CS221 Problem Workout Solutions

Week 5

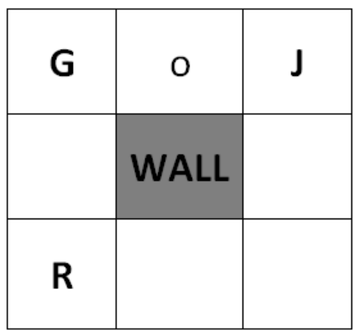
1) [CA session] Problem 1

After finally meeting up, Romeo (R) and Juliet (J) decide to try to catch a goose (G) to keep as a pet. Eventually, they chase it into a 3×3 hedge maze show below. Now they play the following turn-based game:

(a) The Goose moves either Down or Right.

(b) Romeo moves either Up or Right.

(c) Juliet moves either Left or Down.



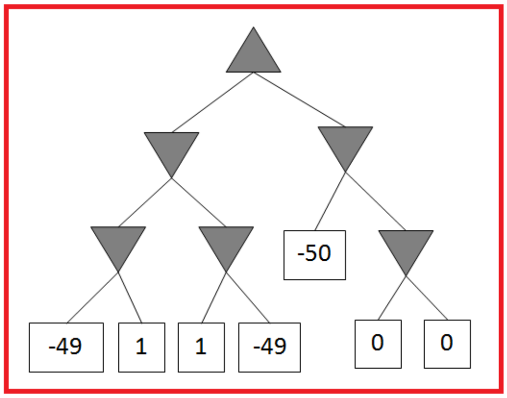
Participants: Goose (G), Romeo (R), Juliet (J), bread (o)

If the Goose enters the square with bread, it gets a reward 1. If either Romeo or Juliet enters the same square as the Goose, they catch it and the Goose gets a reward of -50. The game ends when either the Goose has been caught or everyone has moved once. Note that it is possible for the Goose to get both rewards.

Construct a depth one minimax tree for the above situation, with the Goose as the maximizer and Juliet and Romeo as the minimizers. Use up-triangles Δ for max nodes, down-triangles ∇ for min nodes, and square nodes for the leaves. Label each node with its minimax value.

What is the minimax value of the game if Romeo defects and becomes a maximizer?

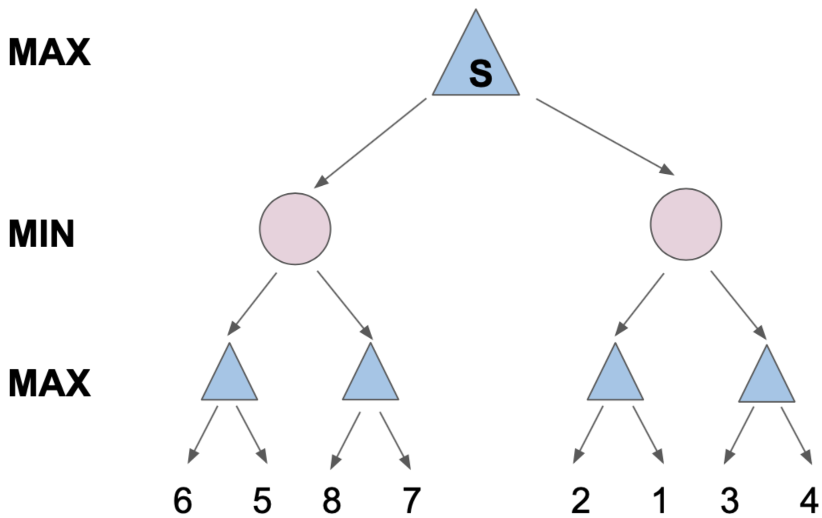
Solution Here is the minimax tree:



