MI Final Credit Fall 2023

Part 1 [5 points]

Q1: Write **T** beside true statements and **F** beside false statements.
Note: Every wrong answer will be penalized by -0.5 points.

1. If H(n) is admissible, is H(n)/2 admissible too?
2. Q-Learning is an off policy Reinforcement Learning
3. DFS never finds an optimal solution
4. Sokoban is a fully observed environment
5. Stochastic beam search with K=10 is the same as Random restart hill climbing when you stop after K restarts
6. Minimax algorithm is always optimal, even if the opponent is not acting 100% optimally
7. It is impossible for an environment to exist when a reflex agent is rational
8. Goal based agents have a model of the world
9. Can not remember
10. Can not remember

Part 2 [3 points]

Given the following equation $x + 2y = 5$.

The objective function is to minimize $f(x, y) = |LHS - RHS|$

The actions are:

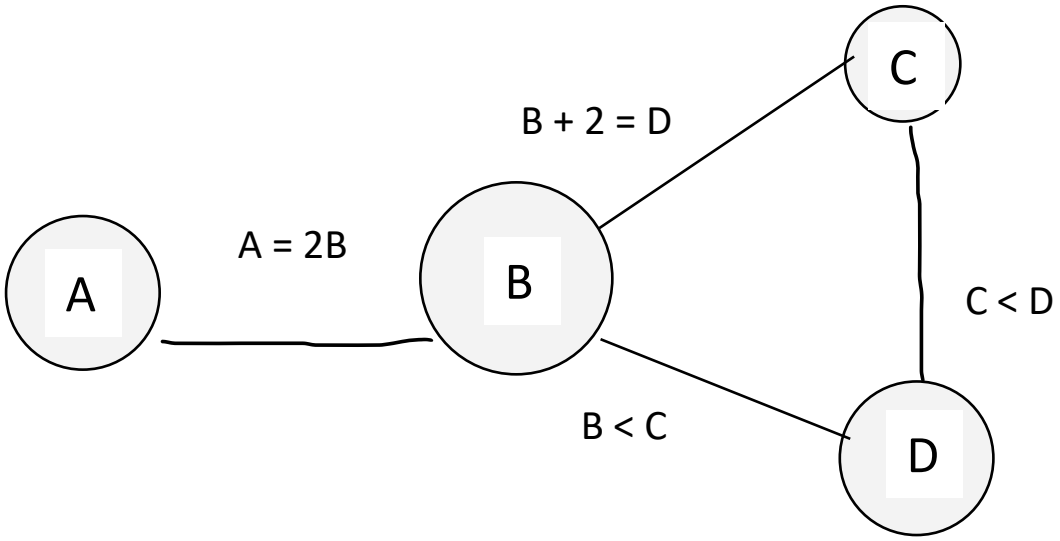
1. Increment X
2. Decrement X
3. Increment Y
4. Decrement Y

Q1. What is the objective function value at (0, 0)? [0.5 points]

Q2. Starting from (0, 0), at which state will it terminate if we used hill climbing? [1.5 points]

Q3. What is the probability to go from (4, 0) to (5, 0) if using simulated annealing with T = 4? [1 points]

Part 3 [5 points]



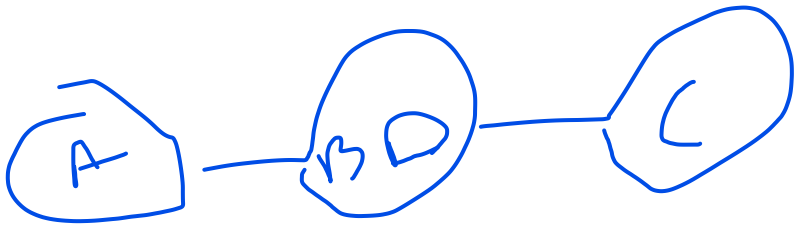
Unary constraint A ≠2

Q1. Apply arc consistency [2 points]

