Office Use Only

MONASH University

Semester One 2020
Final Assessment Period

Faculty of Information Technology

EXAM CODES: **FIT5047**

TITLE OF EXAM: **Fundamentals of Artificial Intelligence**

EXAM DURATION: **2 hours 10 minutes or 130 minutes**

THIS PAPER IS FOR STUDENTS STUDYING AT: (tick where applicable)

O Caulfield O Clayton O Parkville O Peninsula
O Monash Extension O Off Campus Learning O Malaysia O Sth Africa
● **Suzhou (China)**

During an exam, you must not have in your possession any item/material that has not been authorised for your exam. Any authorised items are listed below.

You must not retain, copy, memorise or note down any exam content for personal use or to share with any other person by any means following your exam.

As a student, and under Monash University's Student Academic Integrity procedure, you must undertake your in-semester tasks, and end-of-semester tasks, including exams, with honesty and integrity. In exams, you must not allow anyone else to do work for you and you must not do any work for others. You must not contact, or attempt to contact, another person in an attempt to gain unfair advantage during your exam session. Assessors may take reasonable steps to check that your work displays the expected standards of academic integrity. e.g. perform similarity checking on your submission. If required, you may be contacted after your exam to discuss any concerns.

Failure to comply with the above instructions or attempting to cheat or cheating in an exam is a discipline offence under Part 7 of the Monash University (Council) Regulations, or a breach of instructions under Part 3 of the Monash University (Academic Board) Regulations.

AUTHORISED MATERIALS

OPEN BOOK	● **YES**	□ **NO**
CALCULATORS	● **YES**	□ **NO**
SPECIFICALLY PERMITTED ITEMS	● **YES**	□ **NO**

if yes, items permitted are:
Blank paper, Document Scanner or mobile phone with camera
You need to use camera or scanner to photograph or scan your hand-written answers.

Candidates must complete this section if required to write answers within this paper

STUDENT ID: __ __ __ __ __ __ __ __ __ DESK NUMBER: __ __ __ __ __

Assessment Marks

- This exam accounts for **60% of the total assessment** in FIT5047.
- Total marks for the exam is **100**.

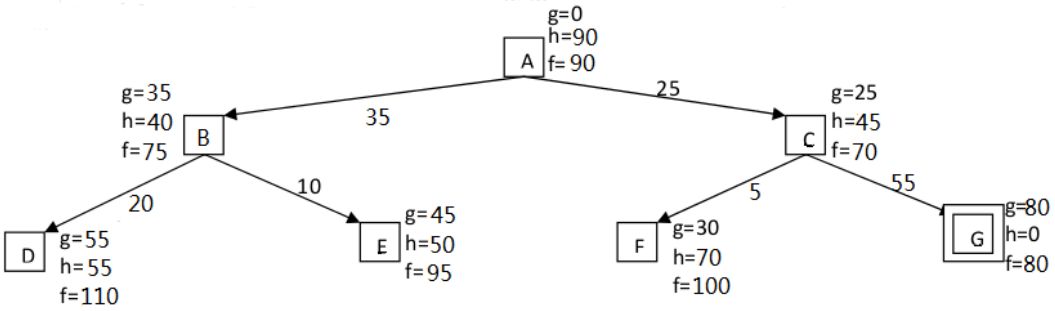
Instructions

- Answer all 10 questions.
- Please create a Microsoft Word (.docx) file as **the answer file** for all your answers.
- Please **type your answers** in the answer file, where appropriate as indicated in the questions.
- If you need to **handwrite your answers** such as drawing diagrams or equations, do this on blank paper.
  Photograph or scan your handwritten answers (using a phone, tablet, camera, or a scanner) and insert them into the answer file.
- Please make sure you **label your answers** to each question/sub-question **with question and sub-question number or letter** (e.g. "Q1 (a)", "Q2").
- Please make sure your answers in the answer file are **in order of question number**.
- For all questions, partial marks are given for working, even if your final answer is incorrect.
- Please name your answer file (in .docx) as

  STUDENT ID - UNIT CODE - Exam S1 2020
  (i.e. 12345678-FIT5047-Exam S1 2020.docx)
- Please **upload your answer file** to the FIT5047 Moodle Exam site.
- It will take time to photograph or scan your hand-written answers and upload your answer file to the Moodle Exam site.

Question 1 **[2 + 6 + 4 + 1 = 13 marks]**

Use the following (fully expanded) search tree to indicate the order in which nodes are expanded for different types of search. Assume that A is the start node and G (double boxed) the only goal node. How many nodes are there in the resultant search tree?
Your answer should be typed and not handwritten.