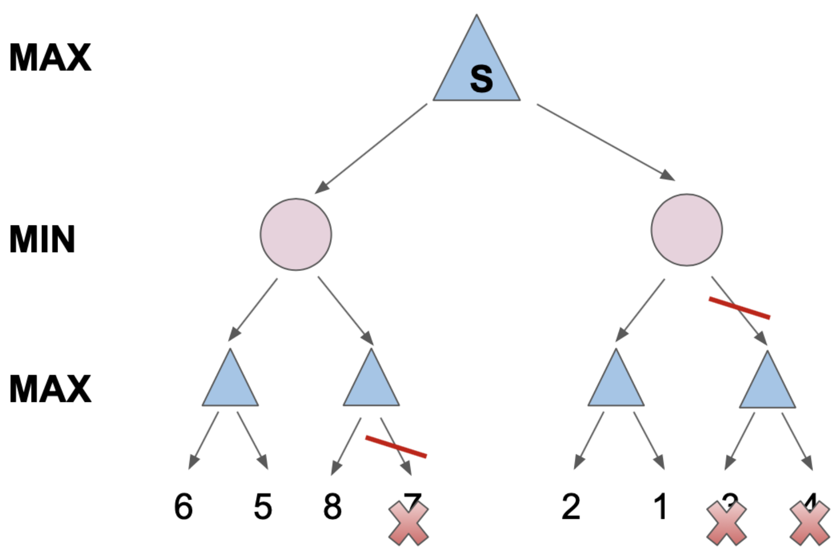
The value of the game is -49 (the goose might as well go for the bread before it gets caught). If Romeo defects, then the value of the game is 0 (the Goose moves towards Romeo).

2) [CA session] Problem 2

Consider running alpha-beta pruning on the following minimax tree. The children of each node will be expanded from left to right. Which nodes will be pruned (thus not being visited)?



Solution The leaves with values 7, 3 and 4 will be pruned.



3) **[Breakouts] Problem 3**

Consider the speed bump problem we did last week:

You're programming a self-driving car that can take you from home (position 1) to school (position n). At each time step, the car has a current position $x \in \{1, \ldots, n\}$ and a current velocity $v \in \{0, \ldots, m\}$. The car starts with $v = 0$, and at each time step, the car can either increase the velocity by 1, decrease it by 1, or keep it the same; this new velocity is used to advance x to the new position. The velocity is not allowed to exceed the speed limit m nor return to 0.

In addition, to prevent people from recklessly cruising down Serra Mall, the university has installed speed bumps at a subset of the n locations. The speed bumps are located at $B \subseteq \{1, \ldots, n\}$. The car is not allowed to enter, leave, or pass over a speed bump with velocity more than $k \in \{1, \ldots, m\}$. **Your goal is to arrive at position n with velocity 1 in the smallest number of time steps.**

Now let's add more information to this problem:

The university wants to remove the old speed bumps and install a single new speed bump at location $b \in \{1, \ldots, n\}$ to maximize the time it takes for the car to go from position 1 to n.

Let $T(\pi, B)$ be the time it takes to get from 1 to n if the car follows policy π if speed bumps B are present. If π violates the speed limit, define $T(\pi, B) = \infty$.

To simplify, assume $n = 6$ and $k = 1$. Again, there is exactly one speed bump. That is, $B = \{b\}$ with $b \in \{1, \ldots, n\}$.

$x = 1$ home	$x = 2$	$x = 3$	$x = 4$	$x = 5$	$x = 6$ school

Figure: The university will add a speed bump somewhere

(i) [5 points] Compute the worst case driving time, assuming you get to adapt your policy to the university's choice of speed bump location b: $\max_b \min_\pi T(\pi, \{b\})$. What values of b attain the maximum?

Solution Note that with $n = 6$, there are only two places where one can travel at a velocity of 2, from 2 to 4 or 3 to 5; in these cases, there can't be any speed bumps there. So if the speed bump is placed at $b \in \{1, 2, 5, 6\}$, the optimal policy has space to speed up to a velocity of 2 around the bump, so the total time is 4. However, if the speed bump is placed at $\boxed{b \in \{3, 4\}}$, then the optimal policy is to travel at a velocity of 1 the whole way which results in a total time of $\boxed{5}$, which is the worst case. Most common error was missing one of the cases for b. Also, there were a number of off-by-one errors (takes only 5 units to get from 1 to 6, not 6).

(ii) [5 points] Compute the best possible time assuming that you have to choose your policy before the university chooses the speed bump: $\min_\pi \max_b T(\pi, \{b\})$. Make sure to explain your reasoning.

Solution If we choose any policy that has velocity of 2, the university can place the speed bump in the appropriate place that results in a time of ∞. Therefore, we must choose a policy that only has velocity 1, which results in a time of $\boxed{5}$. Students should not assume that the university will definitely place speed bumps at $b \in \{3, 4\}$, but it's fine to acknowledge this as a possibility in your reasoning.