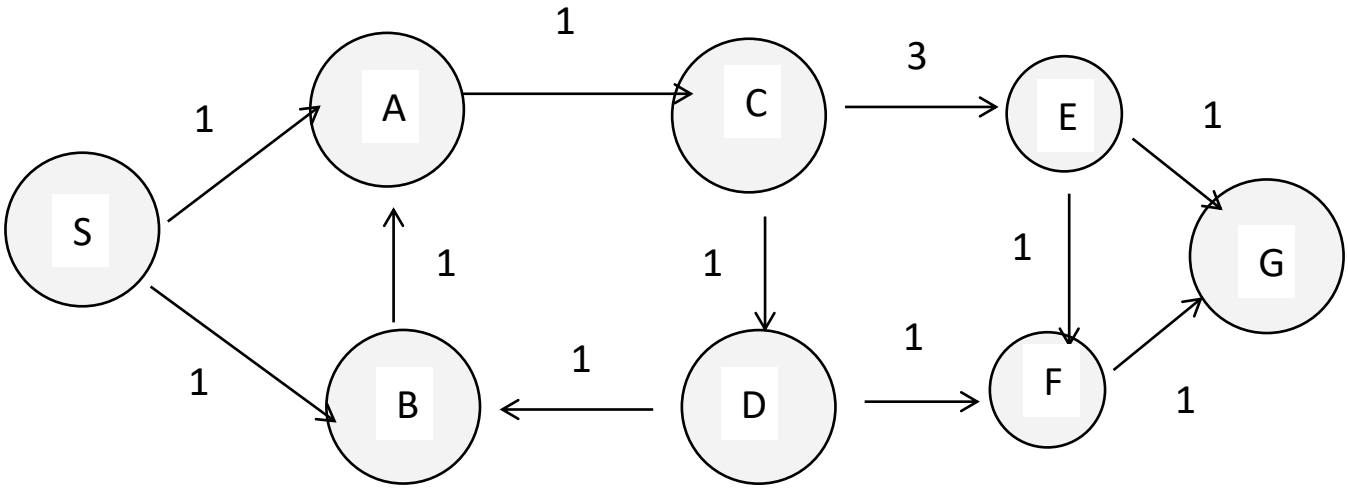
A	~~1~~	2	~~3~~	4	~~5~~
B	1	2	~~3~~	~~4~~	~~5~~
C	~~1~~	2	3	~~4~~	~~5~~
D	~~1~~	~~2~~	3	4	~~5~~

Q2. Is it now node consistent? Justify. [1 point] Yes, because A has {2,4} and 4 satisfing the constraint.

Q3. Convert it to a tree by combining B and D, show the constraint graph after the modification, indicating clearly the constraints and domains. [2 points]



Part 4 [7 points]



N	H(n)
S	5
A	4
B	5
C	3
D	2
E	1
F	1
G	0

Q1. Breadth First Search (Graph Version) [1 point]

SABCDEFG

Q2. Depth First Search (Graph Version) [1 point]

going to the node with smaller alphabetic order

SACDBFG

Q3. Uniform Cost Search (Graph Version) [1 point]

SACDFG

Q4. Greedy Best First Search (Graph Version) [1 point]

focus only on the hurestic

SACEG

Q5. A* Search (Graph Version) [1 point]

