MIDTERM

UNIVERSITY OF SOUTH CAROLINA

COMPUTER SCIENCE AND ENGINEERING

CSCE 580: Artificial Intelligence

Midterm

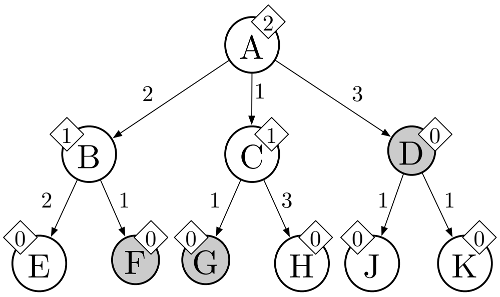
Author:
Your Name (ID: Your USC ID)

Release Date: March 31, 2019
Date&Time: Monday March 19, 4:15 pm – 5:40 pm

THIS PAGE IS INTENTIONALLY LEFT BLANK

Q1. Search: Algorithms

Consider the state space search problem shown to the right. A is the start state and the shaded states are goals. Arrows encode possible state transitions, and numbers by the arrows represent action costs. Note that state transitions are directed; for example, $A \to B$ is a valid transition, but $B \to A$ is not. Numbers shown in diamonds are heuristic values that estimate the optimal (minimal) cost from that node to a goal.



For each of the following search algorithms, write down the nodes that are removed from fringe in the course of the search, as well as the final path returned. Because the original problem graph is a tree, the tree and graph versions of these algorithms will do the same thing, and you can use either version of the algorithms to compute your answer.

Assume that the data structure implementations and successor state orderings are all such that **ties are broken alphabetically**. For example, a partial plan $S \to X \to A$ would be expanded before $S \to X \to B$; similarly, $S \to A \to Z$ would be expanded before $S \to B \to A$.

1. **Depth-First Search (ignores costs) [5 pts]**

   Nodes removed from fringe: A, B, E, F
   Path returned: A, B, F

2. **Breadth-First Search (ignores costs) [5 pts]**

   Nodes removed from fringe: A, B, C, D
   Path returned: A, D

3. **Uniform-Cost Search [5 pts]**

   Nodes removed from fringe: A, C, B, G
   Path returned: A, C, G

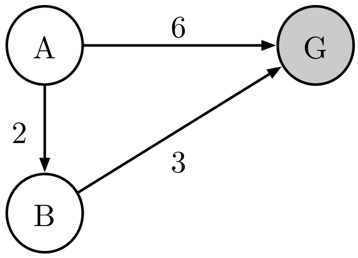
4. **Greedy Search [5 pts]**

   Nodes removed from fringe: A, D
   Path returned: A, D

5. **A* Search [5 pts]**
   Nodes removed from fringe: A, C, G
   Path returned: A, C, G

Q2. Search: Heuristic Function Properties

For the following questions, consider the search problem shown on the left. It has only three states, and three directed edges. A is the start node and G is the goal node. To the right, four different heuristic functions are defined, numbered I through IV.



	h(A)	h(B)	h(G)
I	4	1	0
II	5	4	0
III	4	3	0
IV	5	2	0

Admissibility and Consistency For each heuristic function, circle whether it is admissible and whether it is consistent with respect to the search problem given above.

Table 1

	Admissible?		Consistent?	
I [5 pts]	Yes	No	Yes	No
II [5 pts]	Yes	No	Yes	No
III [5 pts]	Yes	No	Yes	No
IV [5 pts]	Yes	No	Yes	No

