

Calculators may be used in this examination provided they are not capable of being used to store alphabetical information other than hexadecimal numbers

UNIVERSITY OF BIRMINGHAM

School of Computer Science

Artificial Intelligence 1

Main Summer Examinations 2023

Time allowed: 2 hours

[Answer all questions]

Note

Answer ALL questions. Each question will be marked out of 20. The paper will be marked out of 80, which will be rescaled to a mark out of 100.

Question 1 Clustering

Suppose we have five observations, $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \mathbf{x}^{(3)}, \mathbf{x}^{(4)}$ and $\mathbf{x}^{(5)}$, for which we compute the following dissimilarity matrix

$$D = \begin{matrix} & \mathbf{x}^{(1)} & \mathbf{x}^{(2)} & \mathbf{x}^{(3)} & \mathbf{x}^{(4)} & \mathbf{x}^{(5)} \\ \mathbf{x}^{(1)} & 0 & 0.3 & 0.6 & 0.7 & 0.8 \\ \mathbf{x}^{(2)} & & 0 & 0.1 & 0.2 & 0.4 \\ \mathbf{x}^{(3)} & & & 0 & 0.45 & 0.25 \\ \mathbf{x}^{(4)} & & & & 0 & 0.9 \\ \mathbf{x}^{(5)} & & & & & 0 \end{matrix},$$

- (a) Based on the given dissimilarity matrix, hierarchically cluster the observations using **complete linkage**. Sketch the dendrogram, clearly illustrating the height at which each cluster fusion occurs. **[10 marks]**
- (b) Assume that a clustering algorithm produces the following two clusters: $C_1 = (\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ and $C_2 = (\mathbf{x}^{(3)}, \mathbf{x}^{(4)}, \mathbf{x}^{(5)})$. For each of the observations in cluster C_1 , compute the Silhouette coefficient using the information from the dissimilarity matrix. Comment on the suitability of their assignment to cluster C_1 .

Note: The Silhouette coefficient of an observation \mathbf{x} is defined as:

$$SC(\mathbf{x}) = \frac{b(\mathbf{x}) - a(\mathbf{x})}{\max\{b(\mathbf{x}), a(\mathbf{x})\}},$$

where $a(\mathbf{x})$ is the average dissimilarity of \mathbf{x} with respect to other observations in its cluster, and $b(\mathbf{x})$ is the minimum average dissimilarity of \mathbf{x} with respect to all clusters to which it does not belong. **[10 marks]**

Question 2 Supervised Learning

- (a) The following pseudo-code represents one iteration through the training set for gradient descent applied to univariate linear regression.

```

1 cost = 0;
2 w0 = 0;
3 w1 = 0;
4 for j=1 to size(trainingSet) do
5     f = w0 + w1*x(j);
6     cost = cost + (y(j) - f)2;
7     w0 = w0 -  $\alpha$  * (f - y(j));
8     w1 = w1 -  $\alpha$  * (f - y(j)) * x(j);

```

Assume that the value of the learning rate, α is 1.

Give the numerical values of 'cost', 'w0', and 'w1' at the end of the execution of this pseudo-code for the following training set: $\{(-3, -1), (1, 1), (2, 5)\}$. **Show all your working.** [10 marks]

- (b) Suppose that you want to use a k-Nearest Neighbours classification, but the distance metric is not explicitly specified to you. Instead, you are given a "black box" where you input a set of points $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(N)}$ and a new point \mathbf{x} , and the black box outputs the nearest neighbour of \mathbf{x} and its corresponding class label.

Can you construct a k-Nearest Neighbour classification algorithm based on this black box alone? The suggested algorithm should work for any given value of $k \in \{1, 2, \dots, N\}$. If so, describe your algorithm. If not, explain why not? [10 marks]

Question 3 Optimisation

- (a) Give one strength and one weakness of Hill Climbing. **Justify** your answer. **[10 marks]**
- (b) Consider the following optimisation problem and algorithm design to solve it:

Optimisation Problem

A builder has N possible clients. Each client $i \in \{1, \dots, N\}$ would like the builder to undertake a job that takes h_i hours to complete. The builder wishes to decide which clients to accept over the next 30 days so as to maximise their income while ensuring that they work no more than 7.5 hours per day. This problem can be mathematically formulated as follows:

$$\begin{aligned} & \text{maximise} && f(\mathbf{x}) \\ & \text{subject to} && g(\mathbf{x}) = \frac{1}{30} \left(\sum_{i=1}^N x_i h_i \right) - 7.5 \leq 0 \end{aligned}$$

where \mathbf{x} is a vector of size N , and $\forall i \in \{1, \dots, N\}$, $x_i = 0$ if the builder does not accept job i and $x_i = 1$ otherwise; and $f(\mathbf{x})$ is a function that calculates the income.

