Probabilistic Artificial Intelligence
Final Exam
Feb 10, 2017

Time limit:	120 minutes
Number of pages:	19
Total points:	100

You can use the back of the pages if you run out of space. Collaboration on the exam is strictly forbidden. Please show *all* of your work and always *justify* your answers.

Please write your answers with a *pen*.

(1 point) Please fill in your student ID and full name (LASTNAME, FIRSTNAME) in capital letters.

Please leave the table below empty.

Problem	**Maximum points**	**Obtained**
1.	20	
2.	21	
3.	17	
4.	9	
5.	12	
6.	20	
Total	100	

1. Bayesian Networks and d-separation (20 points)

(6 points) (i) Consider the Bayesian Network in Figure 1. Are the following d-separation statements true? If a statement is true, give a blocking node. If it is not true, give an active trail.

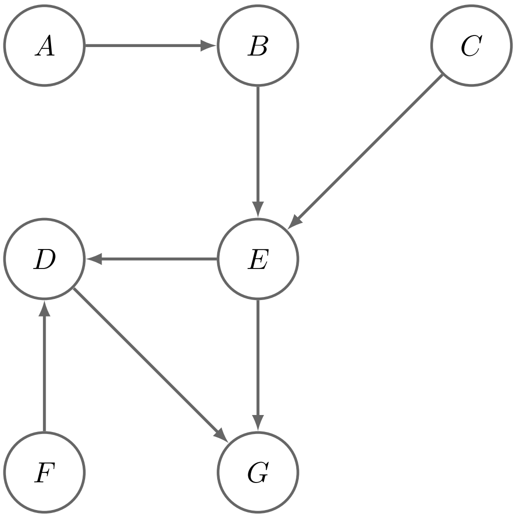


Figure 1

d-sep(A ; C | E)? ☐ True ☐ False

d-sep(A ; C | G)? ☐ True ☐ False

d-sep(A ; D | B)? ☐ True ☐ False

d-sep(A ; F | G)? ☐ True ☐ False

d-sep(F ; E)? ☐ True ☐ False

d-sep(F ; E | G)? ☐ True ☐ False

(4 points) (ii) Let X, Y, and Z be random variables in a Bayesian Network. For each of the statements below, decide whether they are true or false. (Each correct answer gives +0.5 points, each incorrect answer gives -0.5 points. You cannot get less than 0 points.)

- d-separation of X and Y given Z implies conditional independence of X and Y given Z.

  ☐ True ☐ False

- d-separation of X and Y given Z can only be shown with knowledge of the CPTs.

  ☐ True ☐ False

- A fully-connected Bayesian Network means that all variables are independent.

  ☐ True ☐ False

- The Bayesian Network structure in Figure 1 is a DAG.

  ☐ True ☐ False

- The Bayesian Network structure in Figure 1 is a polytree.

  ☐ True ☐ False

- Factor graphs are a particular type of (bipartite) Bayesian Network.

  ☐ True ☐ False

- Belief propagation on a factor graph will always converge but is not guaranteed to converge to the correct marginals.

  ☐ True ☐ False

- Belief propagation will converge to the correct marginals for tree factor graphs (e.g., for polytree Bayesian networks).

  ☐ True ☐ False

(2 points) (iii) Write down the factorized joint probability distribution over the random variables $A, \ldots, G$ according to the Bayesian Network in Figure 1.

(4 points) (iv) Draw a factor graph with exactly two factor nodes representing the same distribution as the Bayesian Network shown in Figure 1 (also define the two factors).

(4 points) (v) Suppose you wish to do variable elimination on the Bayesian Network in Figure 1. Consider the variable ordering D, A, B, C, E, F, G, and the first iteration of the algorithm (eliminating variable D). Write down the equation for the first factor g_1 (which may depend on one or more variables), and give the joint distribution $P(A, B, C, E, F, G)$ in terms of this new factor.

2. A better new year with Bayesian Networks (21 points)

As a New Year's resolution, you decide to work out more frequently. In particular, in the morning you go to the gym with a probability of 0.5. Additionally, you decide to pick up running. Being a good-weather runner, you only go for a run when it is sunny outside. This being Zurich, only a quarter of the days are sunny. On sunny days, you are often too tired if you already went to the gym in the morning and end up having a 0.2 probability of going for a run. However, if you did not go to the gym, you go for a run with probability 0.6.

