CS 747 (Autumn 2025) Week 12 Test (Batch 1)

5.35 p.m. – 6.00 p.m., November 4, 2025, LA 001

Name: ______________________________ Roll number: _______________

Note. There is one question in this test, which has three parts. You can use the space on all pages for your answer. Draw a line (either vertical or horizontal) and do all your rough work on one side of it.

Question 1. This question is about Tic-Tac-Toe, the popular game played on a 3×3 grid. To begin the game, the first player places an "X" in any of the cells; then the second player places an "O" in any other cell. The players alternate placing X's and O's on unoccupied cells until termination. The game terminates either when (1) one of the players has an entire row, column, or diagonal filled with their symbol—in which case that player is the winner, or (2) all 9 cells are filled up such that every row, column, and diagonal contains both players' symbols—in which case the outcome is a draw. For each player, the reward from a win is 1, from a draw is 0, and from a loss is -1.

The figure on the right shows a state s that has been reached after six total moves have been played, with three X's and three O's placed. Clearly it is the first player's turn to make a move from s.

O	X	O
O		X
X		

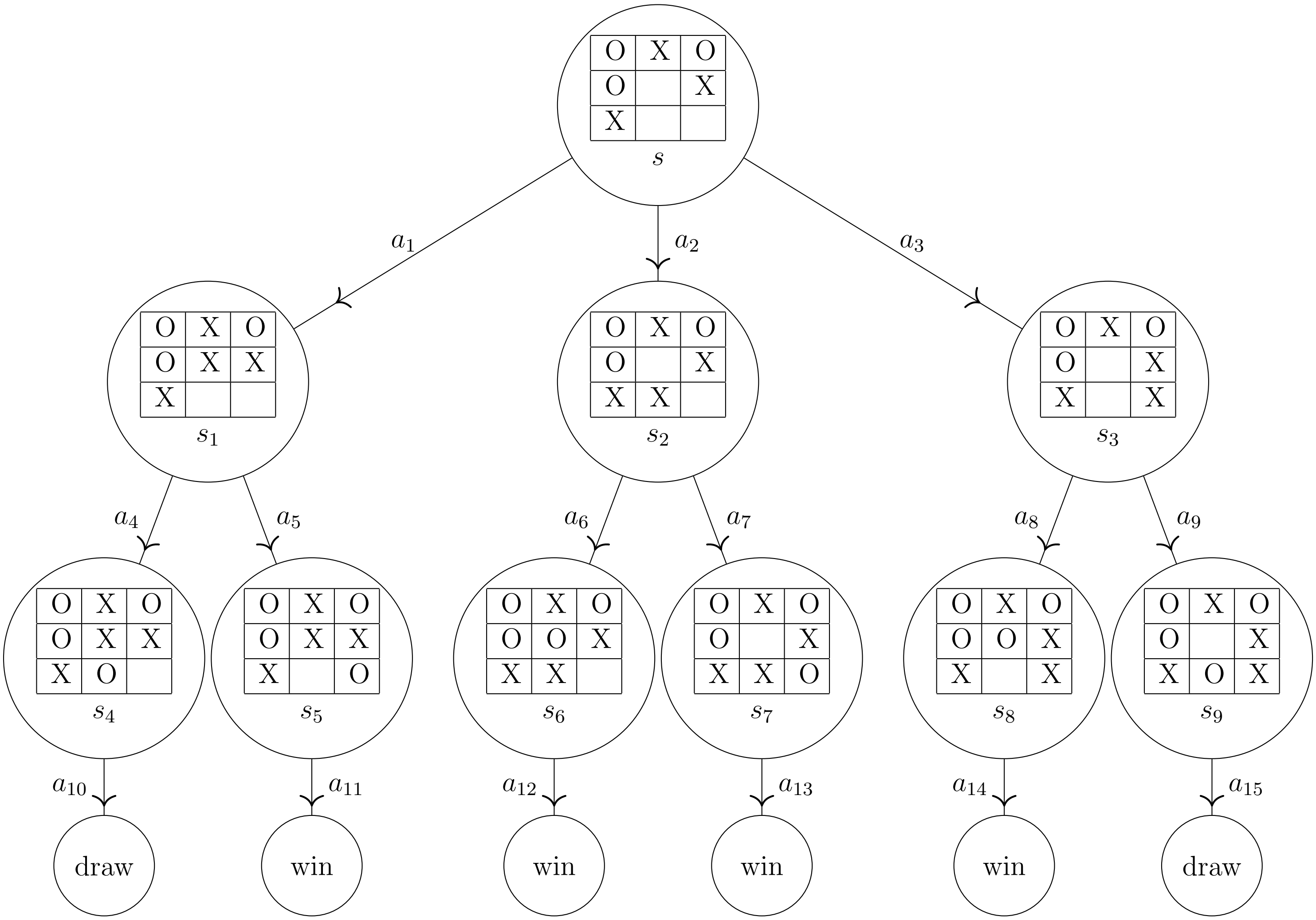
State s

1a. Draw the *game tree* for Tic-Tac-Toe with s as the root: that is, to describe all possible trajectories when starting from s. [1 mark]