both heurestic and actual cost

SACDFG

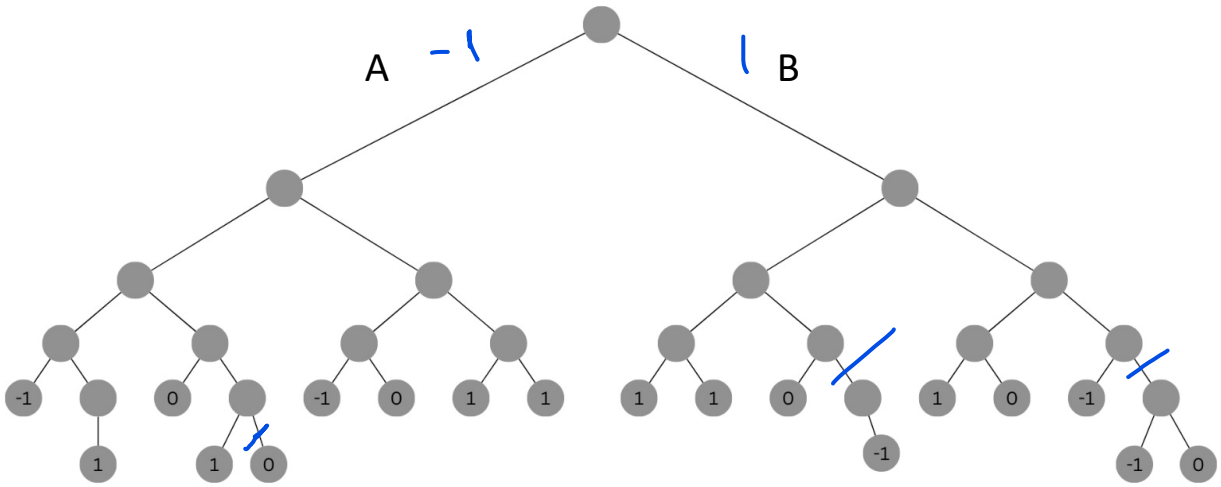
Q6. The given heuristic is admissible, can **another** heuristic function dominate the one given in the table? [2 point]

No

Because the given heuristic already estimates the optimal solution for each node, so there is no better heuristic.

Part 5 [5 points]

This is a zero sum game. You are playing against one opponent. You go first. This is the game tree.

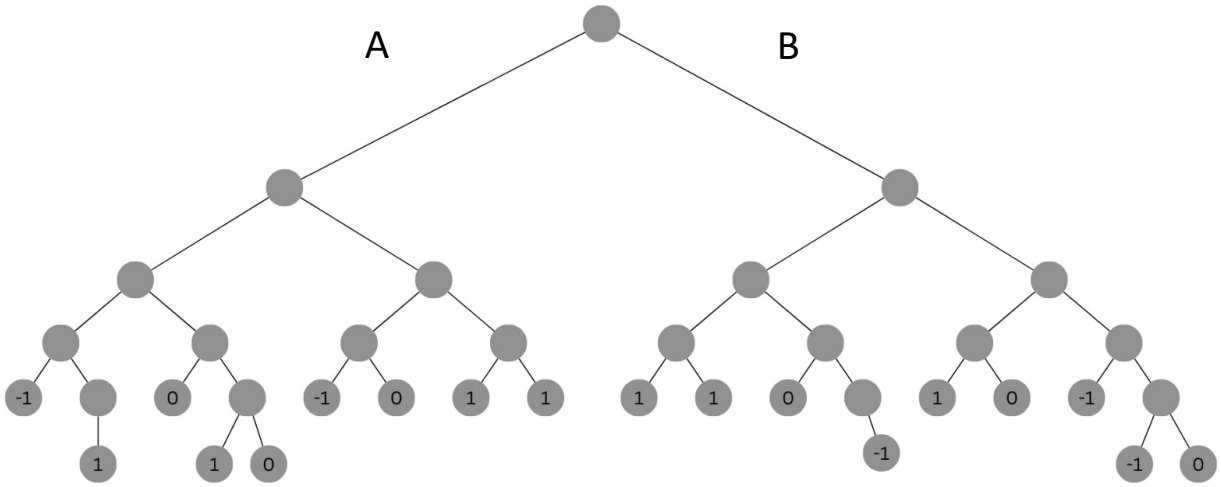


Q1. Apply alpha-beta pruning, and clearly show the pruned nodes. [1 points]

Q2. What is the optimal action to take from the root node and justify your answer (select **all** the correct answers) [2 points]

☐ A ☑ B

This is the same question, but now your opponent does not know how to play the game. He has 75% chance of picking the optimal choice.



Q1. What is the optimal action to take from the root node(select **all** the correct answers) [1 point]

☑ A ☑ B

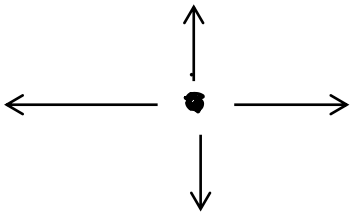
Q2. If you choose the optimal action, what is the expected return value at the root node? [1 point]

7/16

Part 6 [11 points]

MDP

A 1	C 3
B 2	D 4



State D is a terminal state.