(a) Depth-first search

List the nodes according to their order of expansion and brief the logic.

List the nodes in the final search tree (without the nodes deleted by the algorithm).

(b) Uniform cost search

List the nodes according to their order of expansion. For each expansion, list OPEN (with the nodes in the correct order) and CLOSED and brief the logic.

List the nodes in the final search tree

(c) Algorithm A

List the nodes according to their order of expansion. For each expansion, list OPEN (with the nodes in the correct order) and CLOSED and brief the logic.

List the nodes in the final search tree.

(d) Is the heuristic h monotonic? Why or why not? (No marks will be given for an absent or incorrect explanation)

Question 2 **[4.5 marks]**

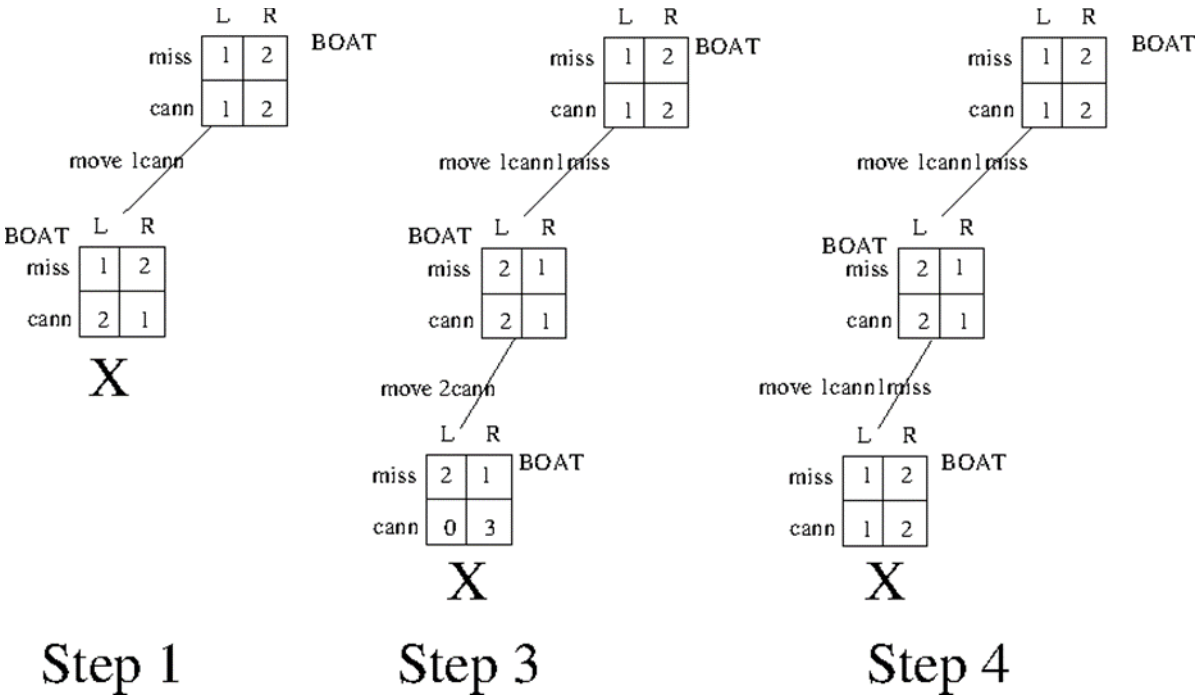
Consider the missionaries and cannibals problem:

Three missionaries and three cannibals come to a river. There is a boat on their side of the river that can be used either by one or two persons. How should they use this boat to cross the river in such a way that cannibals never outnumber missionaries on either side of the river?

The diagram below shows a sequence of partial solutions produced by the Backtrack algorithm, where each state marked with **X** is a state where the algorithm has backtracked. In these solutions, cannibals are labeled "cann", missionaries are labeled "miss", the banks are labeled L (left) and R (right), and the position of the boat is indicated with the label BOAT. The objective is to reach side R.
Label each state marked with **X** with the reason for backtracking, using one or more of the following labels: DEADEND, BOUND REACHED, PREVIOUS STATE, NO MORE APPLICABLE ACTIONS.