1b. Suppose that on the trajectory starting from s, the first player always selects its move uniformly at random from those available, while the second player always plays a move that maximises its expected reward (with arbitrary tiebreaking). What is the expected reward of the first player if starting from s? [1 mark]

1c. Suppose that on the trajectory starting from s, each player always selects a move that maximises its expected reward (with arbitrary tiebreaking). In this case, what is the expected reward of the first player if starting from s? [1 mark]

Answer 1a. The game tree is shown below, with "win", "loss", and "draw" indicating the outcome for the first player upon terminating. States and actions are numbered to support explanations in 1b and 1c. Note that actions from s (level 1) and s_4–s_9 (level 3) are taken by the first player, while actions from s_1–s_3 (level 2) are taken by the second player. The only non-zero rewards are upon reaching the terminal states.



Answer 1b. We use standard MDP notation, leaving implicit the policy being followed. V and Q are defined from the first player's perspective.

$$\begin{aligned}
V(s) &= \frac{1}{3}Q(s,a_1) + \frac{1}{3}Q(s,a_2) + \frac{1}{3}Q(s,a_3) \\
&= \frac{1}{3}\min\{V(s_4),V(s_5)\} + \frac{1}{3}\min\{V(s_6),V(s_7)\} + \frac{1}{3}\min\{V(s_8),V(s)\} \\
&= \frac{1}{3}\times 0 + \frac{1}{3}\times 1 + \frac{1}{3}\times 0 \\
&= \frac{1}{3}.
\end{aligned}$$

Answer 1c. In this case,

$$\begin{aligned}
V(s) &= \max\{Q(s,a_1),Q(s,a_2),Q(s,a_3)\} \\
&= \max\{\min\{V(s_4),V(s_5)\}, \min\{V(s_6),V(s_7)\} \min\{V(s_8),V(s)\}\} \\
&= \max\{0,1,0\} \\
&= 1.
\end{aligned}$$

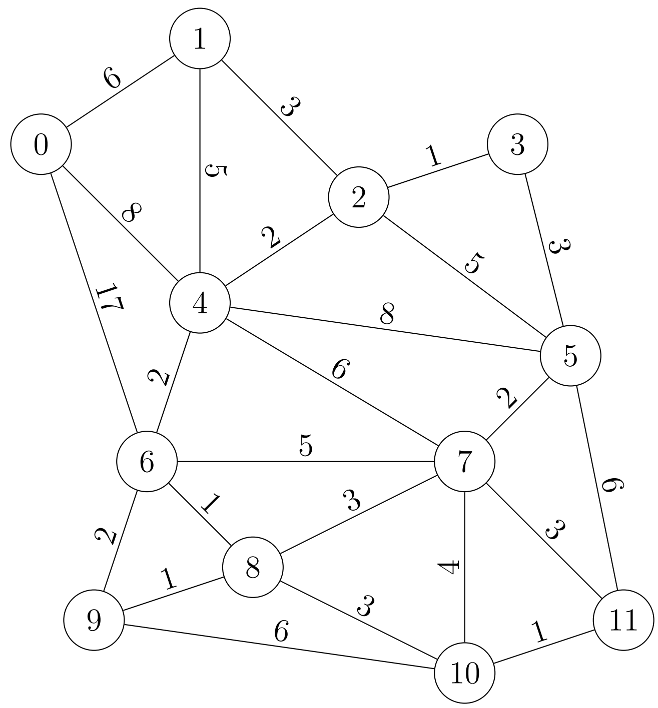
CS 747 (Autumn 2025) Week 12 Test (Batch 2)

6.15 p.m. – 6.40 p.m., November 4, 2025, LA 001

Name: ______________________________ Roll number: _______________

Note. There is one question in this test, which has three parts. You can use the space on all pages for your answer. Draw a line (either vertical or horizontal) and do all your rough work on one side of it.

