CSPB 3202 - Truong - Artificial Intelligence

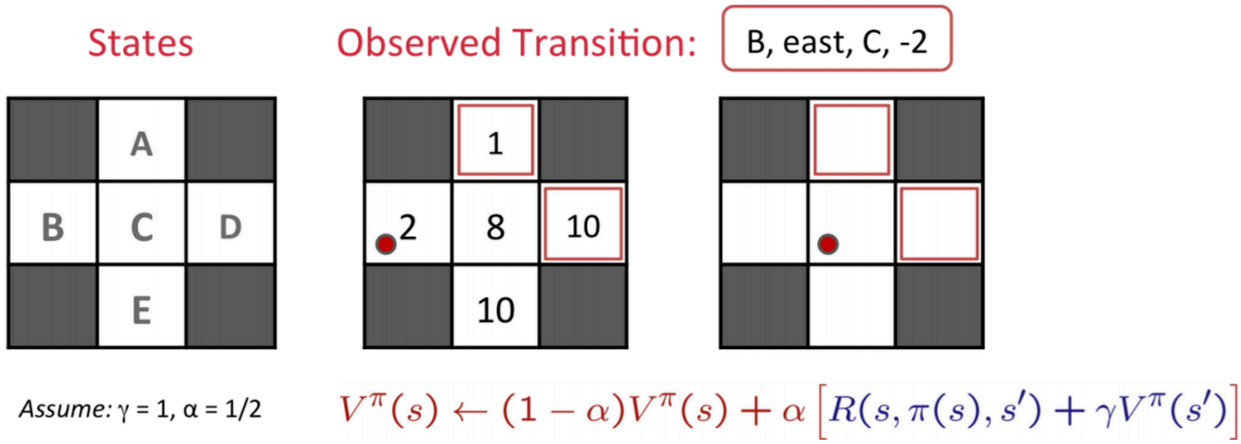
Started on	Friday, 9 August 2024, 8:30 PM
State	Finished
Completed on	Friday, 9 August 2024, 9:38 PM
Time taken	1 hour 7 mins
Marks	73.75/75.00
Grade	**98.33** out of 100.00

Question **1**

Correct

Mark 5.00 out of 5.00

Consider the gridworld shown below. The left panel shows the name of each state A through E. The middle panel shows the current estimate of the value function V^{π} for each state. A transition is observed, that takes the agent from state B through taking action east into state C, and the agent receives a reward of -2. Assuming $\gamma = 1, \alpha = 0.5$, what are the value estimates after the TD learning update? (note: the value will change for one of the states only).



$V^{\pi}(A)$ = 1

$V^{\pi}(B)$ = 4

$V^{\pi}(C)$ = 8

$V^{\pi}(D)$ = 10

$V^{\pi}(E)$ = 10

The only value that gets updated is $V^{\pi}(B)$, because the only transition observed starts in state B.