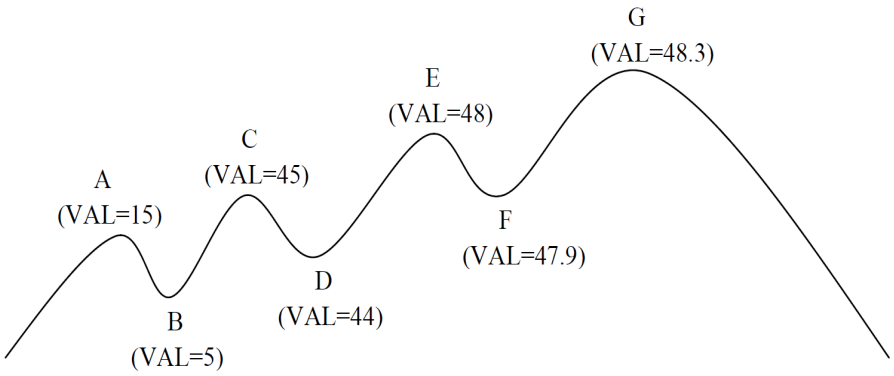
You must also explain why this label happened with reference to the state of the algorithm. An answer with an incorrect label will incur a loss of 1 mark. Marks cannot be below zero.

Your answer should be typed and not handwritten.



Question 3 **[8 marks]**

Perform Simulated Annealing search to maximize value in the following search space.
Recall that a good move (increases value) is always accepted (P = 1.0); a bad move (decreases value) is accepted with probability $P = e^{\Delta VAL/T}$, *where* $\Delta VAL = VAL(Next) - VAL(Current)$.



Use this temperature schedule:

Time Step	*1–100*	*101–200*	*201–300*
Temperature (T)	*10*	*1.0*	*0.1*

This table of values of e may be useful:

x	*0.0*	*−1.0*	*−4.0*	*-4.3*	*−40.0*	*−43.0*
e^x	*1.0*	*≈0.37*	*≈0.018*	*≈0.014*	$\approx 4.0*10^{-18}$	$\approx 2.1*10^{-19}$

(a) Complete the following possible moves in the search. **An answer with an incorrect value will incur a loss of 0.5 mark. Marks cannot be below zero.**

Time	From	To	T	*ΔVAL*	*ΔVAL/T*	*P*
57	A	B				
78	C	B				
132	C	B				
158	C	D				
194	E	D				
194	E	B				
238	E	D				
263	E	F				
289	G	F				
289	G	D				

(b) *At Time=100, is the search more likely to be in state A or in state C? (ignore E, G) And why?*

(c) *At Time=200, is the search more likely to be in state A, C, or E? (ignore G) And why?*

(d) *At Time=300, is the search more likely to be in state A, C, E, or G? And why?*

Your answer should be typed and not handwritten.

Question 4 **[2 + 3 + 1 + 3 + 2 + 1.5 = 12.5 marks]**

Ingrid wants to schedule a day for playing tennis next week. She collects data about some past days and their suitability for playing tennis based on the weather and wind condition.

Your answer should be typed and not handwritten.

Weather (WE)	Wind (WI)	Tennis (T)
rainy (ra)	weak (wk)	no (N)
cloudy (cl)	strong (st)	no (N)
cloudy (cl)	weak (wk)	yes (Y)
sunny (su)	strong (st)	yes (Y)

(a) Given the formula for Entropy:

$$H(X) = -\sum_{i=1}^{n} \Pr(x_i) \log_2 \Pr(x_i)$$

i. What is the entropy of the random variable *Tennis*: *H*(*Tennis*)? **Spell out the formula for this calculation before you plug-in numbers.**

ii. Calculate the Information Gain if the attribute *Weather* is used as the root of a Decision Tree that is used to determine whether Ingrid can play tennis. **Spell out the formula for this calculation before you plug-in numbers.**

iii. Assuming that *Weather* and *Wind* are the only two variables, do you think splitting on *Weather* is a good idea? Why or why not? **(No marks will be given for an absent or incorrect explanation)**.

(b) Given the Na¨ıve Bayes formula

$$\Pr(C|v_{i1}, \ldots, v_{in}) = \alpha \prod_{k=1}^{n} \Pr(v_{ik}|C_i = c) \Pr(C_i = c) ,$$

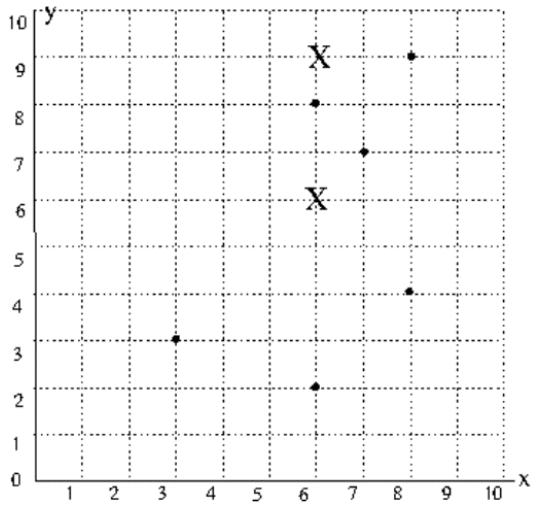
i. Given the four instances in the above table, use Maximum Likelihood Estimation to estimate the probabilities of the parameters required for this formula in order to deter- mine whether Ingrid can plan tennis. **Use the given variable names and values to indicate which parameter you are calculating.**

