INFO8006 Introduction to Artificial Intelligence

Exam of August 2023

Instructions

- The exam lasts for 4 hours.
- You are allowed to use a calculator during the exam, but documents of any kind are forbidden.
- The last two pages can be used for scratch work or for extra space. If you want work done there or on backs of pages to be graded, mention where to look **in big letters with a box around them**, on the page with the question.
- Write your last name, first name, and ULiège ID on the first page. Write only your ULiège ID on all the other pages.
- Before handing in your exam, **sort all the pages according to the page numbers** (even if you used additional pages to answer a question).

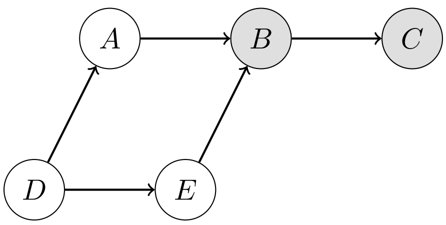
Good luck!

Last name, First name (in uppercase):

ULiège ID (s20XXXX):

Question 1 [4 points] Multiple choice questions. Choose one of the four choices by filling in its circle. Correct answers are graded $+\frac{4}{10}$, wrong answers are graded $-\frac{2}{15}$ and the absence of answers is graded 0. The total of your grade for Question 1 is bounded below at 0/4

1. Upon an observation e, the internal state of a model-based agent can be updated with ...
   - ○ A^*
   - ○ a Bayes filter
   - ○ Policy Evaluation
   - ○ Q-learning
2. Consider the game of Pacman in the absence of ghosts and with a single food dot in the maze. In state s, Pacman can either move up, down, left or right, *but also diagonally.* Pacman is located at position (i_s, j_s) while the food dot is located at (x_s, y_s). At timestep t, the score of the game is $500 - t$. The game ends when the food is eaten. Which of the following heuristics is the best to use?
   - ○ $h_1(s) = |i_s - x_s| + |j_s - y_s|$
   - ○ $h_2(s) = \max(|i_s - x_s|, |j_s - y_s|)$
   - ○ $h_3(s) = \min(|i_s - x_s|, |j_s - y_s|)$
   - ○ $h_4(s) = \sqrt{(i_s - x_s)^2 + (j_s - y_s)^2}$
3. Which of the following is false? In Minimax, pruning ...
   - ○ reduces the computation time of the algorithm.
   - ○ preserves the completeness of the algorithm.
   - ○ may change the computed minimax values.
   - ○ can be done using the α-β algorithm.
4. The Bayes rule states that ...
   - ○ $\mathbf{P}(X|Y) = \frac{\mathbf{P}(Y|X)\mathbf{P}(Y)}{\mathbf{P}(X)}$
   - ○ $\mathbf{P}(X|Y) = \frac{\mathbf{P}(Y|X)\mathbf{P}(X)}{\mathbf{P}(Y)}$
   - ○ $\mathbf{P}(X|Y) = \frac{\mathbf{P}(Y|X)}{\mathbf{P}(X)}$
   - ○ $\mathbf{P}(X|Y) = \frac{\mathbf{P}(X)}{\mathbf{P}(Y)}$
5. Consider the Bayesian network shown below. We want to infer $\mathbf{P}(D|b, c)$ where D is the query variable, B and C are evidence variables, and A and E are hidden variables. Which of the following statements is true?



   - ○ $\mathbf{P}(D|b, c) \propto \sum_a \sum_e P(c)P(b)\mathbf{P}(e|D)\mathbf{P}(a|D)\mathbf{P}(D)$
   - ○ $\mathbf{P}(D|b, c) \propto \sum_a \sum_e P(b|a, e)\mathbf{P}(e|D)\mathbf{P}(a|D)\mathbf{P}(D)$
   - ○ $\mathbf{P}(D|b, c) \propto \sum_a \sum_e P(c)P(b|c)P(a|b)P(e|b)\mathbf{P}(D|a, e)$
   - ○ $\mathbf{P}(D|b, c) \propto \sum_a \sum_e P(c|b)P(b|a)\mathbf{P}(a|D)\mathbf{P}(e|D)\mathbf{P}(D)$

