CS 188 Final

May 19, 2008

NAME:

1 True/False Questions (14 points)

Circle *True* or *False* next to each statement below. Correct answers are worth 1 point each, incorrect answers are worth 0 points, and skipped parts are worth 0.5 points each.

(a) (True/False) Given any probability distribution over N variables, every Bayes' net with N nodes can represent it.

(b) (True/False) All conditional independence properties of the probability distribution represented by a Bayes' net are determined by the graph structure.

(c) (True/False) Given a Bayes' net with N nodes that can represent any probability distribution over N variables, that same network can represent any marginal distribution over M variables where $M \leq N$.

(d) (True/False) In the worst case, the complexity of inference could be exponential in the size (number of variables) of a Bayes' net.

(e) (True/False) The smoothing algorithm for HMMs returns the single most likely sequence of hidden states given the observations.

(f) (True/False) In reinforcement learning, it is useful to sometimes choose an action believed to be suboptimal.

(g) (True/False) Q-learning algorithms require a model of the environment.

(h) (True/False) In a two-player zero-sum game, player A could do worse than the result implied by his minimax strategy if his opponent, player B, plays suboptimally.

(i) (True/False) For tree search problems, if an admissible heuristic is used, the A* search algorithm will never return a suboptimal goal node.

(j) (True/False) Given that $h_1(x)$ and $h_2(x)$ are both admissible heuristics for an A* search problem, $h(x) = h_1(x) + h_2(x)$ is also an admissible heuristic.

(k) (True/False) The result of training both Neural Networks and Support Vector Machines may be a local optimum different from the global optimum.

(l) (True/False) Maximizing information gain is a good heuristic for constructing decision trees.

(m) (True/False) Brightness, color, texture, binocular disparity, and optical flow are all useful cues for detecting boundaries in natural images.

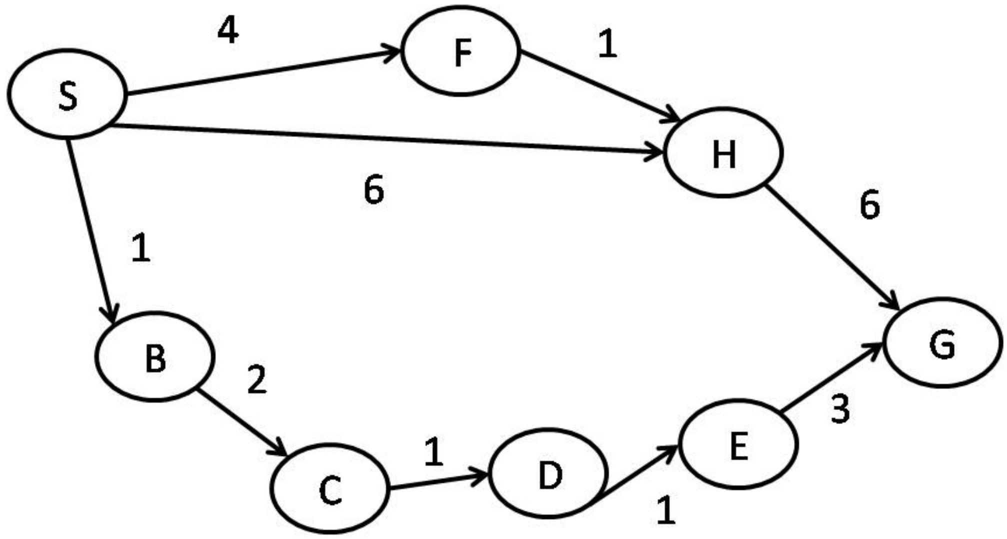
(n) (True/False) A sentence can have multiple syntactically correct parses.

For official use only:

Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Total
/14	/8	/7	/15	/20	/5	/11	/6	/86

2 A* Search (8 points)

Consider the following search graph starting from node S with a goal state G.



(a) (3 points) What is the path that is returned using A* search with the null heuristic, $h(x) = 0$? Also list the remaining states in the priority queue when the goal state is found along with their enqueued values.

(b) (3 points) Create a table and record the best admissible heuristic for each node in the graph above.

