COMP 3220
Final Exam
Summer 2021

Instructions

The paper consists of 3 sections. Answer all questions from all sections. Upload answers in a pdf document.

SECTION 1

AI & Task Environments

20 mks

1. Describe the Turing test and its purpose. **[2 mks]**

2. Describe the difference between strong and weak AI. **[2 mks]**

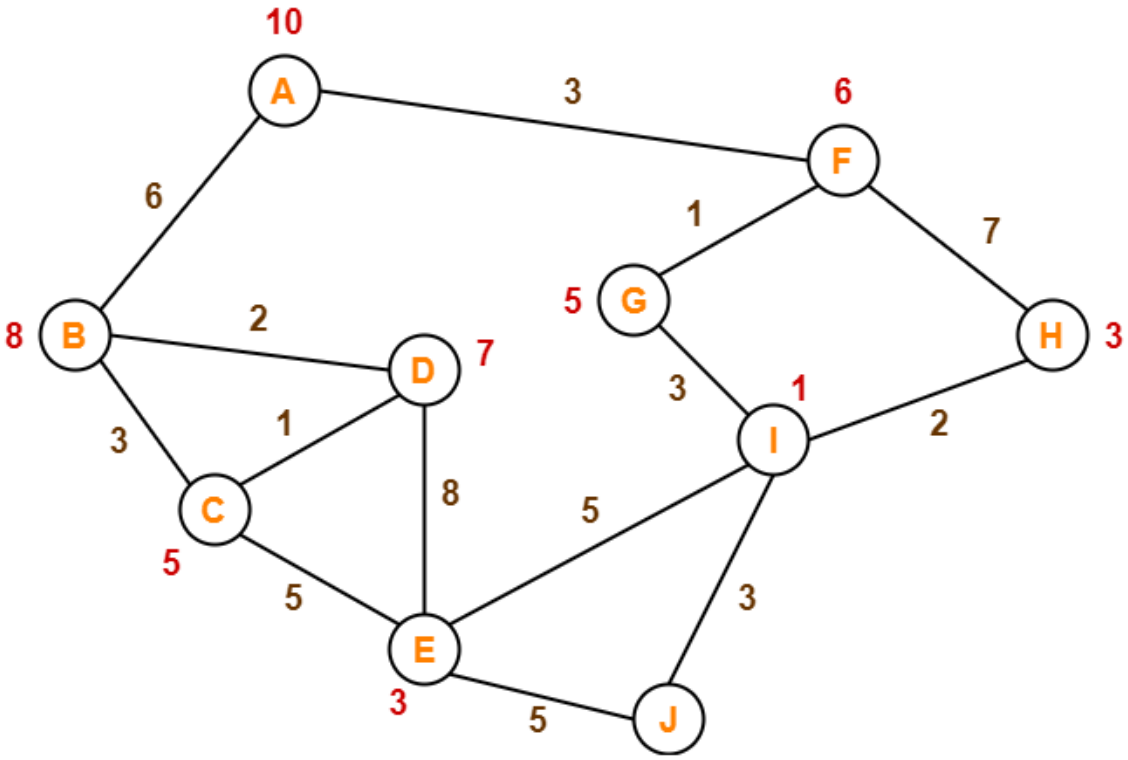
3. Describe the properties of the following environments (deterministic vs stochastic, sequential vs episodic etc.)
    a. Chess **[4 mks]**
    b. Dominoes **[4 mks]**

4. Develop a PEAS(Performance, Environment, Actuator, Sensors) of the following task environments.
    a. Vacuum Cleaning Robot **[4 mks]**
    b. Self Driving Car **[4 mks]**

SECTION 2

Search & Game Playing
20 mks

1. The numbers in red represent the heuristic values to J. For this question you are trying to find a path from **A** to **J**.



a. Develop the search tree for the graph above (must show cumulative path costs). to show all paths from A to J **[3 mks]**