Whether you end up with muscle ache at the end of the day depends on the amount of exercise you have done. If you did not exercise, you will not get muscle ache. If you go both to the gym and for a run, you will always get muscle ache. Otherwise, you have a 0.2 probability of muscle ache independent of which of the two exercises you completed.

(8 points) (i) Draw the Bayesian Network corresponding to the text above and write down the conditional probability distributions in terms of four random variables. Explain the meaning of each random variable that you define.

(2 points) (ii) Write down the joint probability distribution that is induced by the Bayesian network.

(3 points) (iii) What is the probability of going for a run? Show the steps of your computation.

(4 points) (iv) What is the probability of having gone to the gym, given that you went for a run? Show the steps of your computation.

(4 points) (v) What is the probability of a muscle ache, given that you went for a run? Show the steps of your computation.

3. Hidden Markov Models (17 points)

Consider two friends, Alice and Bob, who live far apart from each other and who talk together daily over the telephone about what they did that day. Bob is a student at a university in California and is only interested in two activities: *going to the university* and *going to the beach*. The choice of what to do is determined exclusively by the weather on a given day, which is either *sunny* or *rainy*. Alice has no definite information about the weather where Bob lives, but makes the following assumptions:

- The probability of sun on the first day (with no observations) is 0.95.
- The probability of *sun* at day t is 0.9 given that it was *sunny* the day before, and 0.6 if not.
- The probability of *going to the beach* is 0.8 if the weather is *sunny*, and 0.1 if not.

(5 points) (i) Formulate this information as a Hidden Markov Model (HMM) that Alice could use to predict the weather from a sequence of observations of Bob's activities. Draw the Bayesian Network corresponding to the HMM and give the complete probability tables for the model, i.e. the prior distribution of the initial state, the transition probabilities and the emission probabilities.

(5 points) (ii) Assume that Alice wants to estimate how many days Bob *goes to the university* within the next 30 days. To do this, Alice wants to sample from the HMM. Describe (using pseudo-code) how Alice can obtain a *single* sample from the HMM using *forward sampling*. You may assume that functions for drawing samples according to the prior probabilities, the transition probabilities and the emission probabilities are available.

(4 points) (iii) Describe how Alice can estimate the number of days Bob *goes to the university* within the next 30 days using multiple samples from the HMM.

(3 points) (iv) Let H_t denote the variable representing the hidden state at time t, and let O_t denote the variable representing the observation at time t. Mark all answers that are correct. (Each correct answer gives +0.5 points, each incorrect answer gives -0.5 points. You cannot get less than 0 points.)

- The observation O_t is conditionally independent of all previous observations $O_1, \ldots, O_{t-1}$ and previous hidden states $H_1, \ldots, H_{t-1}$ given the hidden state H_t.

  ☐ True ☐ False

- The hidden state at time $t+1$ is independent of all previous observations $O_1, \ldots, O_{t-1}$ and previous hidden states $H_1, \ldots, H_{t-1}$.

  ☐ True ☐ False

- The hidden state H_{t+5} is conditionally independent of the observations $O_1, \ldots, O_{t-1}$ and previous hidden states $H_1, \ldots, H_{t-1}$ given the hidden state H_t.

  ☐ True ☐ False

- Any k*th*-order HMM can be represented as a 1*st*-order HMM.

  ☐ True ☐ False

- An HMM is a polytree Bayesian network.

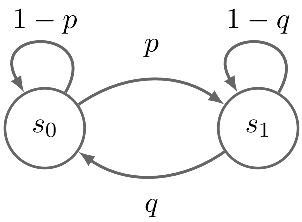
  ☐ True ☐ False

- Given a sequence of observations $O_1, \ldots, O_T$, *filtering* is the task of computing the distribution over states of the hidden variable H_T.

  ☐ True ☐ False

4. Sampling (9 points)

Consider the following Markov Chain with two states and transitions as depicted below, where $0 < p, q \leq 1$.