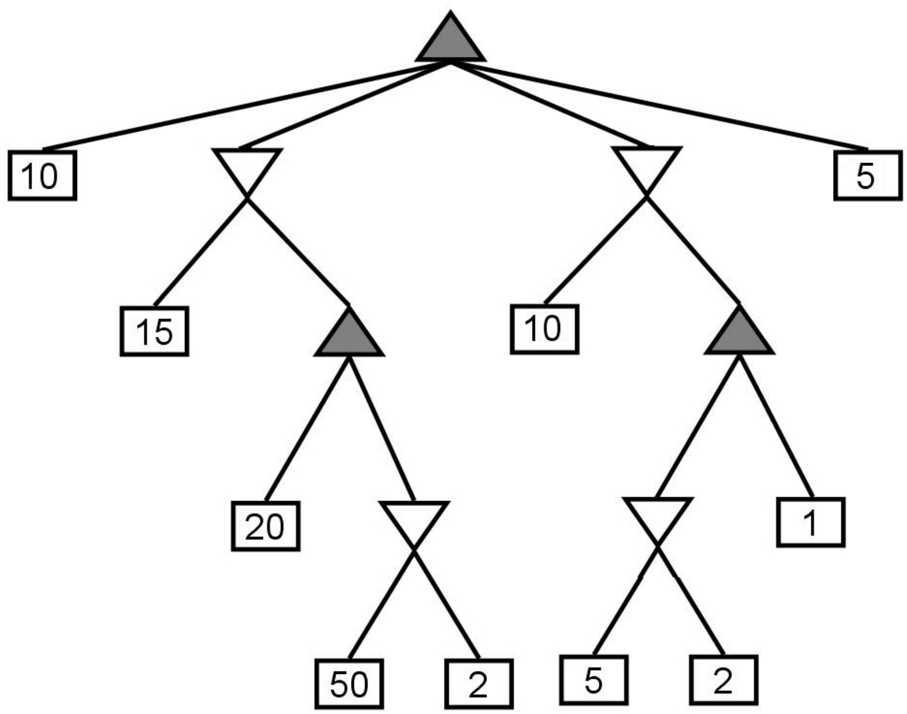
(c) (2 points) Why can't the solution to part (b) be be done practically for most problems?

3 Games and Minimax Search (7 points)

Company A and company B are competitors for the same market. Currently, A makes a profit of \$5 million and B makes a profit of \$3 million. Neither company advertises heavily. If A were to advertise, they would steal part of the market from B, increases their own profit by \$1 million and decreasing B's profit by \$1.5 million. If B were to advertise, they would increase their profit by \$2 million and decrease A's profit by \$1 million. If both advertise, both profits decrease by \$0.5 million.

(a) (2 points) What is the Nash equilibrium for this situation?

Consider the following minimax tree where maximizing nodes are upward triangles that are shaded in and minimizing nodes are downward triangles.

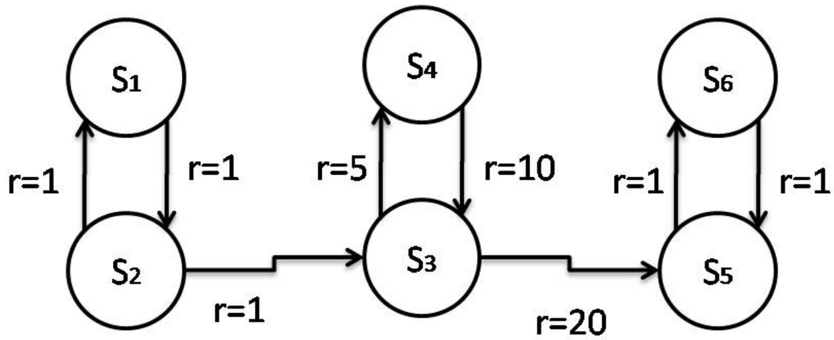


(b) (2 points) What is the minimax value for the root?

(c) (3 points) Draw an X through any nodes which will not be visited by alpha-beta pruning, assuming we explore left subtrees before right subtrees.

4 MDPs and Reinforcement Learning (15 points)

Consider the following Markov Decision Process:



We have states S_1, S_2, S_3, S_4, S_5, and S_6. We have actions *Up*, *Down*, *Left*, and *Right*, and each action deterministically leads to a successor state (with probability 1). Actions not shown are not allowed. For example, in S_1, the only available action is to go to S_2 (*Down*). We follow the Sutton-Barto convention and associate rewards with transitions. The reward for any action is specified above. For taking *Down* from S_4, the reward is 10. Assume a discount factor $\gamma = 0.5$. Recall that $\sum_{i=1}^{\infty} 0.5^i = 1$

(a) (1 point) What is the optimal policy for this MDP?

$$\{S_1: \quad , S_2: \quad , S_3: \quad , S_4: \quad , S_5: \quad , S_6: \quad \}$$

Value Iteration: When performing value iteration, what is the value of state S_3 after

(b) (1 point) 1 iteration:

(c) (1 point) 2 iterations:

(d) (2 points) ∞ iterations:

Policy Iteration: Suppose you run policy iteration. During each iteration, you compute the exact values in the current policy, then update the policy using one-step lookahead given those values. Suppose, too, that when choosing actions when updating the policy, ties are broken by choosing *Up*, *Right*, *Down*, *Left* in that order.

