In the name of God, the Merciful, the Compassionate

Artificial Intelligence Final Exam

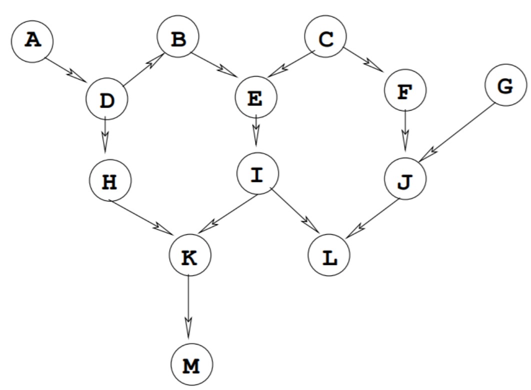
Fall 2020 – Group 2

Time: 5 hours

Note: There are 20 bonus points in the exam.

1) **[30 pts.]** Answer each of these questions briefly. Please justify your answer.
    a. Can all conditional independencies be inferred only from the graph of a Bayes' net?
    b. The columns in the probability table that is used in Gibbs sampling of a variable A given other variables, in a Bayes' net with four nodes are A, B, and C. In addition, by just looking at the Bayes' net graph, one knows that B and C are independent. In addition, one can similarly infer independence of D from A given B. Determine possible graph(s) of the Bayes' net and justify your solution.
    c. Can a hidden Markov model with deterministic emission probabilities be modeled by a Markov model on the observations? Explain. (Deterministic emission probabilities mean that conditional emission probabilities are either 0 or 1).
    d. The classification accuracy of a decision tree on the validation set is much worse than its accuracy on the training set. How would you solve this issue? Explain.
    e. You observed that in the policy evaluation, the value function of states, V(s), diverges. What is the possible cause of this problem? Explain.
    f. Assume that an active RL agent has converged to a suboptimal policy? What has gone wrong? Explain.

2) **[30 pts.]** Consider a Bayes' net with the following structure:



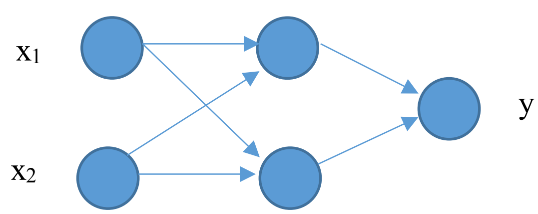
a. **[7 pts.]** What variables are independent of F, if only the variable I is observed. Explain.
b. **[8 pts.]** For A and J to be independent, what variables need to be observed? Explain.
c. **[15 pts.]** Suppose that B, E, and I are the only hidden variables in an exact inference problem. In addition, D is the only query variable in this inference. All variables are

binary. What order of variable elimination would you prefer for better computational efficiency? Justify your answer.

3) **[15 pts.]** Consider a Markov model with 5 states {0, …, 4}. Each state i stays in itself with probability of 0.2 and goes to the state (i-1 mod 5) with probability of 0.3, and to the state (i+1 mod 5) with probability of 0.5. Suppose that the model starts in state 0 at time t=1.
   a. **[6 pts.]** What is the probability that the model stays at the state 2 at time t=3.
   b. **[9 pts.]** If the sequence of states up-to time 3 is (0, 1, 1), what is the probability that we will be at state 2 at time 5?

4) **[15 pts.]** Consider the following multi-layer perceptron, with the dataset consisting of two data points (-1, 2) and (-1, 5), with two features x_1 and x_2. The labels, y, of the two data points are 1 and -1, respectively. Assume that all the weights are initialized to zero, the activation function is sigmoid (see below), and no bias is used. What are the updated weights at the end of epoch 1.
   Note: The sigmoid function is:

$$S(x) = \frac{1}{1+e^{-x}}$$



5) **[30 pts.]** Consider an MDP with three states 0, 1, and 2; and two actions a and b. The action "a" makes the MDP go to a random state (including the current state, and with equal probability of 1/3) when it is taken in any state. The action "b", however, deterministically makes state i go to state (i+1 mod 3). The reward of going into states 0, 1, and 2 are 0, -1, and 1, respectively. Let the discount factor be 0.9. We are aiming at maximizing the expected sum of discounted rewards.
   a. **[8 pts.]** Consider the policy that runs "b" in all the states. Evaluate the value of all the states using this policy.
   b. **[7 pts.]** Use policy improvement to make this policy better.
   c. **[15 pts.]** Now suppose that neither the transition probabilities nor the rewards are known to the agent. Let the observation sequence of the vector (current state, action, next state, reward) be (0, a, 2, 1), (2, a, 1, -1), (1, b, 2, 1), (2, b, 0, 0). What would be the estimated Q table, using the Q-learning algorithm? Let the learning rate be 0.1.