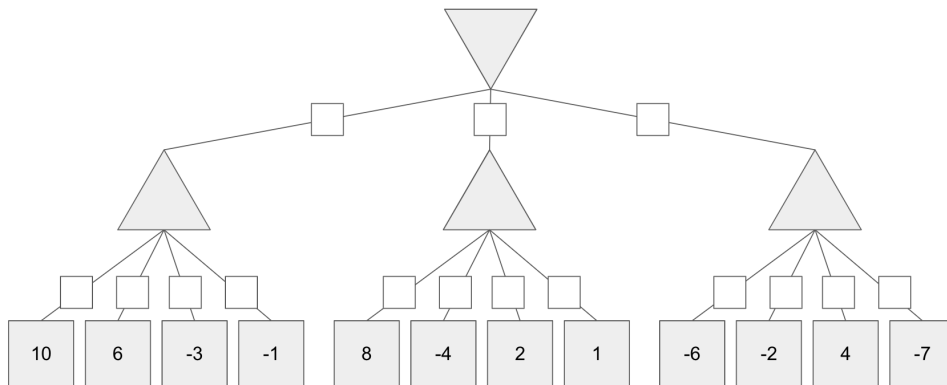


Q1. Coin Stars

In a new online game called Coin Stars, all players are walking around an $M \times N$ grid to collect **hidden coins**, which only appear when you're on top of them. There are also power pellets scattered across the board, which are visible to all players. If you walk onto a square with a power pellet, your power level goes up by 1, and the power pellet disappears. Players will also attack each other if one player enters a square occupied by another player. In an attack, the player with a higher power level will steal all the coins from the other player. If they have equal power levels, nothing happens. Each turn, players go in order to move in one of the following directions: {N, S, E, W}.

In this problem, you and your friend Amy are playing Coin Stars against each other. You are player 1, and your opponent Amy is player 2. Our state space representation includes the locations of the power pellets (x_{p_j}, y_{p_j}) and the following player information: (1) Each player's location (x_i, y_i) ; (2) Each player's power level l_i ; (3) Each player's coin count c_i .

- (a) Suppose a player wins by collecting more coins at the end of a number of rounds, so we can formulate this as a minimax problem with the value of the node being $c_1 - c_2$. Consider the following game tree where you are the maximizing player (maximizing the your net advantage, as seen above) and the opponent is the minimizer. Assuming both players act optimally, if a branch can be pruned, fill in its square completely, otherwise leave the square unmarked.



- None of the above can be pruned
- (b) Suppose that instead of the player with more coins winning, every player receives payout equal to the number of coins they've collected. Can we still use a multi-layer minimax tree (like the one above) to find the optimal action?
- Yes, because the update in payout policy does not affect the minimax structure of the game.
- Yes, but not for the reason above
- No, because we can no longer model the game under the updated payout policy with a game tree.
- No, but not for the reason above

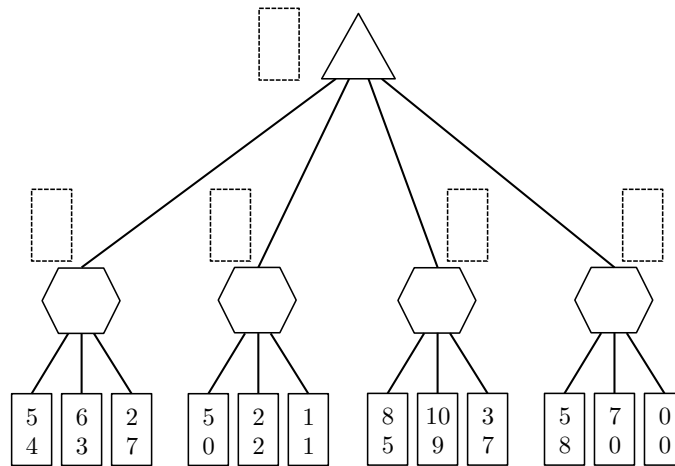
Q2. Game Trees and Pruning

You and one of the 188 robots are playing a game where you both have your own score. In the leaf nodes of the game tree, your score is on top and the robot's score on the bottom.

