INFO8006 Introduction to Artificial Intelligence

Exam of January 2020

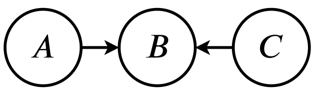
Instructions

- *Duration: 4 hours.*
- *Answer the questions on separate sheets, labeled with the question number, your first name, last name and student id. Answer in English or in French.*
- *Non-programmable calculators are allowed. Notes or documents of any kind are forbidden.*

Question 1 [4 points]

Multiple choice questions. Choose one of the four choices. 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. A rational agent is an agent that chooses whichever action that maximizes the expected value of its performance measure, given the percept sequence to data. Therefore,
   (a) rationality is equivalent to omniscience.
   (b) rationality is equivalent to clairvoyance.
   (c) rationality necessarily leads to success.
   (d) rationality may lead to exploration.
2. Consider the game of Pacman in the absence of ghosts and with a single food dot in the maze. In state s, 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?
   (a) $h_1(s) = (i_s - x_s)^2 + (j_s - y_s)^2$
   (b) $h_2(s) = |i_s - x_s|$
   (c) $h_3(s) = |j_s - y_s|$
   (d) $h_4(s) = \max(h_2(s), h_3(s))$
3. In adversarial search,
   (a) the deeper in the tree the evaluation function is buried, the more the quality of the evaluation matters.
   (b) if not looked deep enough, good moves may appear as bad moves, because their consequences are hidden beyond the search horizon.
   (c) the horizon effect arises when the evaluation is perfect.
   (d) the search horizon is irrelevant if and only if the evaluation function is admissible but not consistent.
4. The Bayes' rule states that ...
   (a) $P(a|b) = P(b|a)P(b) \,/\, P(a)$.
   (b) $P(a|b) = P(b) \,/\, P(b|a)P(a)$.
   (c) $P(b|a) = P(a|b)P(a) \,/\, P(b)$.
   (d) $P(a|b) = P(b|a)P(a) \,/\, P(b)$.
5. Consider the Bayesian network below. Which of the following is always true?



   (a) A and C are causes of B.
   (b) A or C are an effect of B.
   (c) A and C are independent.
   (d) A and C are dependent given B.

6. The task of smoothing consists in computing
   (a) $\arg\max_{\mathbf{x}_{1:t}} P(\mathbf{x}_{1:t}|\mathbf{e}_{1:t})$.
   (b) $\mathbf{P}(\mathbf{X}_{t+k}|\mathbf{e}_{1:t})$, for $k > 0$.
   (c) $\mathbf{P}(\mathbf{X}_t|\mathbf{e}_{1:t})$.
   (d) $\mathbf{P}(\mathbf{X}_k|\mathbf{e}_{1:t})$, for $k < t$.
7. In a Markov decision process, the reward function...
   (a) can shape the optimal agent's behavior from risk-taking to conservative.
   (b) has no effect on the optimal agent's behavior.
   (c) cannot be negative.
   (d) produces sequences of values that sum to zero.
8. If $\mathbf{y}_{1:T}$ is an audio signal and $\mathbf{w}_{1:L}$ represents a sequence of words, then speech recognition can be cast as a Bayesian inference problem which consists in solving
   (a) $\hat{\mathbf{w}}_{1:L} = \arg\max_{\mathbf{w}_{1:L}} P(\mathbf{y}_{1:T}|\mathbf{w}_{1:L})$.
   (b) $\hat{\mathbf{w}}_{1:L} = \arg\max_{\mathbf{w}_{1:L}} P(\mathbf{w}_{1:L}|\mathbf{y}_{1:T})$.
   (c) $\hat{\mathbf{y}}_{1:T} = \arg\max_{\mathbf{y}_{1:T}} P(\mathbf{y}_{1:T}|\mathbf{w}_{1:L})$.
   (d) $\hat{\mathbf{y}}_{1:T} = \arg\max_{\mathbf{y}_{1:T}} P(\mathbf{w}_{1:L}|\mathbf{y}_{1:T})$.
9. The search algorithm in AlphaGo is based on...
   (a) the Minimax algorithm.
   (b) Monte Carlo tree search.
   (c) A*.
   (d) a Markov Chain Monte Carlo method.
10. AlphaGo makes use of...
   (a) a policy network as evaluation function for estimating the state utility and a value network as a cutoff test to decide when to stop expanding a state.
   (b) a value network as evaluation function for estimating the state utility and a policy network as a cutoff test to decide when to stop expanding a state.
   (c) a policy network as evaluation function for estimating the state utility and a value network for guiding the exploration of the search tree.
   (d) a value network as evaluation function for estimating the state utility and a policy network for guiding the exploration of the search tree.