ii. Using these estimates, find the predicted class of a new datapoint with *Weather*=sunny, *Wind*=weak. **Spell out the formula for this calculation before you plug-in numbers.**

iii. Are your results reasonable? Why or why not? And if not, how can we improve it from estimation aspect **(No marks will be given for an absent or incorrect explanation)**

Question 5 **[1 + 2 = 3 marks]**

The six dots (8,9), (6,8), (7,7), (8,4), (3,3), (6,2) in the following plot are data points to be clustered using the K-means algorithm. The two Xs (6,6) and (6.9) indicate initial cluster centroids.

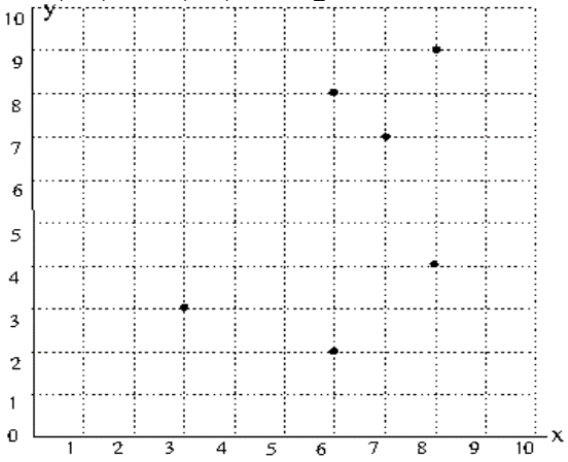


(a) Show the first step of the K-means algorithm by assigning each data point to a cluster centroid, X(6,6) or X(6,9). Use the coordinates of data points.

X(6,6):

X(6,9):

(b) Show the next step of the K-means algorithm by indicating the approximate locations (coordinates) of the two new cluster centroids Xs using the plot below. Assign each data point to a new cluster centroid, X(?,?) or X(?,?), using its coordinates as (a).



X(?,?):

X(?,?):

Your answer should be typed and not handwritten.

Question 6 **[3 + 3 + 3 + 3 = 12 marks]**

(a) In machine learning, a dataset used to build a model is often spilt into a training set, a validation set and a test set. Briefly **explain** the purpose of having a training set, a validation set and a test set for building a neural network (NN) model.

(b) How can you examine if a trained NN model is underfitting? Briefly **explain two** ways to avoid or overcome the underfitting problem?

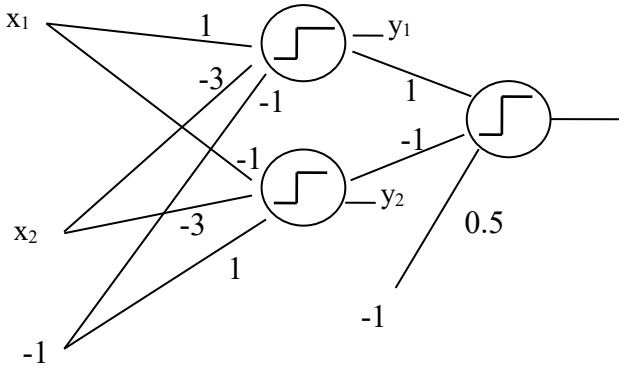
(c) How can you examine if a trained NN model is overfitting? Briefly **explain two** ways to avoid or overcome the overfitting problem?

(d) Apart from the training issues and practical issues in relation to hyperparameter setting and parameter training for building a NN model, what **other practical issues** can be addressed in order to improve the performance of the NN model? Briefly **explain two** such issues.

Your answer should be typed and not handwritten.

