SEMESTER 1 FINAL ASSESSMENT 2021/22

INTELLIGENT SYSTEMS

Duration 2 hrs

Answer 3 Qs only (marks out of 75)

Question 1

A student is writing a program to solve a puzzle. In the puzzle there are 5 moves that can be made from each state. At least one solution exists at a depth of 6 moves and all other solutions are deeper. The rules of the puzzle ensure that after 60 moves from the start position there are no further moves possible (i.e. maximum depth of the search tree=60). Each move has a cost of 8.

For the following questions, where appropriate show the relevant formula and give a numerical answer (i.e. show that you can plug in the correct numbers to the formulae you know).

The first approach the student considers is to use tree search with a breadth first strategy (BFS).

a) How many nodes will this method generate in the worst case? [2 marks]

b) What is the maximum number of nodes that will be stored in memory at any one time? [2 marks]

c) What will be the cost of the solution that it finds (if it finds one)? [2 marks]

The second approach the student considers is to use iterative deepening search (IDS).

d) How many nodes will this method generate in the worst case? [2 marks]

e) What is the maximum number of nodes that will be stored in memory at any one time? [2 marks]

f) What will be the cost of the solution that it finds (if it finds one)? [2 marks]

The student contemplates their options. Their computer can generate 100 nodes per second and store up to 1000 nodes at a time. The student is prepared to wait up to 1 hour for an answer (or 360,000 nodes generated).

g) Explain whether either of these methods can find a solution given the time available. Explain whether either of these methods can find a solution given the memory resources available. Accordingly, which of these methods, if any, is able to find a solution given these resources? [10 marks]

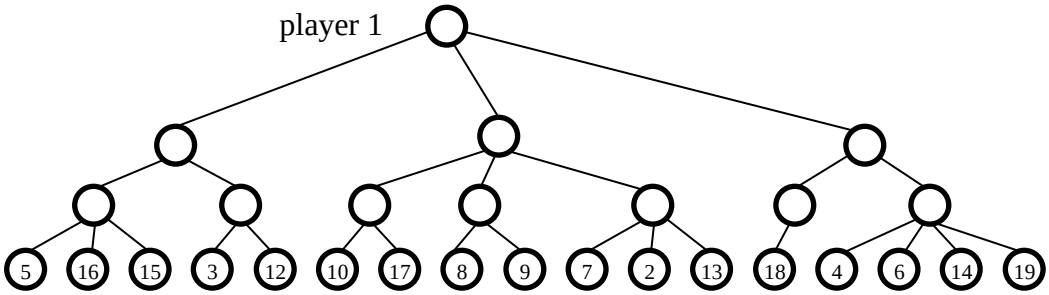
h) If it was possible to compute an estimate of how many actions it takes to find a solution and this method never over-estimates, what tree search method could be used to find optimal solutions? How does this method use the estimate to order nodes for expansion? [3 marks]

Question 2

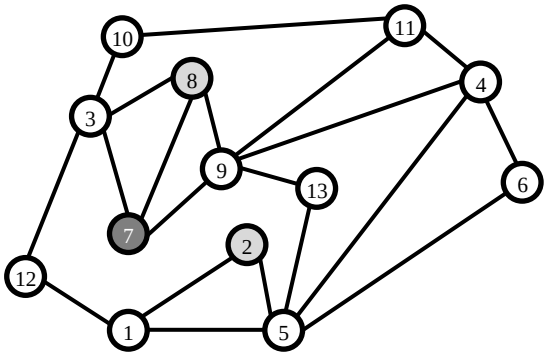
a) The following shows a game tree. Player 1 is maximising and it is player 1's turn to move.

i) What is the value of the game to Player 1? [1 marks]

ii) Using the minimax algorithm with alpha-beta pruning, and assuming conventional left-to-right traversal of the tree, indicate the branches that can be pruned. [9 marks]



b) The following figure shows a partially completed graph-colouring problem using 3 colours. Nodes 8 and 2 have been assigned one colour, and node 7 has been assigned a different colour. Other nodes do not yet have a colour assigned. Using the common heuristics we discussed for this task, which node do you assign a colour to next? Explain briefly. [3 marks]



c) The backtracking search algorithm for constraint satisfaction problems has a successor function that generates daughter nodes in the search tree. Does it generate a node for each of the values that can be assigned to:

- every remaining unassigned variable or
- only one remaining unassigned variable?

Explain briefly. [5 marks]

d) Briefly describe some points for and/or against the statement that "The Turing Test is a good test of intelligence" [7 marks]

Question 3

a) Explain verbally how the perceptron algorithm works. [4 marks]

b) A data set of 8 training points is shown in the table below, e.g., the input feature vector corresponding to ID=6 is $\underline{\boldsymbol{x}}_6 = (0.4, -0.9)^T$ and the output class label is $y_6 = (-1)$.

ID	1	2	3	4	5	6	7	8
$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$	$\begin{pmatrix} -1.2 \\ 1.4 \end{pmatrix}$	$\begin{pmatrix} -1.4 \\ -1.1 \end{pmatrix}$	$\begin{pmatrix} 0.1 \\ 2.8 \end{pmatrix}$	$\begin{pmatrix} -2.7 \\ -2.0 \end{pmatrix}$	$\begin{pmatrix} 1.4 \\ -0.4 \end{pmatrix}$	$\begin{pmatrix} 0.4 \\ -0.9 \end{pmatrix}$	$\begin{pmatrix} 2.6 \\ -0.1 \end{pmatrix}$	$\begin{pmatrix} -3.0 \\ 0.9 \end{pmatrix}$
y	+1	-1	+1	-1	+1	-1	+1	-1

Table 1. Training data for perceptron

Using $\boldsymbol{w} = (w_0, w_1, w_2) = (0.1, 0.3, 0.1)$ as your initial proposal for a separating hyperplane $w_0 + w_1 x_1 + w_2 x_2 = 0$, answer the following questions related to the perceptron.

(i) What is the suitable choice of f in the classification rule $\hat{y} = f((x_1, x_2), \boldsymbol{w})$? [2 marks]

(ii) How many training points in Table 1 are misclassified? [2 marks]

(iii) Using a **learning rate** of 0.05, sequentially perform 1 update step of the perceptron algorithm for **each** of the misclassified points. [11 marks]

(iv) How many training points are now misclassified? [1 mark]

c) Explain why the perceptron architecture cannot represent the XOR function. What elements must be added to the perceptron to solve the XOR problem? [5 marks]

Question 4

a) Describe the key ingredients in multi-armed bandit theory. What is the basis of making rational decisions about which action to take? [4 marks]

b) Explain the basis of the exploration versus exploitation trade-off in the bandit problem. Illustrate your answer with an example task. [4 marks]

c) The following questions are based on the following joint probability distribution for 3 random binary-valued random variables A, B and C.

A	B	C	P (A, B, C)
0	0	0	0.024
0	0	1	0.324
0	1	0	0.056
0	1	1	0.216
1	0	0	0.096
1	0	1	0.036
1	1	0	0.224
1	1	1	0.024

(i) Show that A and B are not independent random variables. [4 marks]

(ii) What does it mean for A and B to be conditionally independent given C? [2 marks]

(iii) Show that A and B are conditionally independent given C. It is sufficient to show this for any one choice of values of A, B and C. [6 marks]

(iv) What are the 3-node Bayesian networks (with each random variable A, B and C represented as a node) that would be consistent with the probabilities shown in the table? How would you justify your answer? [5 marks]

END OF PAPER