Suppose we start with the following policy:

$$\{S_1: \ Down, \ S_2: \ Right, \ S_3: \ Up, \ S_4: \ Down, \ S_5: \ Up, \ S_6: \ Down\}$$

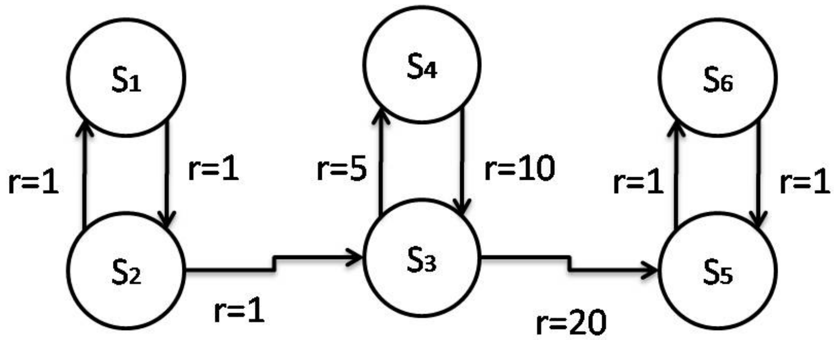
What will be the policy after... (*Note: Fill in with U, D, R, and L for Up, Down, Right, and Left*)

(e) (2 points) 1 iteration: $\{S_1: \quad , S_2: \quad , S_3: \quad , S_4: \quad , S_5: \quad , S_6: \quad \}$

(f) (2 points) 2 iterations: $\{S_1: \quad , S_2: \quad , S_3: \quad , S_4: \quad , S_5: \quad , S_6: \quad \}$

(g) (2 points) 3 iterations: $\{S_1: \quad , S_2: \quad , S_3: \quad , S_4: \quad , S_5: \quad , S_6: \quad \}$

MDP repeated here for convenience:



Q-learning: Consider executing Q-learning on this MDP. Assume the learning rate $\alpha = 0.5$, and that Q-learning uses a greedy policy, meaning that it always chooses the action with maximum Q-value. Suppose the algorithm breaks ties by choosing *Up*, *Right*, *Down*, *Left* in that order.

(h) (2 points) What are the first 10 $(state, action)$ pairs visited if our agent learns using Q-learning and starts in S_3?

(i) (2 points) What is the Q-value of (S_5, Up) after these first 10 actions?

5 Hidden Markov Models (20 points)

Recall that on the midterm, we helped Billy model the state of his health using the following Bayes' net. He could exhibit symptoms coughing (C), sneezing (S), and temperature (T), with underlying cause (X) as sick ($X = sick$), allergic ($X = allergic$), or well ($X = well$).

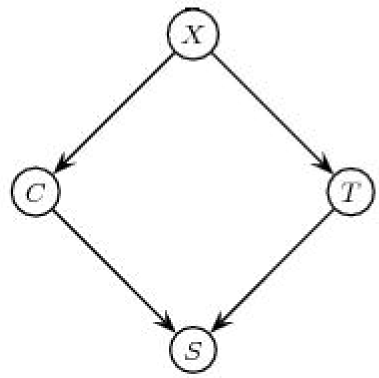


Figure 1: Bayes' net

Billy has now been sick for several days and asks you to formalize the previous Bayes' net as the Hidden Markov Model (HMM) shown in Figure 2.

