QUESTION 1

Compute the *Edit Distance* (also known as *Levenshtein Distance*) between the strings “SATU” and “SUNDA” by completing the Dynamic Programming table below. We assume that all edit operations cost one unit.

(10 marks)

		S	A	T	U
S					
U					
N					
D					
A					

QUESTION 2

Let $h_1(s)$ be an admissible A* heuristic. Let $h_2(s) = 2\,h_1(s)$

(a) Is the solution found by A* tree search with h_2 guaranteed to be an optimal solution? Justify your answer.

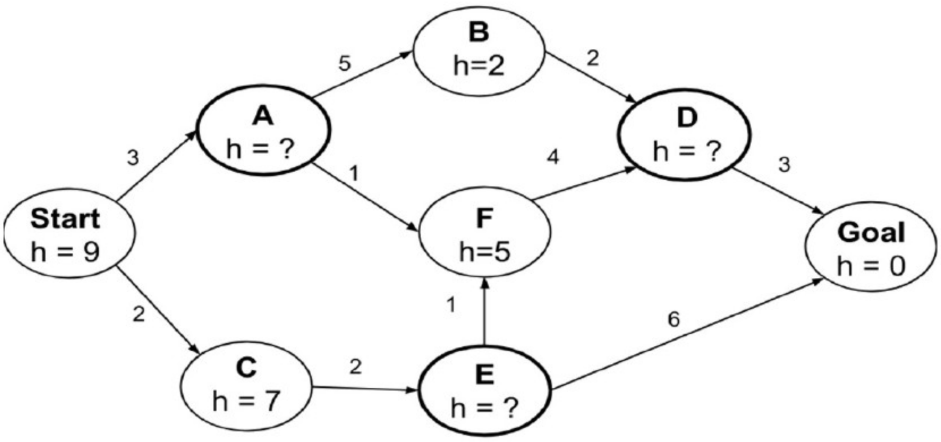
(2 marks)

(b) Is the solution found by A* tree search with h_2 guaranteed to have a cost at most twice as much as the optimal path? Justify your answer.

(2 marks)

(c) Consider the state space graph shown below in which some of the states are missing a heuristic value. Determine the possible range for each missing heuristic value so that the heuristic is admissible and consistent. If this isn't possible, write so.

(6 marks)



State	Range for h(s)
A	$\leq h(A) \leq$
D	$\leq h(D) \leq$
E	$\leq h(E) \leq$

QUESTION 3

Suppose that you have implemented a program for Reinforcement Learning (RL) using the Q-learning algorithm. To check this program, you will compute by hand a test case. Consider a small toy example where the environment can be in four different states ***s*** (1 , 2 , 3 and 4), and the agent in each time step chooses one out of three actions ***a*** (1, 2 and 3). The Q-function is stored as a table in the program, and you initialize it with the values below:

	s = 1	**s = 2**	**s = 3**	**s = 4**
a = 1	0.1	0.2	0.3	0.4
a = 2	0.5	0.6	0.7	0.8
a = 3	0.9	1.0	1.1	1.2

The discount of future rewards γ is set to 0.9. The learning rate λ is 0.2 (that is, the Q-value is adapted by 20% of the temporal difference value).

The system starts in State 3 and there the agent decides to take Action 1. The environment responds by giving a reward of 1 and moves to State 2. There, the agent chooses Action 2 which gives a reward of 0, and moves the agent to State 4.

To check your program, you will compute by hand how the Q-values in the table should change after these two steps.

(a) State which Q-values have changed and by how much after the agent has completed the first step (take Action 1 in State 3 and arrive in State 2 with a reward of 1). Give the relevant formula update(s) for the action values.

(3 marks)

Question 3 continued overleaf

Question 3 continued

(b) State which Q-values have changed and by how much after the agent has completed the second step (take Action 2 in State 2 and arrive in State 4 with a reward of 0). Give the relevant formula update(s) for the action values.

(3 marks)

(c) Explain what happens in reinforcement learning if the agent always chooses the action that maximizes the Q-value.

(2 marks)

(d) Suggest a way to force a RL agent to explore its policy space.

(2 marks)

QUESTION 4

(a) Consider the following cancer testing scenario:

- *1% of women have breast cancer (and therefore 99% do not).*
- *80% of mammograms detect breast cancer when it is there (and therefore 20% miss it).*
- *9.6% of mammograms detect breast cancer when it is not there (and therefore 90.4% correctly return a negative result).*

Now suppose that Alice gets a positive test result. What are the chances Alice has cancer?
Show your work below.

(6 marks)

Question 4 continued overleaf

Question 4 continued

(b) Given the table below,

Ex #	A	B	C	D	Y
1	1	1	1	1	0
2	0	1	1	0	1
3	1	1	0	1	1
4	1	1	0	0	0
5	0	0	1	0	0
6	1	0	1	0	0
7	0	1	0	1	1

Apply Naïve Bayes rule to calculate the ratio $\frac{P(Y=1|A=0,B=1,C=0,D=0)}{P(Y=0|A=0,B=1,C=0,D=0)}$

(6 marks)