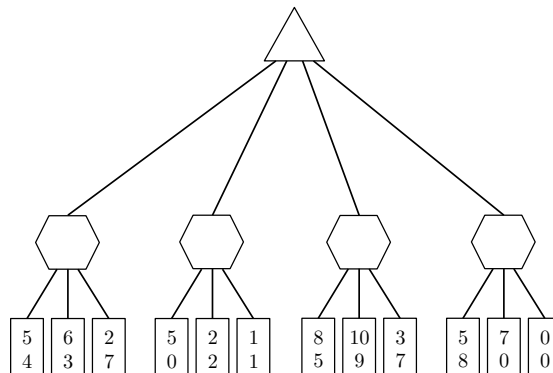
- The maximum possible score for either player is 10.
- You are trying to maximize your score, and you do not care what score the robot gets.
- The robot is trying to minimize the absolute difference between the two scores. In the case of a tie, the robot prefers a lower score. For example, the robot prefers (5,3) to (6,3); it prefers (5,3) to (0,3); and it prefers (3,3) to (5,5).

The figure below shows the game tree of your max node followed by the robots nodes for your four different actions.

- (a) Fill in the dashed rectangles with the pair of scores preferred by each node of the game tree.

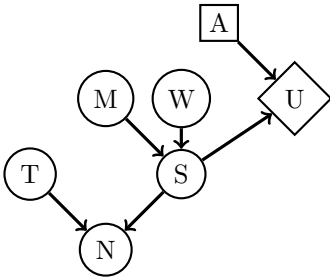


- (b) You can save computation time by using pruning in your game tree search. On the game tree above, put an 'X' on line of branches that do not need to be explored. Assume that branches are explored from left to right.
- (c) You now have access to an oracle that tells you the order of branches to explore that maximizes pruning. On the copy of the game tree below, put an 'X' on line of branches that do not need to be explored given this new information from the oracle.



Q3. Decision Networks and VPI

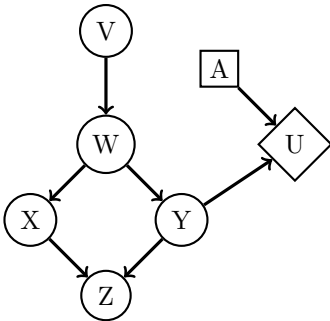
(a) Consider the decision network structure given below:



Mark all of the following statements that **could possibly be true**, for some probability distributions for $P(M), P(W), P(T), P(S|M, W)$, and $P(N|T, S)$ and some utility function $U(S, A)$:

- (i) $VPI(T) < 0$ $VPI(T) = 0$ $VPI(T) > 0$ $VPI(T) = VPI(N)$
(ii) $VPI(T|N) < 0$ $VPI(T|N) = 0$ $VPI(T|N) > 0$ $VPI(T|N) = VPI(T|S)$
(iii) $VPI(M) > VPI(W)$ $VPI(M) > VPI(S)$ $VPI(M) < VPI(S)$ $VPI(M|S) > VPI(S)$

(b) Consider the decision network structure given below.



Mark all of the following statements that are **guaranteed to be true**, regardless of the probability distributions for any of the chance nodes and regardless of the utility function.

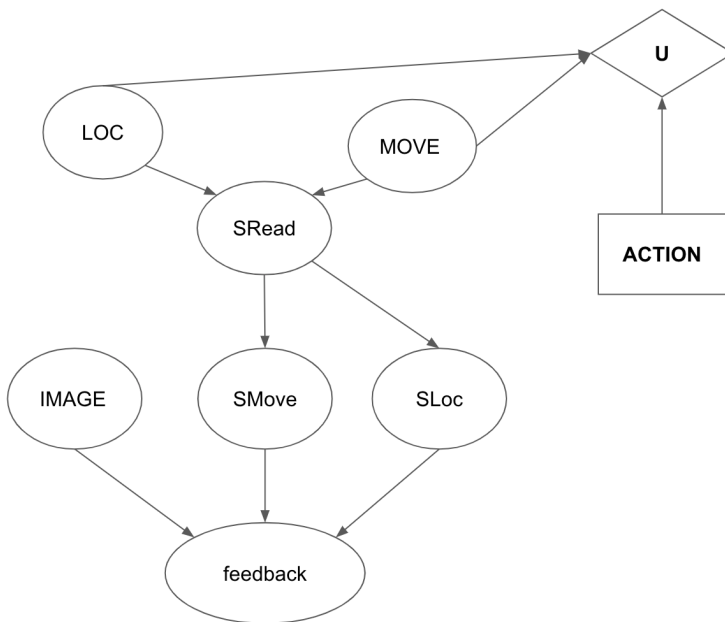
- (i)
 $VPI(Y) = 0$
 $VPI(X) = 0$
 $VPI(Z) = VPI(W, Z)$
 $VPI(Y) = VPI(Y, X)$
- (ii)
 $VPI(X) \leq VPI(W)$
 $VPI(V) \leq VPI(W)$
 $VPI(V | W) = VPI(V)$
 $VPI(W | V) = VPI(W)$
- (iii)
 $VPI(X | W) = 0$
 $VPI(Z | W) = 0$
 $VPI(X, W) = VPI(V, W)$
 $VPI(W, Y) = VPI(W) + VPI(Y)$