Question 2 [4 points]

It is training day for Pacbabies (a.k.a. Hungry Running Maze Games day)! Each of the k Pacbabies starts in its own assigned start location s_i in a large maze of size $M \times N$ and must return to its own Pacdad who is waiting patiently at g_i. Along the way, the Pacbabies must, between them, eat all the food dots in the maze cooperatively.

At each step, all k Pacbabies move one unit to any open adjacent square. The only legal actions are Up, Down, Left or Right. It is illegal for a Pacbaby to wait in a square, attempt to move into a wall, or attempt to occupy the same square as another Pacbaby. To set a record, the Pacbabies must find an optimal collective solution. None of the dots are placed at the start locations s_i or at the Pacdad locations g_i.

(a) Define a minimal state space representation for this problem. How large is the state space?

(b) What is the maximum branching factor for this problem?

(c) Let $d(p, q)$ be the Manhattan distance between positions p and q, F be the set of all positions of remaining food pellets, and p_i be the current position of Pacbaby i. Which of the following are admissible heuristics? Explain and motivate your answers briefly.
   i. $h_A = \frac{1}{k} \sum_i = 1^k d(p_i, g_i)$
   ii. $h_B = \max_{1 \leq i \leq k} d(p_i, g_i)$
   iii. $h_C = \max_{1 \leq i \leq k} [\max_{f \in F} d(p_i, f)]$
   iv. $h_D = \max_{1 \leq i \leq k} [\min_{f \in F} d(p_i, f)]$
   v. $h_E = \min_{1 \leq i \leq k} [\min_{f \in F} d(p_i, f)]$

vi. $h_E = \min_{f \in F} [\min_{1 \leq i \leq k} d(p_i, f)]$

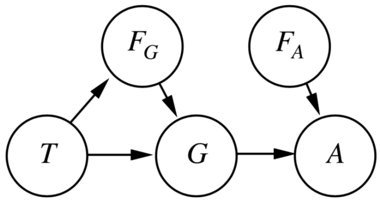
Now suppose that some of the squares are flooded with water. In the flooded squares, it takes two timesteps to travel through the square, rather than one. However, the Pacbabies do not know which squares are flooded and which are not, until they enter them. After a Pacbaby enters a flooded square, its howls of despair instantly inform all the other Pacbabies of this fact.

(d) Define a minimal space of belief states for this problem. How large is the belief state space?

(e) Discuss and motivate the prior you would choose, before the Pacbabies receive any wetness percepts.

Question 3 [4 points]

In your local nuclear power station, there is an alarm that senses when the temperature gauge exceeds a given threshold. The gauge measures the temperature of the core. Consider the Boolean variables A (the alarm sounds), F_A (the alarm is faulty), and F_G (the gauge is faulty), along with the multivalued variables G (the gauge reading) and T (the actual core temperature). Given that the gauge is more likely to fail when the core temperature gets too high, we model the domain using the Bayesian network below.



(a) Is the Bayesian network above the unique Bayesian network that can represent the domain? If yes, explain why. If not, provide an alternative network as counter-example.

(b) From the network topology above, determine which of the following is true, false, or cannot be determined.

i. $T \perp F_G$

ii. $F_G \perp F_A$

iii. $T \perp F_A | A$

iv. $F_G \perp F_A | \{A, G\}$

v. $A \perp T | F_G$

vi. $T \perp F_G | A$

(c) Suppose there are just two possible actual and gauged temperatures, normal and high; the probability that the gauge gives the correct temperature is x when it is working, but y when it is faulty. Give the conditional probability table associated with G.

(d) Suppose the alarm works correctly unless it is faulty, in which case it never sounds. Give the conditional probability table associated with A.

(e) Suppose the alarm and gauge are working and the alarm sounds. Calculate an expression for the probability that the temperature of the core is too high, in terms of the various conditional probabilities in the network.

Question 4 [4 points]

(a) Define mathematically the inference task of filtering. Discuss how it can be useful to an agent.

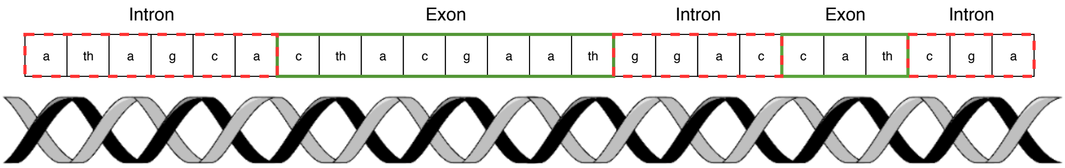
(b) Derive the recursive update equation of the Bayes filter, assuming discrete variables.

(c) Explain how the Bayes filter update equation can be rewritten in terms of matrix and vector operations when the state variable $\mathbf{X}_t$ and the evidence variable $\mathbf{E}_t$ are both single discrete random variables.