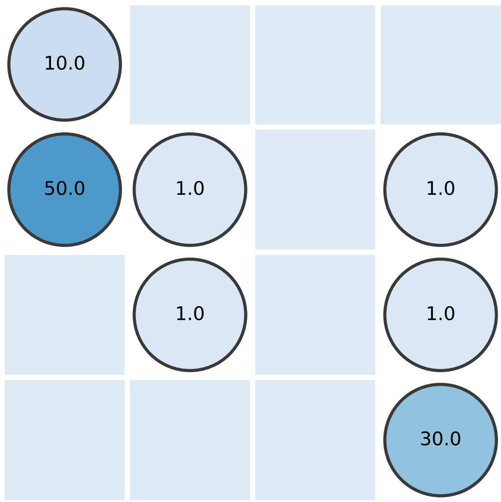
(3 points) (i) Write down the transition matrix T of the above Markov Chain. *[Reminder: The transition matrix is defined by $T(x, y) = \mathbb{P}[S_t = y \mid S_{t-1} = x]$, for all states x, y.]*

(2 points) (ii) If the chain has stationary distribution π with $\pi(s_0) = r$ and $\pi(s_1) = 1 - r$, write down the condition imposed by detailed balance in terms of p, q and r.

(4 points) (iii) If our desired stationary distribution is $\pi(s_0) = 1/3$ and $\pi(s_1) = 2/3$, what values would you choose for p and q in order for the chain to converge to π?

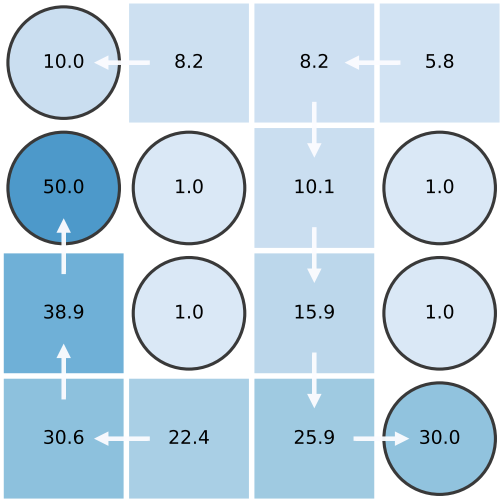
5. Q-learning **(12 points)**

Consider a maze consisting of a two-dimensional grid that contains non-terminal (square) and terminal (circle) states. A robot moving through the maze can take any of four actions (up, down, left, or right) as long as it stays within the grid. However, the actions do not always result in the desired movement. In particular, with probability p_{noise} the robot makes a move chosen uniformly at random among the allowed ones, irrespective of the chosen action. Finally, the robot only receives reward when moving to a terminal state; the reward amounts are shown in the figure below.

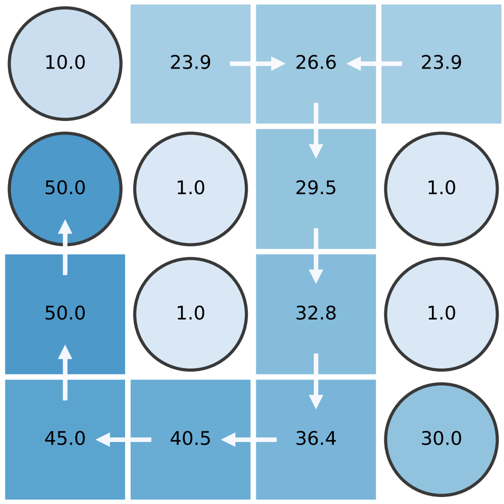


In the following figures, we plot estimates obtained by running $100{,}000$ episodes of the Q-learning algorithm with random starting states and a random exploration policy. In particular, we show the estimated value function $V(x)$ of each non-terminal state, as well as the associated greedy policy, both computed using the obtained Q-function estimates.

(6 points) (i) For this question, we only vary the amount of movement noise. Concretely, we run Q-learning twice, one time using $p_{\text{noise}} = 0$, and the other using $p_{\text{noise}} = 0.5$. In both cases we use the same discount factor γ. Match the resulting figures below to the two noise regimes. Briefly explain your reasoning.