6. The Kalman filter requires the specification of ...
   - ◯ a prior $\mathbf{P}(\mathbf{X}_0)$ and a transition model $\mathbf{P}(\mathbf{X}_{t+1}|\mathbf{x}_t)$.
   - ◯ a transition model $\mathbf{P}(\mathbf{X}_{t+1}|\mathbf{x}_t)$ and an observation model $\mathbf{P}(\mathbf{E}_t|\mathbf{x}_t)$.
   - ◯ a prior $\mathbf{P}(\mathbf{X}_0)$, a transition model $P(\mathbf{X}_{t+1}|\mathbf{x}_t)$, and an observation model $P(\mathbf{E}_t|\mathbf{x}_t)$.
   - ◯ a prior $\mathbf{P}(\mathbf{X}_0)$, a prior $\mathbf{P}(\mathbf{E}_0)$, a transition model $P(\mathbf{X}_{t+1}|\mathbf{x}_t)$, and an observation model $P(\mathbf{E}_t|\mathbf{x}_t)$.
7. Logistic regression models the conditional $P(Y = 1|\mathbf{x})$ as ...
   - ◯ $P(Y = 1|\mathbf{x}) = \sigma(\mathbf{w}^T\mathbf{x} + b)$, where σ is the sigmoid function.
   - ◯ $P(Y = 1|\mathbf{x}) = \sigma(\mathbf{w}^T\mathbf{x} + b)$, where σ is the ReLU activation function.
   - ◯ $P(Y = 1|\mathbf{x}) = \sigma(\mathbf{w}^T\mathbf{x} + b)$, where σ is the standard deviation function.
   - ◯ $P(Y = 1|\mathbf{x}) = \sigma(\mathbf{w}^T\mathbf{x} + b)$, where b is random Gaussian noise.
8. François is trying to perform gradient ascent on a function $f(x)$ using the update

   $$x_{t+1} := x_t - \frac{\partial f}{\partial x}(x_t).$$

   Is this update rule guaranteed to converge to a local maximum of f?
   - ◯ Yes, since he's updating using the gradient of f.
   - ◯ Yes, but not for the reason above.
   - ◯ No, since he is updating updating x in the wrong direction.
   - ◯ No, but not for the reason above.
9. Which of the following is false? In Markov Decision Processes, ...
   - ◯ the closer the discount factor γ to 0, the smaller the utility of future rewards.
   - ◯ the closer the discount factor γ to 0, the shorter Value Iteration may take to converge.
   - ◯ the closer the discount factor γ to 1, the greedier the optimal agent.
   - ◯ the closer the discount factor γ to 1, the longer Value Iteration may take to converge.
10. In reinforcement learning, we assume a Markov Decision process $(\mathcal{S}, \mathcal{A}, P, R)$ where ...
    - ◯ $\mathcal{S}$, $\mathcal{A}$, P and R are known.
    - ◯ $\mathcal{S}$, $\mathcal{A}$, P and R are unknown.
    - ◯ $\mathcal{S}$ and $\mathcal{A}$ are known, but P and R are unknown.
    - ◯ P and R are known, but $\mathcal{S}$ and $\mathcal{A}$ are unknown.

Question 2 [4 points] Imagine a car whose goal is to traverse a city and park at a designated parking spot. We model the city as a grid-like environment and the car as an agent located at a particular grid cell, facing one of the four cardinal directions $d \in \{N, S, E, W\}$, and moving forward with an adjustable velocity v. The car can simultaneously take turning and moving actions. The turning actions are either `left` or `right`, which change the car's direction by 90 degrees, or `none`, which keeps the car's direction the same. The moving actions are `fast` or `slow`, which change the car's velocity by $+1$ or -1 respectively, or `none`, which keeps the car's velocity the same. Any action that would result in the car moving off the grid is not permitted. Any action that would reduce v below 0 or above a maximum speed V_{max} is also not permitted. The agent's goal is to find a plan which parks the car at the designated parking spot, facing any direction, with velocity 0.

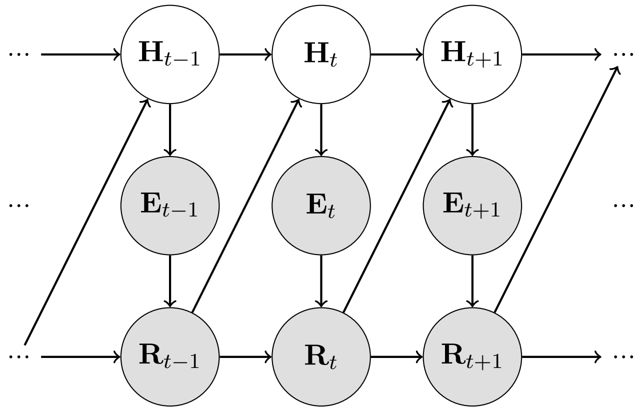
(a) If the grid is $M \times N$, what is the size of the state space? Explain your answer.

(b) Is the Manhattan distance a good heuristic function for this search problem? Explain your answer.

(c) If we used A^* *graph* search with an inadmissible heuristic, would the search be complete? Would it be optimal? Explain your answers.

(d) If we used A^* *graph* search with an admissible heuristic, would the search be optimal? If yes, prove it formally. If not, explain why not and provide a solution that would make it optimal.