$V^{\pi}(A)$ = 1

$V^{\pi}(B)$ = .5 $*$ 2 + .5 $*$ (−2 + 8) = 4

$V^{\pi}(C)$ = 8

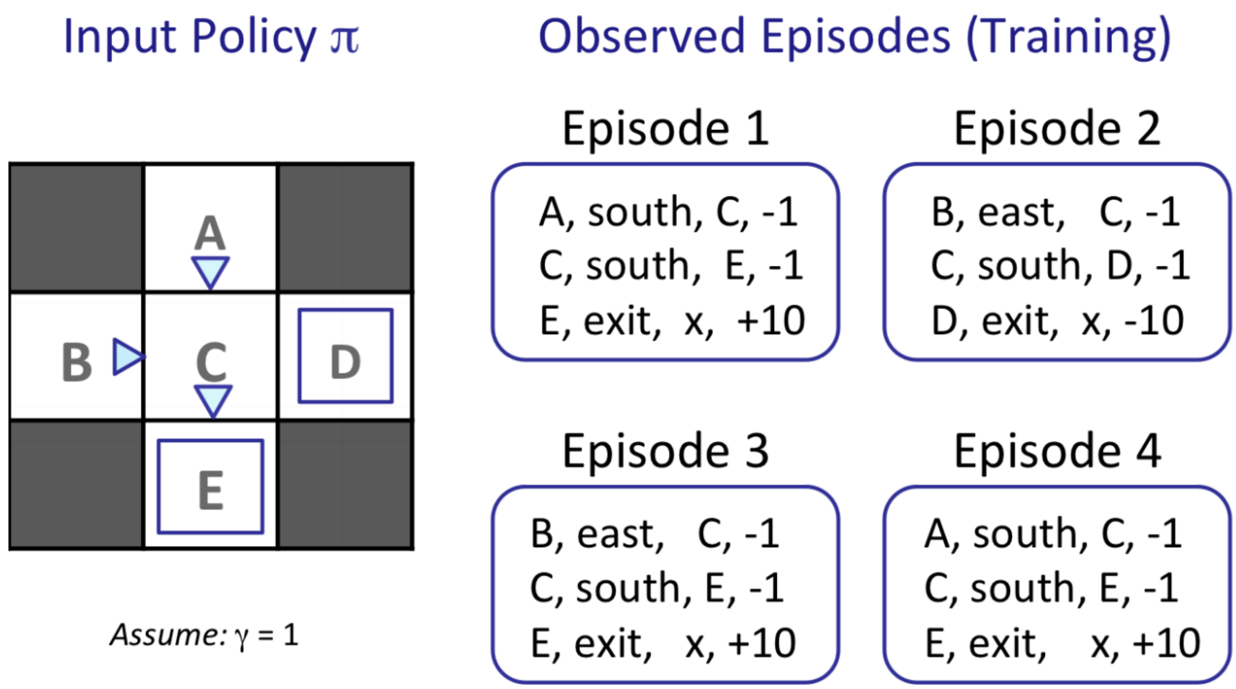
$V^{\pi}(D)$ = 10

$V^{\pi}(E)$ = 10

Question 2

Correct

Mark 5.00 out of 5.00



What are the estimates for the following quantities as obtained by direct evaluation:

$V^{\pi}(A)$ = 8

$V^{\pi}(B)$ = -2

$V^{\pi}(C)$ = 4

$V^{\pi}(D)$ = -10

$V^{\pi}(E)$ = 10

Question **3**

Partially correct

Mark 3.75 out of 5.00

For each of the statements, choose **True** or **False**.

1. In Q-learning, you do not learn the model. True ✔
2. For TD Learning, if I multiply all the rewards in my update by some nonzero scalar p, the algorithm is still guaranteed to find the optimal policy. True ✖
3. In Direct Evaluation, you recalculate state values after each transition you experience. False ✔
4. Q-learning requires that all samples must be from the optimal policy to find optimal q-values. False ✔

Your answer is partially correct.

3 of your answers are correct.

For 1: Q-learning is model-free, you learn the optimal policy explicitly, and the model itself implicitly.

For 2: If p is positive then yes, the discounted values, relative to each other, are just scaled. But if p is negative, you will be computing negating the values for the states, but the policy is still chosen on the max values.

For 3: In order to estimate state values, you calculate state values from episodes of training, not single transitions.

For 4: Q-learning is off-policy, you can still learn the optimal values even if you act sub-optimally sometimes.

Question **4**

Correct

Mark 5.00 out of 5.00

Alpha-beta pruning will prune the same branches after a monotonically increasing function is applied to the value of the leaves.

Select one:

- (•) True ✔
- ( ) False

correct

The correct answer is 'True'.

Question **5**

Correct

Mark 5.00 out of 5.00

What is the maximum number of times a backtracking search algorithm might have to backtrack in a general CSP, if it is running arc consistency and applying the MRV and LCV heuristics?

Select one:

- a. 0
- b. $O(1)$
- c. $O(nd^2)$
- d. $O(n^2 d^3)$
- e. $O(d^n)$ ✔
- f. ∞

Your answer is correct.

\(O(d^n)\). The MRV and LCV heuristics are often helpful to guide the search, but are not guaranteed to reduce backtracking in the worst case.

