RULES

- **No books**, **no notes**, and **no calculators**.
- **No bathroom breaks** until after you have completed and submitted the exam.
- **All phones must be completely silent** for the duration of the exam, so please *turn off your phone now!*
- Out of consideration for your classmates, do not make disturbing noises during the exam.

Cheating will not be tolerated. If there is any indication that a student may have given or received unauthorized aid on this test, the case will be referred to the Office of the Dean of Students. When you finish the exam, you must sign the following pledge:

"On my honor as a student I, ______________________________, have neither given nor received unauthorized aid on this exam." (print name clearly)

Signature: ______________________________________ Date: March 26, 2022

Question:	1	2	3	4	5	6	7	8	9	10	Total
Points:	10	8	10	4	5	14	4	10	6	4	75
Score:											

Part 1. Search

1. (10 pts) We call a search algorithm a *graph search* if it checks for redundant paths and a *tree-like search* if it does not.

   Best-First Search employs a general *priority queue* data structure to store the frontier, and we can use this algorithm to implement various search strategies—e.g., Breadth-First Search, Depth-First Search, Uniform-Cost Search, etc.—by specifying the type of queue used to store the frontier (or "fringe") nodes of the search. Recall, the three basic types of queues: (1) *first-in-first-out* (FIFO) *queue* whose pop method returns the node that was added to the queue first; (2) *last-in-first-out* (LIFO) *queue* (a *stack*), where pop returns the node that was most recently added; (3) *priority queue*, where pop returns the node n with $\mathrm{f}(n) = \min_m \mathrm{f}(m)$, for some evaluation function f.

   Select the answer that correctly completes the following sentences.

   (a) Breadth-First Search uses
   - ☐ a priority queue for some evaluation function, f.
   - ☐ a FIFO queue. ☐ a LIFO queue. ☐ None of these.

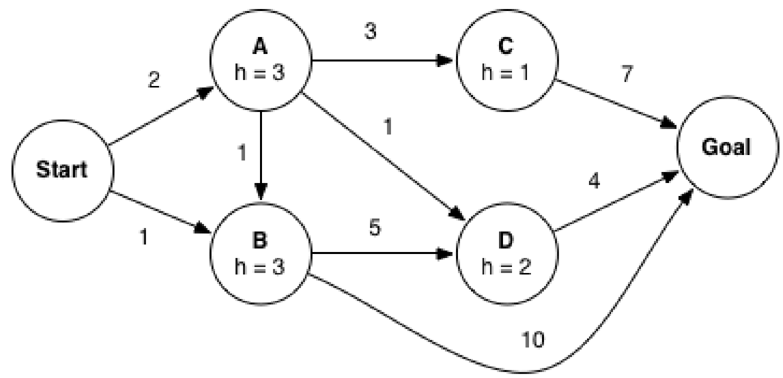
   (b) Depth-First Search uses
   - ☐ a priority queue for some evaluation function, f.
   - ☐ a FIFO queue. ☐ a LIFO queue. ☐ None of these.

   (c) Uniform-Cost Search uses
   - ☐ a priority queue for some evaluation function, f.
   - ☐ a FIFO queue ☐ a LIFO queue ☐ None of these.

   (d) Suppose the search space is infinite. Which of the following search strategies is complete? (Select all that apply.)
   - ☐ Breadth-First Search
   - ☐ Depth-First Search
   - ☐ Iterative Deepening Search
   - ☐ None of the above.

   (e) In general, Iterative Deepening Search is the preferred uninformed search method when the search state space is larger than can fit in memory and the depth of the solution is not known. ☐ True ☐ False

2. (8 pts) Consider the graph below. Arcs are labeled with their weights. Values of an heuristic function are shown, though you may or may not need these values to answer some of the question below. Assume that ties are broken alphabetically (so a partial plan S→X→A would be expanded before S→X→B and S→A→Z would be expanded before S→B→A). You may find it helpful to execute your searches on the blank pages provided at the end of the exam.



(a) In what order are states expanded by Uniform Cost Search?

☐ Start, A, B, C, D, Goal ☐ Start, A, C, Goal ☐ Start, B, A, D, C, Goal

☐ Start, A, D, Goal ☐ Start, A, B, Goal ☐ Start, B, A, D, B, C, Goal

(b) What path does uniform cost search return?

