

1. Construct by hand a 2-dimensional linear classifier that is consistent with the positive examples (0,3), (1,4), (2,5) (2,3) and the negative examples (3,2), (1,1),(2,2), (4,3). Give the equation defining this classifier. Y-X>1
2. For the above example, consider the linear classifier -2\*x + y + 2. What is the geometric margin of each of the data points? (If an example is misclassified, then its margin is negative)
3. Consider the following 2D dataset (filled circles indicate positive examples, empty circles indicate negative ones).

x1

x2

How might you transform the data into a higher-dimensional feature space in order to get a good linear classifier in that space? Hint: consider that the positive examples seem to be above a parabola.

**V. Reasoning with Uncertainty Over Time**

1. Suppose you wanted to build a second order Markov model of English sentences, i.e., a probabilistic model P(wk|wk-1,wk-2), by learning from a dataset D of sentences w1,…,wn. Considering there are approximately 170,000 words in the English language, describe two problems with maximum likelihood estimation of this conditional probability distribution. Describe a remedy for at least one of these problems.

2. A professor wants to know if students are getting enough sleep. Each day the professor observes whether the students sleep in class, and whether they have red eyes. The professor believes the following:

* + The prior probability of getting enough sleep, with no observations, is 0.7
  + The probability of getting enough sleep on night t is 0.8 given that the student got enough sleep the previous night, and 0.3 if not.
  + The probability of having red eyes is 0.2 if the student got enough sleep, and 0.7 if not.
  + The probability of sleeping in class is 0.1 if the student got enough sleep, and 0.3 if not.

Formulate this information as a hidden Markov model with an observation variable O=(RedEyes, Sleeping) that consists of two binary random variables. Give the probability tables for this model.

Then, given the HMM and observations O1 = not red eyes or sleeping, O2 = red eyes and not sleeping, and O2 = red eyes and sleeping, describe how to perform the computations for:

1. Filtering: Compute P(EnoughSleept|O1:t) for each of t=1,2,3.
2. Smoothing: Compute P(EnoughSleep2|O1:3)

(For your answers, just show the computations needed to compute these quantities in terms of the professor’s beliefs. You do not need to evaluate them numerically, and you may use the quantities you computed in part A in part B.)

How does the smoothed estimate P(EnoughSleep2|O1:3) differ from P(EnoughSleep2|O1:2)?

3. Match the *problems* on the left to the *algorithms* on the right.

|  |  |
| --- | --- |
| A. Maintain a probability distribution over time as new observations arrive. | Recursive filtering |
| B. Learn the parameters of a probabilistic model from data using a prior | Viterbi algorithm |
| C. Compute the utility values of an optimal policy. | Maximum a posteriori estimation |
| D. Maintain a continuous probability distribution as new observations arrive, assuming a linear Gaussian process. | Value iteration |
| E. Find the sequence of states that is most likely given a sequence of observation | Kalman filtering |