CS-E4800 Artificial Intelligence

Examination, August 6, 2020

Wrong answers to yes/no questions will give half the points negatively, e.g. for a 2-point question this means -1 points. If you are not sure about the answer, it's best to answer 'DontKnow'! Give your real valued answers with at least to two decimal places (rounded in the last digit) unless the question says otherwise. Answers to numeric questions 'close enough' will yield full points, and answers a bit off still yield some points. Note that questions may continue on the following page. Please read carefully.

Question 1 (12.00 pts) True or false?

(a) If h_1 and h_2 are admissible heuristics, then also $h(s) = (1 - \lambda)\, h_1(s) + \lambda h_2(s)$, $0 \leq \lambda \leq 1$ is admissible.
(b) For a network on the Euclidean plane, if n is the number of nodes in the network and s is the cost of the lowest cost arc in the network, then ns is an upper bound on the cost of the shortest path between any given nodes.
(c) The WA* algorithm with $W = 3$ performs at most $\frac{1}{3}$ of the expansions performed by A*.
(d) The Minimax procedure is depth-first.
(e) For a system of N states, the number of discrete belief-states is $2^N - 1$.
(f) When applying the Bellman equation to a reward sequence, short-term rewards is emphasized by setting the parameter γ, close to 1.0.

Question 2 (10.00 pts) Consider the following two games.

	C	D
A	$9, -6$	$5, -7$
B	$-2, 1$	$-5, 3$

	G	H
E	$4, -4$	$0, 0$
F	$1, -1$	$3, -2$

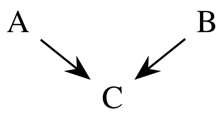
Both games have exactly one Nash equilibrium (either in pure or in mixed strategies). Find those Nash equilibria. What are the probabilities of playing the strategies A, C, E and G? Return your answer as the vector $(P(A), P(C), P(E), P(G))$.

Question 3 (10.00 pts) Explain and contrast MDPs, POMDPs, and Mote-Carlo Tree Search as decisions-making approaches.

Show how the techniques differ in their approaches to exploring state spaces. For which kind of problems is an approach insufficient? Preferable over the others? Why? (Be precise.)

Brief technical explanations and short examples suffices, but the answer should be in your own words.

Question 4 (10.00 pts) Consider the following Bayesian network and the CPTs for all nodes.



$P(A)$
0.50

$P(B)$
0.70

		$P(C\|A,B)$
A	B	0.80
A	$\neg B$	0.30
$\neg A$	B	0.90
$\neg A$	$\neg B$	0.20

Answer the following questions.

(a) What is the probability $P(A, B, C)$?
(b) What is the probability $P(\neg A, B, \neg C)$?
(c) What is the conditional probability $P(C|\neg B)$?
(d) What is the conditional probability $P(C|\neg A, B)$?

Question 5 (10.00 pts) Represent the following either in propositional logic or in the predicate logic (appropriate level of abstraction).

1. "A black cat sees a grey mouse."

2. "Werewolves only come out at full moon."
3. "All mammals are vertebrates, and some mammals lay eggs."
4. A system can be configured with some or all of components A, B, C and D. At least one of A and B must be included. Component B requires also installing component C. If component C is not included, then component D must be included. Components D and B are incompatible.

Question 6 (10.00 pts) Create and describe a state space model of the following:

You are building the world's smallest automated snack bar. The bar can seat up to three customers at the same time along a small counter in front of the kiosk window (at which your mechanized butler will serve them). Your plan to have five different drinks and eight different snacks on the menu.

Because your establishment will be so fancy, a customer can order and consume only one snack at a time. Moreover, a customer may only order a drink once they already have ordered something to eat. However, while a customer can only drink one thing at a time, they may keep ordering drinks as long as they haven't finished their snack (some of your snacks are quite spicy). For simplicity you are coding up your butler to be snobbish: the bar doesn't take reservations, if someone arrives and a place is free they may take it, and if not you don't serve them. The rest (e.g. queuing) is not your problem.

The events in your system is a customer arriving, leaving, ordering and finishing food and/or drink. (You can also assume that your mechanized butler is very efficient - customers won't have to wait for their orders, and you won't run out of any item. Also, you do not have to worry about payment, or keeping a tab on the orders.)

Considering the constraints given in the question. What state-space variables do you need to represent this problem? What conditions and effects are associated with the actions transitioning between states? Use a (semi-)formal notation of your choice to model and describe this system.

Question 7 (5.00 pts) Give a formula ϕ that satisfies the following conditions.

- ϕ is a logical consequence of $(C \to B) \leftrightarrow \neg A$.
- ϕ is not logically equivalent to $(C \to B) \leftrightarrow \neg A$.
- ϕ is not valid.

Question 8 (10.00 pts) Consider an agent using Q-learning to optimize its behaviour in an unknown environment.

The states consist of some finite set $\mathcal{S}$ and the actions available to the agent are $\mathcal{A} = \{a_1, a_2, a_3, a_4, a_5\}$.

The agent has been executing the Q-learning algorithm a few iterations with parameters leaning rate $\lambda = 0.10$, and discount factor $\gamma = 0.80$.

The current state is $x \in \mathcal{S}$.

The Q-table entries relevant in this question are:

- $Q(x, a_1) = 7.45$, $Q(x, a_2) = 20.41$, $Q(x, a_3) = 30.06$, $Q(x, a_4) = 28.38$, $Q(x, a_5) = 37.66$
- $Q(y, a_1) = 7.21$, $Q(y, a_2) = 15.27$, $Q(y, a_3) = 20.83$, $Q(y, a_4) = 24.44$, $Q(y, a_5) = 26.91$

Moreover, the agent also keeps a record of rewards for previous actions. For state x these are:

$a_1 : (1, 1, 0, 0);\ a_2 : (5, 3);\ a_3 : (1, 0, 2, 0, 0, 0);\ a_4 : (1, 10, -1);\ a_5 : (0, 3, 0, 5, 2)$.

Answer the following (a-d):

(a) If the agent was to use pure exploitation in order to select its next action at this stage of the learning process, which action would that be? (1/10 pts)

An alternative strategy to pure exploitation is for the agent to base its choice of action on the record of previous rewards.

So, instead, let the agent pick the action using UCB1.

(b) What is the best UCB1 score? (4/10 pts)

(c) Which action does this correspond to? (1/10 pts)

(d) Call this action $\alpha \in \mathcal{A}$. The agent executes α with the result that it transitions to state $y \in \mathcal{S}$ and receives reward $r = 11$. Calculate value of the updated entry $Q(x, \alpha)$ in the Q-table. (4/10 pts)

Exam Rules

1. Any **communication** with other people by any means is **not allowed**.
2. You are allowed to use the CS-E4800 course material and general sources such as Wikipedia or the Russell-Norvig textbook, but no other written sources.
3. Use of calculator is allowed.

Grading

Maximum from the exam (excluding bonus questions) is 12.00+10.00+10.00+10.00+10.00+10.00+5.00+10.00 = 77.00 points. These points are not directly comparable to the exercise points, and will be scaled and aggregated with the exercise points to determine the course grade.