☐ Start-A-C-Goal ☐ Start-B-Goal ☐ Start-A-D-Goal

☐ Start-A-B-Goal ☐ Start-A-B-D-Goal

(c) In what order are states expanded by A* graph search?

☐ Start, A, B, C, D, Goal ☐ Start, A, C, Goal ☐ Start, B, A, D, C, Goal

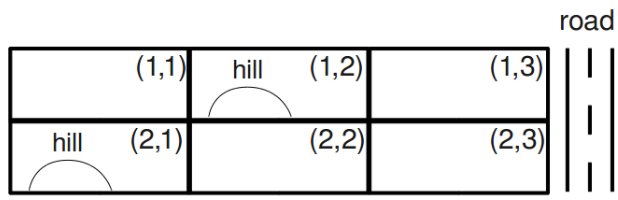
☐ Start, A, D, Goal ☐ Start, A, B, Goal ☐ Start, B, A, D, B, C, Goal

(d) What path does A* graph search return?

☐ Start-A-C-Goal ☐ Start-B-Goal ☐ Start-A-D-Goal

☐ Start-A-B-Goal ☐ Start-A-B-D-Goal

3. (10 pts) (Campus Layout) You are asked to determine the layout of a new, small college. The campus will have four structures: an administration structure (A), a bus stop (B), a classroom (C), and a dormitory (D). Each structure (including the bus stop) must be placed somewhere on the grid shown below.



The layout must satisfy the following constraints:

(i) The bus stop (B) must be adjacent to the road.

(ii) The administration structure (A) and the classroom (C) must both be adjacent to the bus stop (B).

(iii) The classroom (C) must be adjacent to the dormitory (D).

(iv) The administration structure (A) must not be adjacent to the dormitory (D).

(v) The administration structure (A) must not be on a hill.

(vi) The dormitory (D) must be on a hill or adjacent to the road.

(vii) All structures must be in different grid squares.

Here, adjacent means that the structures must share a grid edge, not just a corner. We recommend you work out the solutions to the following questions on a sheet of scratch paper, and then enter your results below.

(a) Which of the constraints above are unary constraints?

☐ (i) ☐ (ii) ☐ (iii) ☐ (iv) ☐ (v) ☐ (vi) ☐ (vii) ☐ None of these

(b) Select the domains of all variables after unary constraints have been applied.

A: ☐ (1,1) ☐ (1,2) ☐ (1,3) ☐ (2,1) ☐ (2,2) ☐ (2,3)
B: ☐ (1,1) ☐ (1,2) ☐ (1,3) ☐ (2,1) ☐ (2,2) ☐ (2,3)
C: ☐ (1,1) ☐ (1,2) ☐ (1,3) ☐ (2,1) ☐ (2,2) ☐ (2,3)
D: ☐ (1,1) ☐ (1,2) ☐ (1,3) ☐ (2,1) ☐ (2,2) ☐ (2,3)

(c) Let's start from the table above (the answer to Part b) and enforce arc consistency. Initally, the queue contains all arcs (in alphabetical order). Let's examine what happens when enforcing A $\rightarrow$ B. After enforcing unary constraints, the domains of A and B are:

A	B
(1,1)	
(1,3)	(1,3)
(2,2)	
(2,3)	(2,3)

Which of the following contains the correct domains after enforcing A $\rightarrow$ B? Pay attention to which variable's domain changes and which side of the arc it's on.

A	B
(1,1)	
	(1,2)
(1,3)	(1,3)
(2,2)	
(2,3)	(2,3)

i

A	B
(1,3)	(1,3)
(2,2)	
(2,3)	(2,3)

ii

A	B
(1,1)	
(1,3)	
(2,2)	
(2,3)	(2,3)

iii

A	B
(1,1)	
(1,3)	(1,3)
(2,2)	
(2,3)	

iv

☐ i ☐ ii ☐ iii ☐ iv

(d) Starting from the answer to Part b (in which unary constraints are enforced), select the domains of all variables after A $\rightarrow$ B is enforced.