(d) Assume a caricature of the DNA molecule (see the figure below) consisting of alternating introns and exons. In short, introns and exons are both nucleotide sequences in DNA and RNA with the difference that introns do not directly code for proteins whereas exons do. The nucleotide sequences of introns and exons have different nucleotide compositions. From the observed sequence of nucleotides we wish to infer the unobserved sequence of introns/exons. Hereto we assume that the nucleotide sequence can be modeled by the following hidden Markov model. Let $D_X = \{i, e\}$ be the set of state values (corresponding to intron and exon) of the hidden Markov chain, $\mathbf{f}_0$ the prior state probabilities, $\mathbf{T} = P(X_t | X_{t-1})$ the transition matrix, $D_E = \{a, c, g, th\}$ the emission alphabet, and $\mathbf{B} = \mathbf{P}(E_t | X_t)$ emission matrix such that:

$$\mathbf{f}_0 = (0.5, 0.5)^T, \quad \mathbf{T} = \begin{pmatrix} P(X_t = i | X_{t-1} = i) & P(X_t = e | X_{t-1} = i) \\ P(X_t = i | X_{t-1} = e) & P(X_t = e | X_{t-1} = e) \end{pmatrix} = \begin{pmatrix} 0.6 & 0.4 \\ 0.1 & 0.9 \end{pmatrix},$$

$$\mathbf{B} = \begin{pmatrix} P(E_t = a | X_t = i) & P(E_t = c | X_t = i) & P(E_t = g | X_t = i) & P(E_t = th | X_t = i) \\ P(E_t = a | X_t = e) & P(E_t = c | X_t = e) & P(E_t = g | X_t = e) & P(E_t = th | X_t = e) \end{pmatrix} = \begin{pmatrix} 0.49 & 0.01 & 0.49 & 0.01 \\ 0.01 & 0.49 & 0.01 & 0.49 \end{pmatrix}.$$



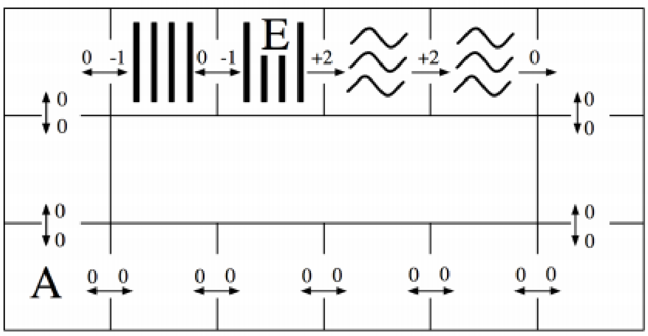
i. Derive the stationary distribution of the transition model, if any.
ii. Given the sequence of observations $E_1 = a$, $E_2 = a$ and $E_3 = th$, compute the probability of the last nucleotide $E_3 = th$ being part of an exon.
iii. A new scientific paper pretends to have discovered better parameter values of the hidden Markov chain. Suppose that you are given a long sequence of nucleotides and their associated hidden states (introns/exons). How would you compare the two models? Explain in details.
iv. The conclusion you reached above with only one DNA sequence could be criticized. Discuss what would be the advantages, if there are some, of using more than one sequence.

Question 5 [4 points]

(a) Sometimes MDPs are formulated with a reward function $R(s, a)$ defined as the reward obtained by performing the action a from state s. Write the Bellman equations for this formulation.

(b) Describe (with pseudo-code) the Value Iteration algorithm for solving these new Bellman equations.

(c) Consider the MDP drawn below. The state space consists of all squares in a grid-world water park. There is a single waterslide that is composed of two ladder squares and two slide squares (marked with vertical bars and squiggly lines respectively). An agent in this water park can move from any square to any neighboring square, unless the previous or the current square is a slide in which case it must move forward one square. The actions are denoted by arrows between squares on the map and all deterministically move the agent in the given direction. The agent cannot stand still: it must move on each time step. Rewards $R(s, a)$ are also shown below: the agent feels great pleasure as it slides down the water slide (+2), a certain amount of discomfort as it climbs the rungs of the ladder (-1), and receives rewards of 0 otherwise.



i. How many deterministic policies π are possible for this MDP?
ii. Fill in the blank cells of this table with the state values associated with the optimal policy which depends on the discount and horizon. The horizon is to the total number of actions that the agent can take before the MDP terminates.

	γ	horizon	$s = A$	$s = E$
V(s)	1.0	3		
V(s)	1.0	10		
V(s)	0.1	10		
V(s)	1.0	∞		
V(s)	0.1	∞		

iii. For $\gamma = 0.99$ and an infinite horizon, provide the optimal policy π^*.