UP: 80% chance the agent moves up, 20% chance the agent moves left

LEFT: 80% chance the agent moves left, 20% chance the agent moves down

DOWN: 80% chance the agent moves down, 20% chance the agent moves right

RIGHT: 80% chance the agent moves right, 20% chance the agent moves up

γ = 0.5

assume $U_{1A} = 1, U_{1B} = 2, U_{1C} = 3, U_{1D} = 4$

Q1. What is the value of U_{2A} after 1 value iteration? [1 point] 2.3

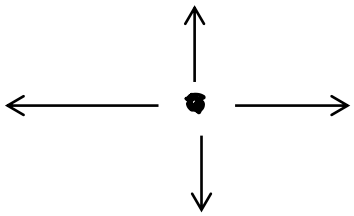
Q2. Equations of policy iteration if π_A = right, π_B= down, π_C = up [1 point]

Q3. Given $U_{1A} = 1, U_{1B} = 2, U_{1C} = 3, U_{1D} = 4.$

Applying policy improvement for state A, what is the action to take? [1 point]

Reinforcement Learning

A 1	C 3
B 2	D 4



State D is a terminal state.

UP: 80% chance the agent moves up, 20% chance the agent moves left

LEFT: 80% chance the agent moves left, 20% chance the agent moves down

DOWN: 80% chance the agent moves down, 20% chance the agent moves right

RIGHT: 80% chance the agent moves right, 20% chance the agent moves up

γ = 0.5

Episodes:

1. A -> B -> A -> C -> D
2. C -> D
3. C -> A -> B -> D

Q1. Apply direct estimated utility of A. [2 points]

Q2. Get transitional models (ADP) [2 points]

	P(A | S, π_O)	P(B | S, π_O)	P(C | S, π_O)	P(D | S, π_O)
S = A	0	2 / 3	1 / 3	0
S = B	1 / 2	0	0	1 / 2
S = C	1 / 3	0	0	2 / 3

Q3. Given table of Qs

	Q(S, UP)	Q(S, LEFT)	Q(S, DOWN)	Q(S, RIGHT)
S = A	1	1	2	1
S = B	1	2	3	2
S = C	3	3	2	1

Given the following Episode

So = A, ao = DOWN

S1 = B, a1 = RIGHT

Q1. Update the Q value for state A, DOWN using Q-Learning, clearly state the equation and show your steps. [2 points]

Q2. Update the Q value for state A, DOWN using SARSA, clearly state the equation and show your steps. [2 points]

Part 7 [4 points]

Ahmed took money to buy books, the probability that he will buy books is 60%. If he buys the books, the probability that he masters the course is 80%. If he does not buy the books, the probability that he masters the course is 40%.

This is an open-book exam.

The probability that he will pass is as follows:

Bought the books and mastered the course: 90%

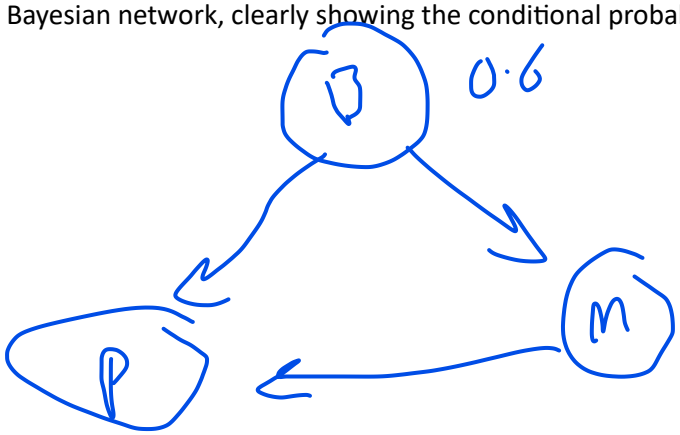
Did not buy the books and mastered the course: 80%

Bought the books and did not master the course: 20%

Did not buy the books and did not master the course: 10%

Let B be variable that he bought the books, M that he mastered the course, P that he passes the exam.

Q1. Draw the Bayesian network, clearly showing the conditional probability tables. [4 points]



B	
T	0.6
F	0.4