A: ☐ (1,1) ☐ (1,2) ☐ (1,3) ☐ (2,1) ☐ (2,2) ☐ (2,3)
B: ☐ (1,1) ☐ (1,2) ☐ (1,3) ☐ (2,1) ☐ (2,2) ☐ (2,3)
C: ☐ (1,1) ☐ (1,2) ☐ (1,3) ☐ (2,1) ☐ (2,2) ☐ (2,3)
D: ☐ (1,1) ☐ (1,2) ☐ (1,3) ☐ (2,1) ☐ (2,2) ☐ (2,3)

(e) After enforcing these arcs, the next is C $\rightarrow$ B. Continuing from the previous parts, select the domains of all variables after C $\rightarrow$ B is enforced.

A: ☐ (1,1) ☐ (1,2) ☐ (1,3) ☐ (2,1) ☐ (2,2) ☐ (2,3)
B: ☐ (1,1) ☐ (1,2) ☐ (1,3) ☐ (2,1) ☐ (2,2) ☐ (2,3)
C: ☐ (1,1) ☐ (1,2) ☐ (1,3) ☐ (2,1) ☐ (2,2) ☐ (2,3)
D: ☐ (1,1) ☐ (1,2) ☐ (1,3) ☐ (2,1) ☐ (2,2) ☐ (2,3)

4. (4 pts) Assume you are given a CSP and you enforce arc consistency. Which of the following are true?

- ☐ If the CSP has no solution, it is guaranteed that enforcement of arc consistency resulted in at least one domain being empty.
- ☐ If the CSP has a solution, then after enforcing arc consistency, you can directly read off the solution from resulting domains.
- ☐ In general, to determine whether the CSP has a solution, enforcing arc consistency alone is not sufficient; backtracking may be required.

5. (5 pts)

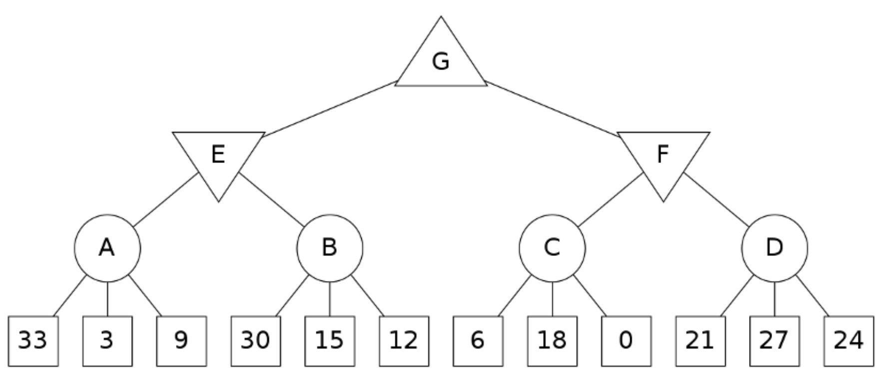
   (a) Assuming a Markov decision process has a finite number of actions and states, and that the discount factor satisfies $0 < \gamma < 1$, which of the following are true about *value iteration*?

   - ☐ Value iteration is guaranteed to converge.
   - ☐ Value iteration will converge to the same vector of values (V^*) no matter what values we use to initialize V.
   - ☐ None of the above

   (b) Which of the following statements are true for a Markov Decision Process?

   - ☐ If the only difference between two MDPs is the value of the discount factor then they must have the same optimal policy.
   - ☐ For an infinite horizon MDP with a finite number of states and actions and with a discount factor γ that satisfies $0 < \gamma < 1$, value iteration is guaranteed to converge.
   - ☐ When running value iteration, if the policy (the greedy policy with respect to the values) has converged, the values must have converged as well.
   - ☐ If one is using value iteration and the values have converged, the policy must have converged as well.
   - ☐ Expectimax will generally run in the same amount of time as value iteration on a given MDP.
   - ☐ For an infinite horizon MDP with a finite number of states and actions and with a discount factor γ that satisfies $0 < \gamma < 1$, policy iteration is guaranteed to converge.

6. (14 pts) Consider the game tree shown below. Triangles that point up, such as the top node (root), represent choices for the maximizing player; triangles that point down represent choices for the minimizing player. The circular nodes represent chance nodes in which each of the possible actions may be taken with equal probability. The square nodes at the bottom represent leaf nodes. Assuming both players act optimally, carry out the expectiminimax search algorithm. Enter the values for the letter nodes in the boxes below the tree.