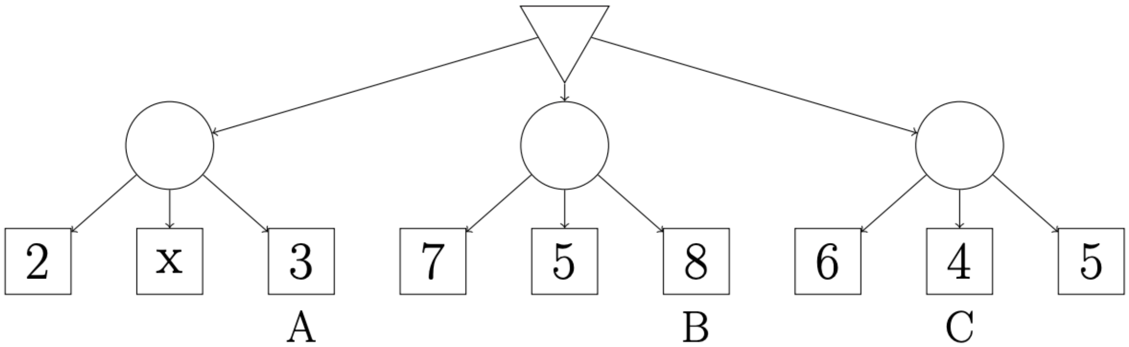
The correct answer is: \(O(d^n)\)

Question 6

Correct

Mark 5.00 out of 5.00

For each of the leaves labeled A, B, and C in the tree below, determine which values of x will cause the leaf to be pruned, given the information that all values are non-negative and all nodes have 3 children. Assume we do not prune on equality.



Check and fill the right answer. For the inequality item, choose "none" if no value of X will cause pruning, or "any" if the node will be pruned for all values of X. If you choose "none" or "any" for the inequality, fill 0 for the numeric box.

For A to be pruned: X needs to be ☐smaller ☐larger ☑none ✔ ☐any than

0

✔

For B to be pruned: X needs to be ☑smaller ✔ ☐larger ☐none ☐any than

7

✔

For C to be pruned: X needs to be ☑smaller ✔ ☐larger ☐none ☐any than

1

✔

Your answer is correct.

Question 7

Correct

Mark 5.00 out of 5.00

What is necessarily true regarding iterative deepening on any search tree?

a. Complete as opposed to DFS tree search

b. Strictly faster than DFS tree search

c. Strictly faster than BFS tree search

d. More memory ecient than BFS tree search

e. None of above

Your answer is correct.

The correct answers are: Complete as opposed to DFS tree search, More memory ecient than BFS tree search

Question 8

Correct

Mark 5.00 out of 5.00

First consider general search problems, which of the following are true?

Select one or more:

a. The max of two admissible heuristics is always admissible.

b. A* tree search is complete with any heuristic function

c. A heuristic that always evaluates to h(s) = 1 for non-goal search nodes s is always admissible.

d. Code that implements A* tree search can be used to run UCS

e. A* tree search is optimal with any heuristic function.

f. A* graph search is guaranteed to expand no more nodes than DFS.

Your answer is correct.

The correct answers are: Code that implements A* tree search can be used to run UCS, The max of two admissible heuristics is always admissible.

Question **9**

Correct

Mark 5.00 out of 5.00

When enforcing arc consistency in a CSP, the set of values which remain when the algorithm terminates does not depend on the order in which arcs are processed from the queue.

Select one:

- (•) True ✔
- ( ) False

The correct answer is 'True'.

Question **10**

Correct

Mark 5.00 out of 5.00

Given two admissible heuristics \(h_A\) and \(h_B\), consider performing A* tree search. Which is generally best to use if we want to expand the fewest number of nodes?

Select one:

- ( ) a. \(h_A+h_B\)
- ( ) b. \(\frac{1}{2}h_A\)
- ( ) c. \(\frac{1}{2}h_B\)
- ( ) d. \(\frac{1}{2}(h_A+h_B)\)
- ( ) e. \(h_A \times h_B\)
- (•) f. max(\(h_A\),\(h_B\)) ✔
- ( ) g. min(\(h_A\),\(h_B\))

Your answer is correct.