Question 1. The figure on the right represents a road network in a city, with each vertex a place and each edge a road. The edge weight shown corresponds to a positive "cost" to travel the road. Roads are bidirectional and the cost is the same in each direction. An agent that is at place 0 needs to travel to place 11. It uses A* search to plan its path, with the aim of identifying a route that minimises the total cost. Recall that a node in the search is a triple (state, path from start state, path cost). Our agent's states are places. The root node in its search is $n_0 = (0, 0, 0)$.



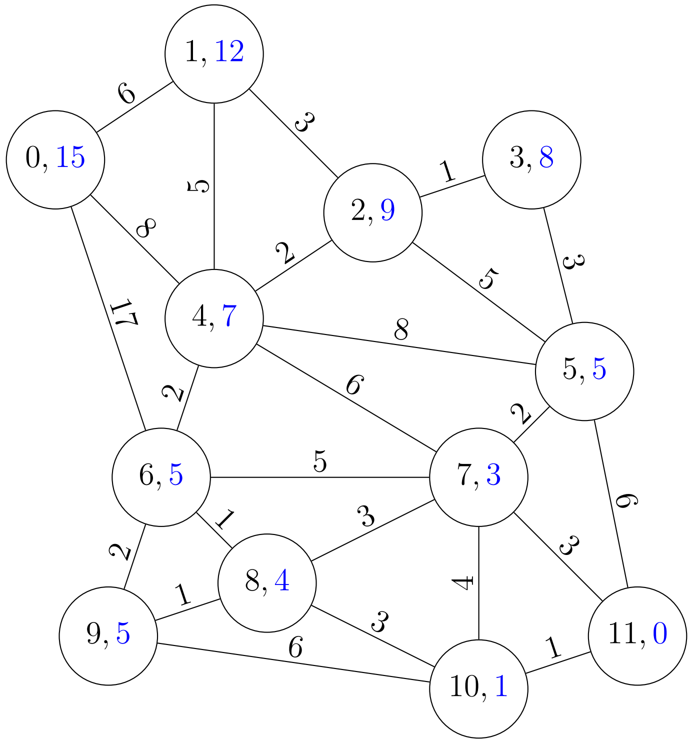
1a. When the agent performs its first expansion (of the root), it puts three nodes on the frontier: $n_1 = (1, 0 \to 1, 6)$, $n_2 = (4, 0 \to 4, 8)$, and $n_3 = (6, 0 \to 6, 17)$. If the heuristic function h used by the agent is admissible, what it the set of all possible values that $h(n_2)$ can take? [1 mark]

1b. Describe an admissible heuristic h such that the A* search procedure expands all three nodes n_1, n_2, and n_3 before termination; or explain why such a heuristic does not exist. [1 mark]

1c. Suppose that the A* search procedure is carried out using a heuristic h which takes value 1 for n_2, and is 0 for all other nodes encountered during the procedure. What is total cost along the path from the start state to the goal state that is found by this procedure? [1 mark]

For your convenience, the figure is reproduced on each subsequent page. You are free to annotate and include these reproductions in your final answers. Clearly distinguish between rough and fair work.

1. The first step to attempt any of the questions is to compute the optimal cost-to-goal $c^\star(s)$ for each state s. The cost-to-goal for a node is the cost-to-goal of the state in the node. This computation is easily done by starting at the target 11 and proceedings backwards, to states whose all actions lead to states with known cost-to-goal. The optimal cost-to-goal values are shown in blue within each vertex.



1a. Since h is admissble, it must satisfy $h(s_1) \geq 0$ and $h(s_1) \leq c^\star(s_1)$. Therefore, the set of admissible heuristic functions h satisfy $h(s) \in [0, 7]$.

1b. The cost-to-goal of state 0 is 15, which is realised along the path $0 \to 4 \to 6 \to 8 \to 10 \to 11$. Consider the node $n_4 = (11, 0 \to 4 \to 6 \to 8 \to 10 \to 11, 15)$. If h is admissible, we have $f(n_4) = g(n_4) + h(n_4) = 15 + 0 = 15$, whereas $f(n_2) = g(n_2) + h(n_2) = 17 + h(n_2) \geq 17$. Since A* expands the frontier node n that minimises $g(n) + h(n)$, it will expand n_4 before it expands n_2. Since the search will terminate when it discovers the state of n_4 to be a goal state, the search will never expand n_2. In short, there is no admissible heuristic h under which n_2 will be expanded.

1c. The heuristic specified is admissible, so we are assured that it will find an optimal path to the goal state. The only optimal path is $0 \to 4 \to 6 \to 8 \to 10 \to 11$, and its total cost is 15.