A: ____________ B: ____________ C: ____________ D: ____________

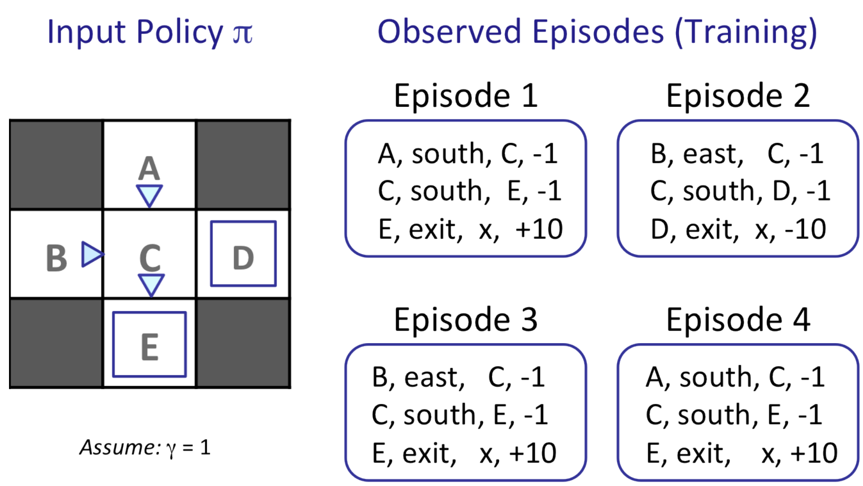
E: ____________ F: ____________ G: ____________

7. (4 pts) Q-LEARNING PROPERTIES

In general, for Q-Learning to converge to the optimal Q-values...

- ☐ It is necessary that every state-action pair is visited infinitely often.
- ☐ It is necessary that the learning rate α (weight given to new samples) is decreased to 0 over time.
- ☐ It is necessary that the discount γ is less than 1/2.
- ☐ It is necessary that actions get chosen according to $\text{argmax}_a\ Q(s, a)$.

8. (10 pts) DIRECT EVALUATION



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

(Write you answers on the lines provided.)

(a) $\hat{V}^\pi(\text{A}) =$

(a) ____________

(b) $\hat{V}^\pi(\text{B}) =$

(b) ____________

(c) $\hat{V}^\pi(\text{C}) =$

(c) ____________

(d) $\hat{V}^\pi(\text{D}) =$

(d) ____________

(e) $\hat{V}^\pi(\text{E}) =$

(e) ____________

9. (6 pts) Select all expressions that are equivalent to the specified probability using the given independence assumptions.

(a) Given no independence assumptions, $P(A, B \mid C) =$

☐ $\frac{P(C|A)P(A|B)P(B)}{P(C)}$ ☐ $\frac{P(B,C|A)P(A)}{P(B,C)}$ ☐ $P(A \mid B, C)P(B \mid C)$ ☐ $\frac{P(A|C)P(B,C)}{P(C)}$

(b) Given that A is independent of B given C, $P(A, B \mid C) =$

☐ $\frac{P(C|A)P(A|B)P(B)}{P(C)}$ ☐ $\frac{P(B,C|A)P(A)}{P(B,C)}$ ☐ $P(A \mid B, C)P(B \mid C)$ ☐ $\frac{P(A|C)P(B,C)}{P(C)}$

(c) Given that A is independent of B given C, $P(A \mid B, C) =$

☐ $\frac{P(C|A)P(A|B)P(B)}{P(C)}$ ☐ $\frac{P(B,C|A)P(A)}{P(B,C)}$ ☐ $\frac{P(A|C)P(C|B)P(B)}{P(B,C)}$ ☐ $\frac{P(C|A,B)P(B|A)P(A)}{P(B|C)P(C)}$

10. (4 pts) Given are the prior distribution $P(X)$ and two conditional distributions $P(Y \mid X)$ and $P(Z \mid Y)$ shown below. Also, assume Z is independent of X given Y. All variables are binary (0-1 variables). Compute the following joint distributions based on the chain rule.

X	$P(X)$
0	0.500
1	0.500

Y	X	$P(Y\|X)$
0	0	0.600
1	0	0.400
0	1	0.900
1	1	0.100

Z	Y	$P(Z\|Y)$
0	0	0.100
1	0	0.900
0	1	0.700
1	1	0.300

(a) $P(X = 1, Y = 1)$

(a) ____________

(b) $P(X = 1, Y = 1, Z = 1)$

(b) ____________