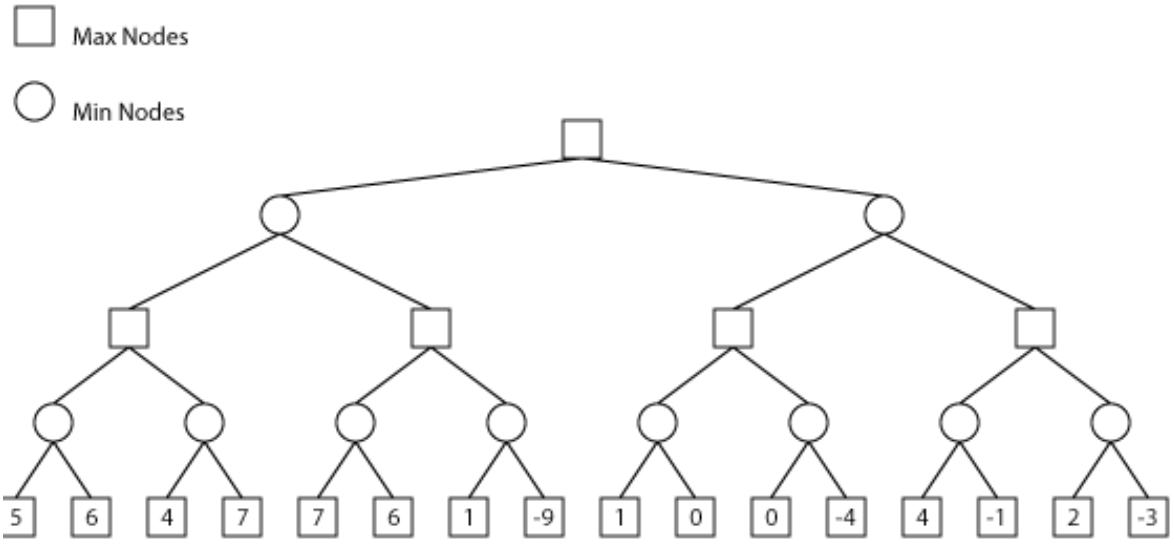
b. State the path and the nodes expanded(**show state of fringe on each step**) for the following algorithms for **A** to **J**.
   i. Greedy **[3mks]**
   ii. A-Star **[5 mks]**

c. Which of the two algorithms above is guaranteed to give you the optimal path. **[1 mk]**

d. What is the main difference between A-Star and UCS and why is A-Star better? **[2 mk]**

2.

a. Use minmax to find out which move max would make (**Left, Right**). **[2 mks]**
b. Use alpha beta pruning to cut branches(**show working**) **[4 mks]**



Section 3
Machine Learning
20mks

1. What is the difference between supervised learning and unsupervised learning and state one example of each. **[2 mk]**

2.

X	Y
95	85
85	95
80	70
70	65
60	70

$$w = \frac{\sum(x - \hat{x})(y - \hat{y})}{\sum(x - \hat{x})^2}$$

$$y = wx + b$$

$$R^2 = 1 - \frac{\sum(y_i - \hat{y})^2}{\sum(y_i - \bar{y})^2}$$

Where,

$\hat{y}$ – *predicted value of* y

$\bar{y}$ – *mean value of* y

a. Using the **OLS(Ordinary Least Square)** Regression algorithm and the data above , calculate the values of **w** and **b**. **Show full working. [6 mk]**

b. Calculate the R2 value and describe what this value represents. **[2 mk]**

3. Given the normalized dataset below and w = **0.2 and** b = **0.1,** our objective is to develop a model that can predict house price based on house size.

$$y_{pred} = wX + b$$

$$\frac{\partial L}{\partial w} = (y_{pred} - y)X$$

$$\frac{\partial L}{\partial b} = (y_{pred} - y)$$

$$L = \frac{1}{2}(y_{pred} - y)^2$$

$$w_{new} = w_{old} - \alpha\frac{\partial L}{\partial w}$$

$$b_{new} = b_{old} - \alpha\frac{\partial L}{\partial b}$$

House Size(X)	House Price(y)
0.22	0.22
0.24	0.58
0.33	0.2
0.37	0.55
0.44	0.39
0.44	0.54

Using a learning rate(α) of 0.5 and the MSE(L) loss function as your loss function. After **one step** of stochastic gradient descent (Show all working), answer the questions below. ***Note you can use any row as the sample***

i. Calculate the loss **[1 mk]**
ii. Calculate new values for w and b **[3 mk]**

b. How many parameter updates would be done for the dataset above in the following scenarios
   i. Stochastic gradient descent and 200 epochs. **[1 mk]**
   ii. Mini batch stochastic gradient descent batch size of 2 and 300 epochs. **[1 mk]**

4. A data scientist creates a neural net with two hidden layers each with 500 neurons. He uses a dataset of size 50 to train a neural net. He then evaluates the neural net. He notices that he gets an accuracy 99% on the training set but on the test set the accuracy went down to 67%.
   a. What is the phenomenon that the data scientist might be experiencing? **[1 mk]**
   b. What can be done to overcome this problem? **[2 mk]**
   c. What activation function could be used on the output layer if the output value must be between 0 and 1. **[1 mk]**

END OF QUESTION PAPER