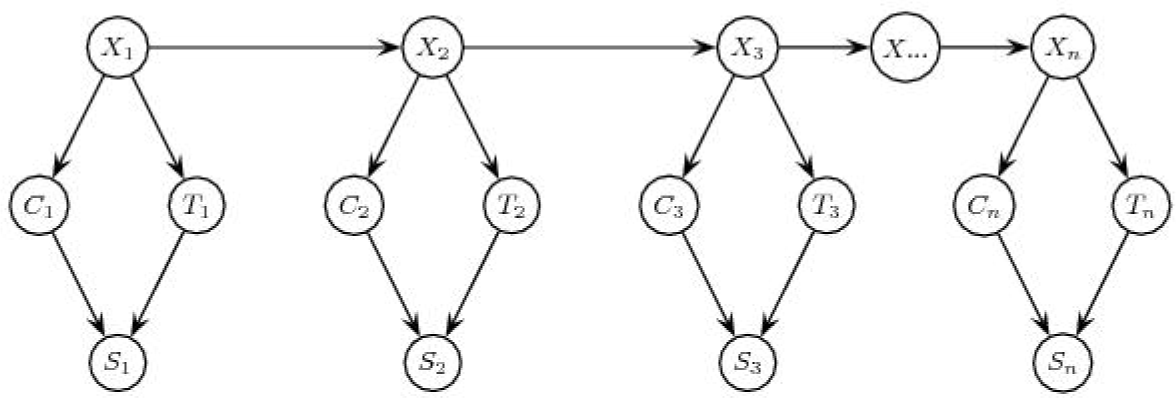


Figure 2: HMM

Billy has not been keeping track of his temperature, but does remember the history of his coughing and sneezing symptoms. Hence, although we modeled his temperature in creating the diagnostic criteria, the only observations we have are whether Billy is coughing or sneezing on each day.

Recall that we parameterize the distribution $P(X_k, S_1, ..., S_n, C_1, ..., C_n)$ with $\alpha(X_k), \beta(X_k)$, so

$$P(X_k, S_1, ..., S_n, C_1, ..., C_n) = \alpha(X_k)\beta(X_k)$$

$$\alpha(X_k) = P(X_k, S_1, ..., S_k, C_1, ..., C_k)$$
$$\beta(X_k) = P(S_{k+1}, ..., S_n, C_{k+1}, ..., C_n | X_k)$$

Also recall the following equations from the midterm, which may be useful in answering the questions below:

$$P(C,S) = \sum_x \sum_t P(X=x, C, T=t, S)$$

$$P(X|C,S) = \frac{\sum_t P(X, C, T=t, S)}{\sum_x \sum_t P(X=x, C, T=t, S)}$$

$$P(C,S|X) = \frac{\sum_t P(X, C, T=t, S)}{\sum_c \sum_s \sum_t P(X, C=c, T=t, S=s)}$$

(a) (2 points) Derive the probability distribution for $P(X_k, S_1, ..., S_k, C_1, ..., C_k)$ in terms of $P(X_i|X_{i-1})$, and $P(X,C,T,S)$ using marginalization and conditional independence.

(b) (2 points) Derive the probability distribution for $P(S_{k+1}, ..., S_n, C_{k+1}, ..., C_n|X_k)$ in terms of $P(X_i|X_{i-1})$, and $P(X,C,T,S)$ using marginalization and conditional independence.

(c) (2 points) Derive the $\alpha(X_k)$ update in terms of $\alpha(X_{k-1})$, $P(X_k|X_{k-1})$, and $P(X,C,T,S)$. Remember that you only have observations for C and S.

(d) (2 points) Derive the $\beta(X_k)$ update in terms of $\beta(X_{k+1})$, $P(X_{k+1}|X_k)$, and $P(X, C, T, S)$. Remember that you only have observations for C and S.

(e) (2 points) Derive the equation for filtering using $\alpha(X_k)$, to find $P(X_k|S_1, ..., S_k, C_1, ..., C_k)$

(f) (2 points) Derive the equation for prediction using $\alpha(X_k)$, to find $P(X_{k+t}|S_1, ..., S_k, C_1, ..., C_k)$. Hint: it may help to draw this Bayes' net.

(g) (2 points) Derive the equation for smoothing using $\alpha(X_k)$ and $\beta(X_k)$, to find $P(X_k|S_1, ..., S_n, C_1, ..., C_n)$.

(h) (2 points) Derive the Viterbi update ($\alpha^*(X_k)$) for this HMM.

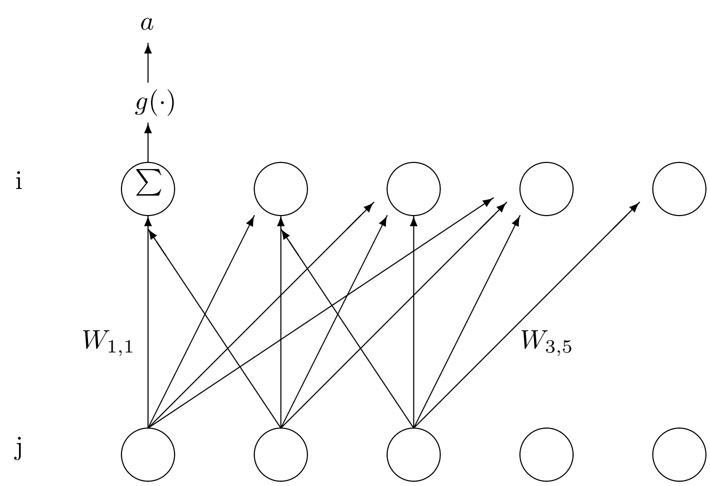
(i) (2 points) Briefly describe how you would extract the most-likely sequence from $\alpha^*(X_k)$. Would you need any data structures? If so, specify which and what you would keep track of.