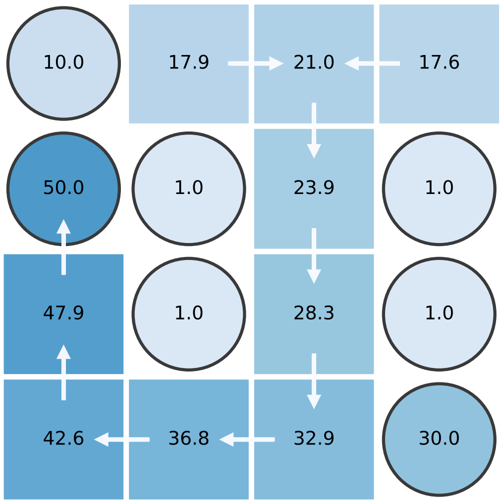


(a)

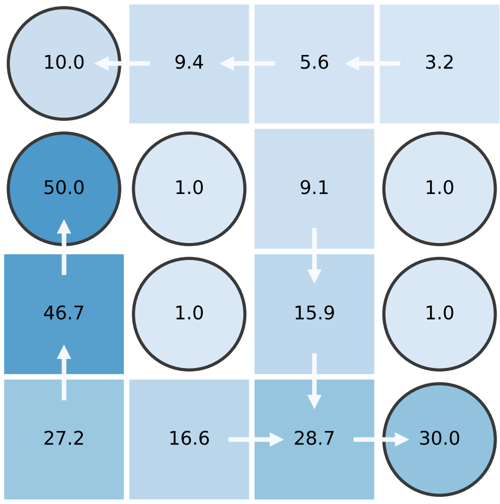


(b)

(6 points) (ii) For this question, we only vary the discount parameter γ. Again, we run Q-learning twice, one time using $\gamma = 0.6$, and the other using $\gamma = 0.9$. In both cases we use the same movement noise. Match the resulting figures below to the two choices of γ. Briefly explain your reasoning.



(a)



(b)

6. Managing a shop with MDPs (20 points)

Imagine that you are the manager of a shop. Your shop can have up to two items in stock. Every day you carry out one of the following actions:

- *order* one item from the production plant,
- *do not order* anything from the production plant.

Every day your clients either buy exactly one item from your shop or do not buy anything. This depends on the current stock according to the following distribution:

$$P(buy \mid stock = 0) = 0.0,$$

$$P(buy \mid stock = 1) = 0.8,$$

$$P(buy \mid stock = 2) = 0.6.$$

The number of items in stock on the next day depends on whether you have placed an order or not, and on whether your clients have bought an item or not. For example, if you have 1 item in stock and you have placed an order, next day there will be 1 item in stock with probability 0.8 and 2 items with probability 0.2. Placing an order when there are 2 items in stock always results in having 2 items in stock on the next day.

For simplicity, we assume that items can be ordered for free and that there is no profit gained by selling items. We only consider storage costs, which depend on the amount of items in stock on any given day:

$$\text{cost}(stock = 0) = 35\,\text{CHF},$$

$$\text{cost}(stock = 1) = 4\,\text{CHF},$$

$$\text{cost}(stock = 2) = 13\,\text{CHF}.$$

(6 points) (i) Draw the MDP corresponding to this sequential decision problem. [*Hint: Costs can be interpreted as negative rewards.*]

(8 points) (ii) In order to manage your center efficiently, you come up with the following policy:

$$\pi(0) = \text{order},$$
$$\pi(1) = \text{order},$$
$$\pi(2) = \text{do not order}.$$

Recall that the value function induced by a policy π is given by

$$V^\pi(x) = r(x) + \gamma \sum_{x'} P(x'|x, \pi(x)) V^\pi(x').$$

Compute the value function induced by the policy given above for $\gamma = 0.5$.

(6 points) (iii) Bellman's theorem states that a policy is optimal if and only if it is greedy with respect to its induced value function. Recall that the greedy policy π_V with respect to a given value function V is given by

$$\pi_V(x) = \arg\max_{a \in \mathcal{A}} \Big[r(x) + \gamma \sum_{x'} P(x'|x,a) V(x') \Big],$$

where $\mathcal{A}$ is the set of possible actions. Compute the greedy policy with respect to the value function computed in (ii).

Finally, use Bellman's theorem and the greedy policy you have computed to assess whether the policy given in (ii) is optimal or not.