The correct answer is: max(\(h_A\),\(h_B\))

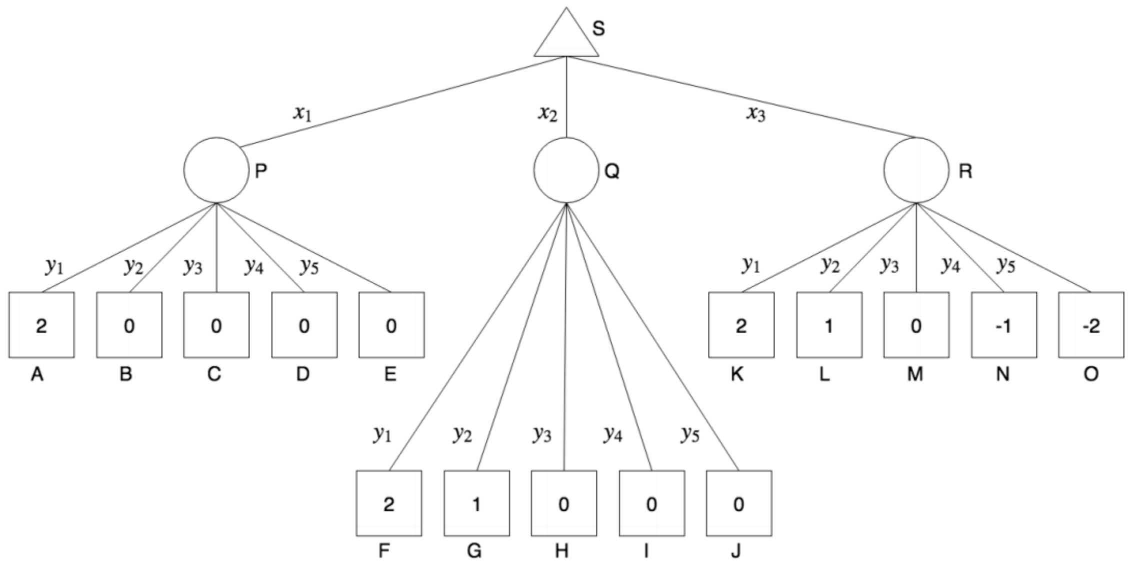
Question **11**

Correct

Mark 5.00 out of 5.00

Alyssa P. Hacker and Ben Bitdiddle are bidding in an auction at Stanley University for a bike. Alyssa will either bid x1, x2, or x3 for the bike. She knows that Ben will bid y1, y2, y3, y4, or y5, but she does not know which. All bids are nonnegative.

Alyssa wants to maximize her payoff given by the expectimax tree below. The leaf nodes show Alyssa's payoff. The nodes are labeled by letters, and the edges are labeled by the bid values xi and yi . The maximization node S represents Alyssa, and the branches below it represent each of her bids: x1, x2, x3. The chance nodes P, Q, R represent Ben, and the branches below them represent each of his bids: y1, y2, y3, y4, y5.



Suppose that Alyssa believes that Ben would bid any bid with equal probability. What are the values of the chance (circle) and maximization (triangle) nodes?

Node R:

Answer: 0

The correct answer is: 0

Question **12**

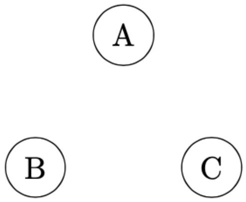
Correct

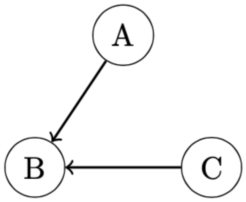
Mark 5.00 out of 5.00

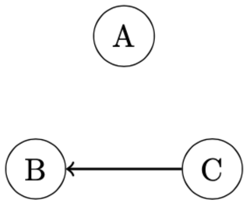
Given the joint probability table below, choose all corresponding Bayes Nets (BNs) options that can correctly represent the distribution on the right. If no such Bayes Nets are given, clearly select *None of the above*.