Suppose that we knew the reward of taking Robitassium DM when Billy has a allergy ($X = allergy$) or when he is well ($X = well$) is -1. Additionally, if Billy takes Benedryll when he is sick or when he is well, he receives a reward of -1. If Billy uses the medicine correctly, taking Robitassium DM for a sickness or Benedryll for his allergies, or nothing when he is well, he will receive a reward of $+1$.

(j) (2 points) Assuming that Billy observes his symptoms (S and C) when he wakes up each day before (possibly) taking medicine, explain how he should decide which (if any) medicine to use so as to maximize his utility. Assume that Billy has a record of his symptoms from previous days.

6 Neural Networks (5 points)

Consider the Neural Network shown below. We've drawn the output layer as well as the layer immediately before it. Nodes in the output layer are indexed by i and nodes in the previous layer are indexed by j. The outputs are $a_i = g(in_i)$ where g is the sigmoid function and $in_i = \sum_j W_{ji} a_j$ is the input to node i.



Neural Network. For clarity, only a subset of connections are shown.

(5 points) Given desired output vector y_i, we want to adjust weights so as to minimize the error $E = \frac{1}{2}\sum_i (y_i - a_i)^2$. Derive the expression for $\frac{\partial E}{\partial W_{ji}}$, the partial derivative of the error E with respect to weight W_{ji}.

7 Classification (11 points)

We are designing an intrusion detection system using three sensors, a sound level meter (S), an infrared camera (C), and a floor pressure sensor (F). Unfortunately, each sensor is relatively unreliable and hence we want to base the decision on whether or not an intruder (I) is present on the output of all three.

The sound meter can indicate one of two states (S = quiet or loud).
The infrared camera can indicate one of three conditions (C = cold, warm, or hot).
The floor pressure sensor can indicate one of three conditions (F = light, medium, or heavy).

We want to decide whether or not an intruder is present (I = true or false).

(a) (2 points) Suppose we choose to use a naive Bayes' classifier. Draw the corresponding Bayes' net.

To collect training data for our classifier, we open the room in which our sensors are set up and observe the sensor values as people enter and leave the room. We obtain the following observations:

S = quiet, C = hot, F = heavy, I = true
S = loud, C = warm, F = medium, I = true
S = quiet, C = hot, F = light, I = false
S = quiet, C = warm, F = light, I = false
S = loud, C = warm, F = medium, I = true
S = quiet, C = warm, F = medium, I = false
S = quiet, C = warm, F = medium, I = true
S = loud, C = hot, F = medium, I = true

(b) (3 points) Using this training data, compute estimates for the parameters of the naive Bayes' classifier, without smoothing.

(c) (2 points) Are all entries in the conditional probability tables in part (b) reasonable? If not, describe how you'd use smoothing to correct the problem.

(d) (2 points) Assume that we do not make any corrections and use the parameters as they are in part (b). Our sensors read $S =$ quiet, $C =$ warm, and $F =$ medium. What is the probability an intruder is present?

(e) (2 points) Without making any changes to the sensors, what additional information could we use and how would we model it?

8 Utility (6 points)

A gambler has the opportunity to make bets on the outcome of a sequence of coin flips. If the coin comes up heads, she wins as many dollars as she has staked on that flip; if it is tails she loses her stake. The game ends when the gambler wins by reaching her goal of attaining \$4, or loses by running out of money. On each flip, the gambler must decide what portion of her capital to stake, in integer number of dollars. The state is the gambler's capital $s \in \{1, 2, 3\}$. The reward is zero on all transitions except ones in which the gambler reaches her goal, when it is $+1$.

Let the probability of the coin coming up heads be p. Solve for the utilities in two cases:

(a) (3 points) For $p = 0.7$, the optimal policy is to always bet \$1. Calculate the utilities for each of the three non-terminal states.

(b) (3 points) For $p = 0.4$, the optimal policy is non-obvious. Implement value iteration to calculate the utilities. Start with an initial guess of $U_1 = 0.25$, $U_2 = 0.5$, $U_3 = 0.75$