Simulated Annealing Algorithm Design

Representation: a direct representation of the design variable \mathbf{x} . In other words, a vector \mathbf{x} of size N , where $\forall i \in \{1, \dots, N\}$, $x_i = 0$ if the builder does not accept job i and $x_i = 1$ otherwise.

Algorithm 1: Initialisation Procedure.

Input: Number of possible clients N .

Output: Candidate solution \mathbf{x} .

- 1 \mathbf{x} = new vector of size N ;
 - 2 **for** $i=1$ to N **do**
 - 3 x_i = value picked uniformly at random from $\{0, 1\}$;
 - 4 **return** \mathbf{x}
-

Algorithm 2: Neighbourhood Operator.

Input: Current solution \mathbf{x} ; number of possible clients N .

Output: Neighbour \mathbf{x}' .

- 1 \mathbf{x}' = copy of \mathbf{x} ;
 - 2 i = value picked uniformly at random from $\{1, \dots, N\}$;
 - 3 $x'_i = 1 - x_i$;
 - 4 **if** $g(\mathbf{x}') > 0$ **then**
 - 5 $x'_i = 0$;
 - 6 **return** \mathbf{x}'
-

Question 3 continued over the page

Are the representation, initialisation and neighbourhood operators correctly designed (i.e., suitable) for this problem? Assume that we wish to deal with the constraints of this problem based on the design of the algorithm's operators. **Justify** your answer by explaining either why all three operators are suitable, or what is wrong.

Note:

- The problem formulation correctly reflects the intended problem, i.e., you do not need to check whether the problem formulation itself is correct.
- You do not need to consider how efficient the design of the operators is, just whether or not it is a correct design for the problem.
- It is acceptable for the neighbourhood operator to sometimes generate a neighbour that is the same as the current solution.

[10 marks]

Question 4 Search

- (a) Consider a problem with two possible actions: a and b . The cost of a is 2 and the cost of b is 1. The shallowest goal node can be at any level of the tree and the problem solution may involve a sequence of these actions. Is breadth-first search optimal to solve this problem, given this cost function? **Justify** your answer.

[10 marks]

- (b) An agent must traverse a maze in order to find and collect a treasure. This task can be formulated as a search problem as in the following:

- Initial state: the agent is in $A1$.
- Goal test: the agent is in $E5$ (automatically collects the treasure).
- Actions: $move(n, n')$ moves the agent from state n to a state n' that is to the left of, or the right of, or above or below state n .
- Transition model: see figure below. States depicted in black, e.g., $A2$, are considered obstacles and cannot be traversed by the agent.
- Path cost: each action has cost 1.

Generate the A^* tree until the goal node is found.

	1	2	3	4	5
A	o				
B					
C					
D					
E					X

Initial State

	1	2	3	4	5
A					
B					
C					
D					
E					o

Goal State

Question 4 continued over the page

The heuristic to be used is the following:

$$h(n) = d_M(n, n_{goal}) = \sum_{i=1}^3 |p_i - q_i|,$$

where $d_M(n, n_{goal})$ represents the Manhattan distance between node n and the goal node n_{goal} . When comparing the first coordinate of each state, assume the following values for each of the possible letters in the grid: $A = 1$, $B = 2$, $C = 3$, $D = 4$ and $E = 5$. Order of expansion: if $f(n)$ produces the same lowest value for two nodes, expand the nodes in alphanumerical order (e.g., $A2$ expanded before $B1$). Additionally, if $g(n)$ also produces the same lowest value for two nodes, keep the node that has been added last into the frontier. Recall that the evaluation function $f(n) = g(n) + h(n)$ is the sum of the cost to reach node n and the heuristic calculated at node n .

Write down the following:

- Search tree produced by A^* , indicating the $f(n)$, $g(n)$ and $h(n)$ values of each node n ; which nodes are in the frontier when the algorithm terminates; and which nodes have been pruned as a result of the above instructions.
- The solution retrieved by A^* and its cost.
- The sequence of nodes visited by A^* in the order they are visited. Note: you can identify a node through its state, e.g., $A1$ or $B1$.

[10 marks]

Do not complete the attendance slip, fill in the front of the answer book or turn over the question paper until you are told to do so

Important Reminders

- Coats/outwear should be placed in the designated area.
- Unauthorised materials (e.g. notes or Tippex) must be placed in the designated area.
- Check that you do not have any unauthorised materials with you (e.g. in your pockets, pencil case).
- Mobile phones and smart watches must be switched off and placed in the designated area or under your desk. They must not be left on your person or in your pockets.
- You are not permitted to use a mobile phone as a clock. If you have difficulty seeing a clock, please alert an Invigilator.
- You are not permitted to have writing on your hand, arm or other body part.
- Check that you do not have writing on your hand, arm or other body part – if you do, you must inform an Invigilator immediately
- Alert an Invigilator immediately if you find any unauthorised item upon you during the examination.

Any students found with non-permitted items upon their person during the examination, or who fail to comply with Examination rules may be subject to Student Conduct procedures.