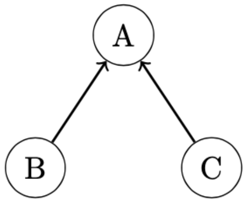
$\mathbf{A}$	$\mathbf{B}$	$\mathbf{C}$	$\mathbf{P(A,B,C)}$
0	0	0	.22
0	0	1	.08
0	1	0	.22
0	1	1	.08
1	0	0	.09
1	0	1	.11
1	1	0	.09
1	1	1	.11

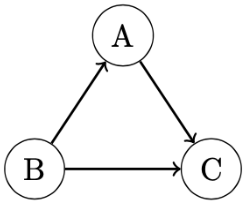
Select one or more:

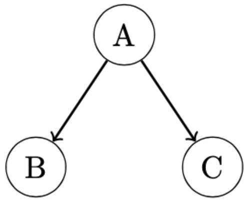






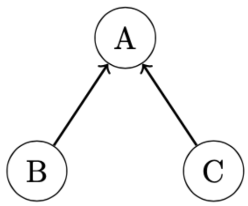
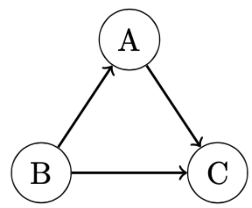
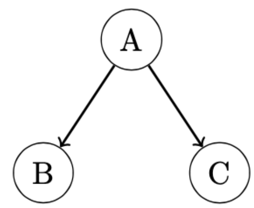




- [x] 
- [ ] None of the above

Your answer is correct.

From the table we can see that the values of \(P(A, B, C)\) repeat in two blocks. This means that the value of \(B\) does not matter to the distribution. The values are otherwise all unique, meaning that there is a relationship between \(A\) and \(C\). Together, this means that \(B⊥A\), \((B⊥A | C)\), \(B⊥C\), and \((B⊥C | A)\) are the only independence relationships in the distribution.

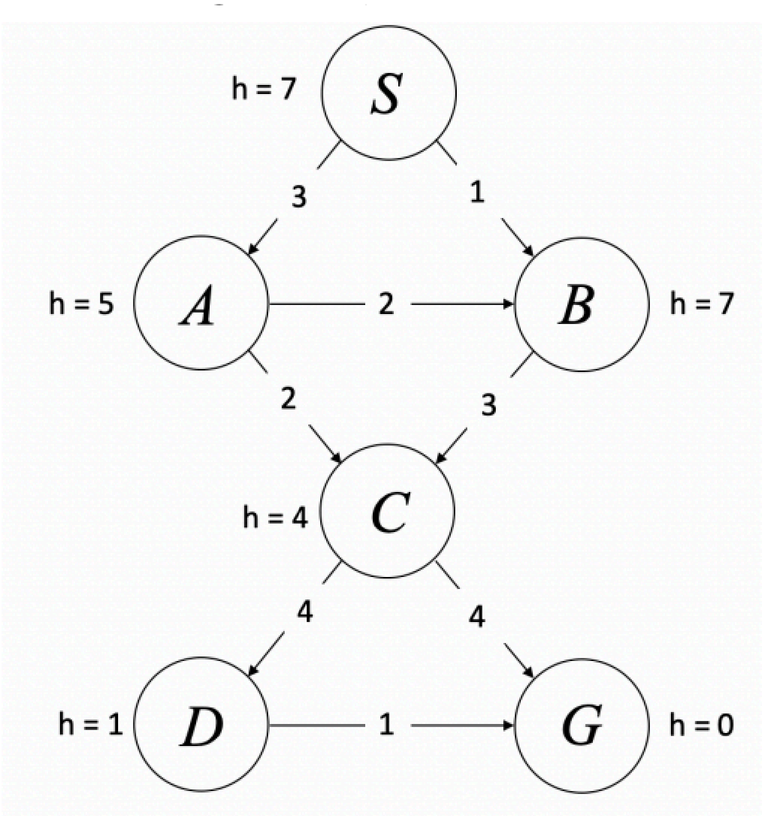
The correct answers are:  ,  , 

Question **13**

Correct

Mark 5.00 out of 5.00

We will investigate various search algorithms for the following graph. Edges are labeled with their costs, and heuristic values h for states are labeled next to the states. S is the start state, and G is the goal state. In all search algorithms, assume ties are broken in alphabetical order.