Question 3 [4 points] In the near future, autonomous robots will live among us. Therefore, the robots need to know how to act appropriately in the presence of humans. In this question, we explore a simplified model of this interaction. We assume that we can observe the robot's actions at time t, $\mathbf{R}_t$, and an evidence observation, $\mathbf{E}_t$, directly caused by the (unobserved) human action, $\mathbf{H}_t$. Human actions and robot actions from the past time-step affect the human and the robot actions in the next-time step, as illustrated in the dynamic Bayesian network given below.



Assuming discrete variables and given $\mathbf{P}(\mathbf{H}_0|\mathbf{e}_0, \mathbf{r}_0)$, $\mathbf{P}(\mathbf{E}_t|\mathbf{H}_t)$, $\mathbf{P}(\mathbf{R}_t|\mathbf{R}_{t-1}, \mathbf{E}_t)$ and $\mathbf{P}(\mathbf{H}_t|\mathbf{H}_{t-1}, \mathbf{R}_{t-1})$, our goal is to derive a procedure to maintain a belief state $\mathbf{P}(\mathbf{H}_t|\mathbf{e}_{0:t}, \mathbf{r}_{0:t})$ about the state of the human at time t.

(a) Derive the prediction $\mathbf{P}(\mathbf{H}_{t+1}|\mathbf{e}_{0:t}, \mathbf{r}_{0:t})$ of the future state $\mathbf{H}_{t+1}$ of the human at time $t+1$ given a belief state $\mathbf{P}(\mathbf{H}_t|\mathbf{e}_{0:t}, \mathbf{r}_{0:t})$.

(b) From the previous prediction equation, derive a recursive update equation of the belief state $\mathbf{P}(\mathbf{H}_t|\mathbf{e}_{0:t}, \mathbf{r}_{0:t})$, as observations $(\mathbf{e}_t, \mathbf{r}_t)$ are collected and time passes.

(c) Let us now assume that all variables are continuous. Discuss how you would compute or approximate the belief state on a computer.

Question 4 [4 points] A new virtual escape game came out, and you decide to play it. You arrive in a 5×5 grid world where each cell (x, y) is a room with doors leading to the adjacent rooms. The game's goal is to reach the exit room as fast as possible, but its position is unknown. Furthermore, some regions of the world are full of riddles, and crossing rooms in these regions takes longer. Fortunately, a leaderboard with the players' best times is provided, starting from a few different rooms. Due to rounding errors, you assume that the best times reported in the leaderboard are measurements affected by additive Gaussian noise $\mathcal{N}(0, 1)$.

i	Starting room	Measured best time
1	$(4, 5)$	4
2	$(5, 3)$	6
4	$(4, 1)$	6
3	$(3, 3)$	8
5	$(1, 2)$	9

From the leaderboard, you wish to learn a heuristic approximating the best time to get to the exit, starting from room (x, y). You decide to use a small neural network as approximator, described by the following parametric function,

$$\begin{aligned} h(x, y; \phi) &= \text{ReLU}(xw_1 + yw_2 + w_3) + \text{ReLU}(xw_4 + yw_5 + w_6) \\ \text{ReLU}(x) &= \max(x, 0), \end{aligned}$$

where $\phi = (w_1, w_2, w_3, w_4, w_5, w_6)$ is the set of parameters/weights of the neural network.

(a) Among the following sets of parameters (A, B or C), which one would you use? Motivate your answer.

Set	w_1	w_2	w_3	w_4	w_5	w_6
ϕ_A	-1	0	4	-2	1	5
ϕ_B	-1	2	3	0	-1	4
ϕ_C	-2	1	4	1	-1	6

(b) You now assume a Gaussian prior $\mathcal{N}(0, 1)$ on each parameter. Which set of parameters in the table above would you now choose? Motivate your answer.

(c) Discuss the procedure you would implement on a computer to find the optimal set of parameters, had the table above not been provided. Note that the leaderboard remains available.

Question 5 [4 points] An agent is in an unknown environment where there are three states $\{A, B, C\}$ and two actions $\{0, 1\}$. We are given the following tuples (s, a, r, s'), generated by taking actions in the environment.

s	a	r	s'
A	0	$+2$	A
C	1	-2	A
B	1	$+1$	B
A	0	-1	B
B	1	-2	C
C	0	$+4$	B
B	0	$+1$	A

Assuming a discount factor $\gamma = 0.5$ and a learning rate $\alpha = 0.75$,

(a) Apply the Q-learning algorithm to obtain state-action-value $Q(s, a)$ estimates. Estimates are initialized to 0.

(b) We now switch to a feature-based estimator $\hat{Q}(s,a) = w_0 + w_1 f_1(s,a)$, with $f_1(s,a) = 2a - 1$. Starting from weights $w_0 = w_1 = 0$, update the weights according to the approximate Q-learning algorithm.

Extra page 1 / 2.

Extra page 2 / 2.