Question 7 **[4 + 8 = 12 marks]**

Consider the multi-layered Perceptron below, designed to classify the points A, B, C and D which are *linearly non-separable* in the x_1 - x_2 plane according to the desired outputs in the table below:



Point	x_1	x_2	Desired final output
A	**0**	**0**	**1**
B	**1**	**1**	**0**
C	**1**	**-1**	**0**
D	**2**	**0**	**1**

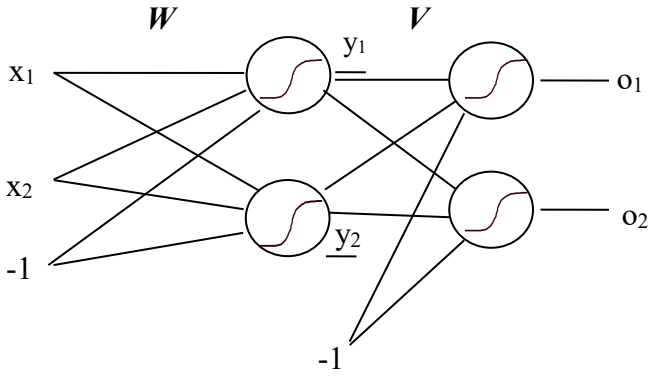
(a) Verify that the weights shown correctly classify the problem by calculating the final outputs for each of the inputs A, B, C, and D (using discrete binary activation functions with threshold = 0)

(b) Show that the 2 hidden neurons perform the task of separating the data into linearly separable regions by plotting the points A, B, C and D in the y_1 - y_2 plane (i.e. show where the points A, B, C and D get translated to when they pass through the hidden neurons).

Question 8 **[2 + 14 = 16 marks]**

(a) In a multilayered feedforward neural network (MFNN), the Delta rule is not applied to modify the hidden layer weights. Why is this?

(b) Consider the MFNN shown below:



The initial values of the weights to the hidden layer and output layer are given (respectively) by:

$$W = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \quad \text{and} \quad V = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 2 & 1 \end{pmatrix}.$$

Assume the learning rate c=1 and that the equation of the activation function is $\frac{1}{1+e^{-net}}$, where *net* is the total input to a neuron.

Perform the steps of the backpropagation learning algorithm when the network is presented with the input (1, -1) which has a known (desired) output of (0.5, 0).
What are the weights (*W* and *V*) after presentation of this input?

You will need the formulae:

$$r_k^o = o_k(d_k - o_k)(1 - o_k)$$

$$r_j^h = y_j(1 - y_j)(\sum_{k=1}^{K} r_k^o v_{kj})$$

$$v_{kj} \leftarrow v_{kj} + r_k^o y_j$$

$$w_{ji} \leftarrow w_{ji} + r_j^h x_i$$

If you do not have a scientific calculator, some of the following values may be useful:

$$e^{-3.26} = 0.04,\ e^{-2} = 0.14,\ e^{-1} = 0.37,\ e^{-0.88} = 0.42,\ e^{-0.5} = 0.61,$$
$$e^{0} = 1.0,\ e^{0.88} = 2.41,\ e^{2} = 7.39,\ e^{3.26} = 26.05$$

Question 9 **[2 + 2 + 3 + 3 = 10 marks]**

Consider the following confusion matrix for classifying 300 good cases and 300 bad cases using a neural network with a decision threshold of 0.5. The case is classified as good if the output of the network is equal to or greater than 0.5 and as bad if the output is less than 0.5.

	Classified good	Classified bad	Row accuracy
Actually good	270	30	(1) ? %
Actually bad	270	30	(2) ? %
Column accuracy	(3) ? %	(4) ? %	

(a) What are the row accuracies and column accuracies for this confusion matrix? Calculate the 4 percentages as indicated by (1) (2) (3) and (4) in the table.

(b) What is the precision and recall of the neural network?

(c) What is your interpretation of the results shown in the confusion matrix above?

(d) What can be done to the decision threshold to improve the accuracy of classifying bad cases?

Your answer should be typed and not handwritten.

Question 10 **[4 + 5 = 9 marks]**

(a) What is the distinctive property of Kohonen's self-organising map (SOM) algorithm, as compared to traditional clustering techniques such as K-means? Briefly explain.

(b) Suppose a self-organising map consists of a linear array of 5 neurons [N_1, N_2, N_3, N_4, N_5]. When a certain input pattern is presented, the winning neuron is found to be N_4. How much learning does each neuron in the array receive, if the neighbourhood size is 2, α=0.5, σ^2=2 for all time *t,* and the equation for the amount of learning received by neuron *i* (given winning neuron *m*) is:

$$c(N_i(t)) = \alpha(t)\, exp(-||r_i - r_m||/\sigma^2(t))$$

If you do not have a scientific calculator, some of the following values may be useful: *exp*(0) = 1, *exp*(-0.5) = 0.61, *exp*(-1) = 0.37, *exp*(-1.5) = 0.22, *exp*(-2) = 0.14, *exp*(-2.5) = 0.08, *exp* (-3) = 0.05

Your answer should be typed and not handwritten.