Assuming we run A* graph search with the heuristic values provided, what path is returned?

Select one:

- S → A → B → C → D → G
- S → A → C → G
- S → A → C → D → G
- S → B → C → G ✔ Correct
- S → A → B → C → G
- None of the above

Your answer is correct.

The correct answer is: S → B → C → G

Question **14**

Correct

Mark 5.00 out of 5.00

In this MDP, available actions are left, right, and stay. Stay always results in the agent staying in its current square. Left and right are successful in moving in the intended direction half of the time. The other half of the time, the agent stays in its current square. An agent cannot try to move left at the leftmost square, and cannot try to move right on the rightmost square. Staying still on a square gives a reward equivalent to the number on that square and all other transitions give zero reward (meaning any transitions in which the agent moves to a different square give zero reward).

4	0	0	36

- The value of $V_0(s)$ is 0 for all s. Perform one step of value iteration with $\gamma = 1/2$ and write $V_1(s)$ in each corresponding square:

4	0	0	36

- Perform another step of value iteration with $\gamma = 1/2$, and write $V_2(s)$ in each corresponding square:

6	1	9	54

- Using $\gamma = \frac{1}{2}$ again, write $V^{*}(s)$ in each corresponding square (Hint: think about states where the optimal action is obvious and work from there):

8	8	24	72

Your answer is correct.

For **Part A**:

$V_1(s)$ is equal to the optimal one-step rewards from each state. This can be achieved from each state by staying, which gives reward equal to the number on each square. (In the middle squares, moving in either direction can do no better than 0 in one step.)

For **Part B**:

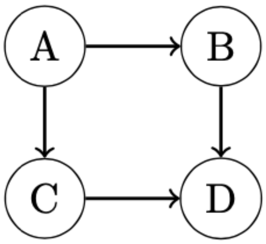
From the two squares on the side, the agent can stay twice, giving $36 + \gamma \times 36 = 36 + 18 = 54$ and $4 + \gamma \times 4 = 4 + 2 = 6$. From the middle squares, the optimal policy for a two-step horizon is to move outward toward the closer side. From the middle-right, it will move right, which in two steps will give 0 reward if the first step is unsuccessful and $0 + \gamma \times 36 = 18$ if the first step is successful, giving an expectation of $\frac{1}{2} \times 0 + \frac{1}{2} \times 18=9$. By an identical argument, V_2 on the middle-left state is $\frac{1}{4} \times 4 = 1$.

Question **15**

Correct

Mark 5.00 out of 5.00

Consider the Bayes’ net shown below:



From the given choices below, choose an expression for \(P(A | B, C, D)\)

Select one:

- \(P(A)P(B | A)P(C | A)\)
- \(\frac{P(A)P(B | A)P(C | A)}{\Sigma_{a}(P(B | a)P(C | a)P(a))}\) ✔
- \(\frac{P(B | A)P(A)}{P(C | A)}\)
- \(\frac{P(A)}{\Sigma_{a}(P(B | a)P(C | a)P(D | a))}\)

Your answer is correct.

\(P(A | B, C, D) = \frac{P(A, B, C, D)}{P(B, C, D)} = \frac{P(D | B, C)P(B | A)P(C | A)P(A)}{P(D | B, C)P(B, C)}\) = \(\frac{P(A)P(B | A)P(C | A)}{\Sigma_{a}(P(B | a)P(C | a)P(a))}\)

The correct answer is: \(\frac{P(A)P(B | A)P(C | A)}{\Sigma_{a}(P(B | a)P(C | a)P(a))}\)