Q3. Games

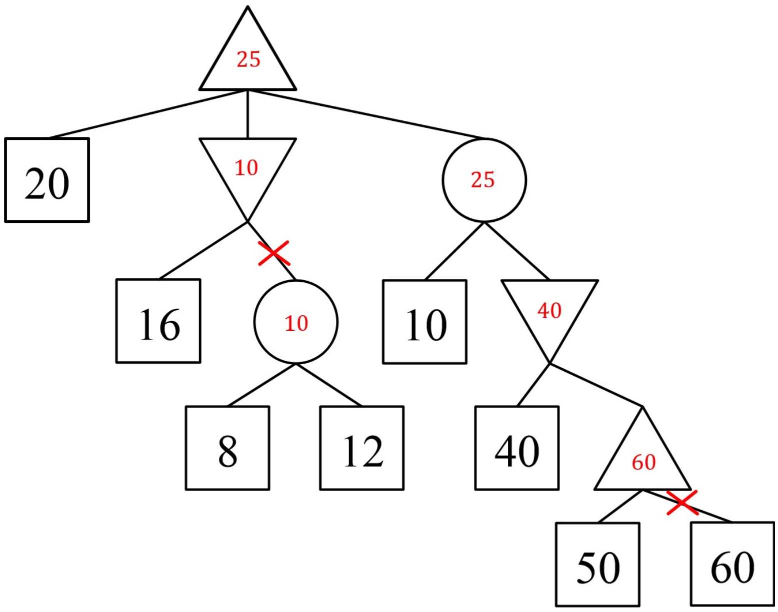
1. **Games.** Consider the game tree below, which contains maximizer nodes, minimizer nodes, and chance nodes. For the chance nodes the probability of each outcome is equally likely.

   (a) Fill in the values of each of the nodes. [10 pts]

   (b) Is pruning possible? [10 pts]

      i. No. Brief justification:

      ii. Yes. Cross out the branches that can be pruned.



2. **Utilities.** Pacman's utility function is $U(\$x) = \sqrt{x}$. He is faced with the following lottery: $[0.5, \$36\ ;\ 0.5, \$64]$. What is Pacman's expected utility? [10 pts]

   The utility of a lottery can be computed via $EU(L) = \sum_{x \in L} p(x)U(x)$
   Hence, $EU([0.5, \$36\ ;\ 0.5,\ \$64]) = 0.5U(\$36) + 0.5U(\$64) = 0.5(6) + 0.5(8) = 7$

Q4. MDPs: Mini-Grids

The following problems take place in various scenarios of the gridworld MDP. In all cases, A is the start state and double-rectangle states are exit states. From an exit state, the only action available is **Exit**, which results in the listed reward and ends the game (by moving into a terminal state X, not shown).
From non-exit states, the agent can choose either **Left** or **Right** actions, which move the agent in the corresponding direction. There are no living rewards; the only non-zero rewards come from exiting the grid.
Throughout this problem, assume that value iteration begins with initial values $V_0(s) = 0$ for all states s.
First, consider the following mini-grid. For now, the discount is $\gamma = 1$ and legal movement actions will always succeed (and so the state transition function is deterministic).

+1	A			+10

1. What is the optimal value $V^*(A)$? [4 pts] 10

   Since the discount $\gamma = 1$ and there are no rewards for any action other than exiting, a policy that simply heads to the right exit state and exits will accrue reward 10. This is the optimal policy, since the only alternative reward if 1, and so the optimal value function has value 10.

2. When running value iteration, remember that we start with $V_0(s) = 0$ for all s. What is the first iteration k for which $V_k(A)$ will be non-zero? [4 pts] 2
   The first reward is accrued when the agent does the following actions (state transitons) in sequence: Left, Exit. Since two state transitions are necessary before any possible reward, two iterations are necessary for the value function to become non-zero.

3. What will $V_k(A)$ be when it is first non-zero? [4 pts] 1
   As explained above, the first non-zero value function value will come from exiting out of the left exit cell, which accrues reward 1.

4. After how many iterations k will we have $V_k(A) = V^*(A)$? If they will never become equal, write **never**. [4 pts] 4
   The value function will equal the optimal value function when it discovers this sequence of state transitions: Right, Right, Right, Exit. This will obviously happen in 4 iterations.

Now the situation is as before, but the discount γ is less than 1.

5. If $\gamma = 0.5$, what is the optimal value $V^*(A)$? [4 pts]

   The optimal policy from A is Right, Right, Right, Exit. The rewards accrued by these state transitions are: 0, 0, 0, 10. The discount values are $\gamma^0, \gamma^1, \gamma^2, \gamma^3$, which is 1, $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$. Therefore, $V^*(A) = 0 + 0 + 0 + \frac{10}{8}$.

6. For what range of values γ of the discount will it be optimal to go **Right** from A? [5 pts]

   The best reward accrued with the policy of going left is $\gamma^1 * 1$. The best reward accrued with the policy of going right is $\gamma^3 * 10$. We therefore have the inequality $10\gamma^3 \geq \gamma$, which simplifies to $\gamma \geq \sqrt{1/10}$. The final answer is $1/\sqrt{10} \leq \gamma \leq 1$