Q4. Vehicle Perception Indication

A vehicle is trying to identify the situation of the world around it using a set of sensors located around the vehicle.

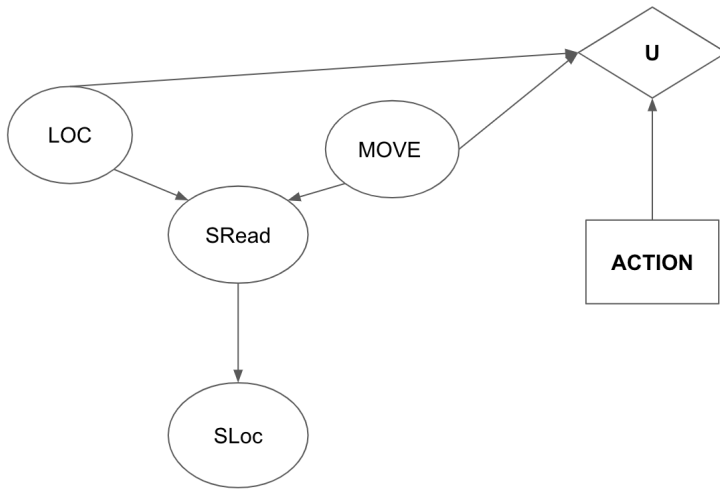
Each sensor reading (SRead) is based off of an object's location (LOC) and an object's movement (MOVE). The sensor reading will then produce various values for its predicted location (SLoc) and predicted movement (SMove). The user will receive these readings, as well as the the image (IMAGE) as feedback.

- (a) The vehicle takes an action, and we assign some utility to the action based on the object's location and movement. Possible actions are MOVE TOWARDS, MOVE AWAY, and STOP. Suppose the decision network faced by the vehicle is the following.



- (i) Based on the diagram above, which of the following **could possibly be true**?
- VPI (Image) = 0
 - VPI (SRead) < 0
 - VPI (SMove, SRead) > VPI (SRead)
 - VPI (Feedback) = 0
 - None of the above
- (ii) Based on the diagram above, which of the following **must necessarily be true**?
- VPI (Image) = 0
 - VPI (SRead) = 0
 - VPI (SMove, SRead) = VPI (SRead)
 - VPI (Feedback) = 0
 - None of the above

Let's assume that your startup has less money, so we use a simpler sensor network. One possible sensor network can be represented as follows.



You have distributions of $P(\text{LOC})$, $P(\text{MOVE})$, $P(\text{SRead}|\text{LOC}, \text{MOVE})$, $P(\text{SLoc}|\text{SRead})$ and utility values $U(a, l, m)$.

(b) Complete the equation for determining the expected utility for some ACTION a .

$$EU(a) = \left(\text{(i)} \quad \text{(ii)} \quad \text{(iii)} \right) U(a, l, m)$$

- (i) $\sum_l P(l)$ $\sum_{sloc} P(sloc|l)$ $\sum_l \sum_{sloc} P(sloc|l)$ 1
- (ii) $\sum_m P(m)$ $\sum_m P(sloc|m)$ $\sum_l \sum_m \sum_{sloc} P(sloc|l)P(sloc|m)$ 1
- (iii) $\sum_l \sum_m \sum_{sloc} P(sloc|l)P(sloc|m)$ $+\sum_l \sum_m \sum_{sloc} P(sloc|l)P(sloc|m)$
- $+\sum_l \sum_m \sum_{sloc} P(sloc|l, m)P(l)P(m)$ *1

(c) Your colleague Bob invented a new sensor to observe values of $S\text{Loc}$.

(i) Suppose that your company had no sensors till this point. Which of the following expression is equivalent to $VPI(S\text{Loc})$?

- $VPI(S\text{Loc}) = (\sum_{sloc} P(sloc) \text{MEU}(S\text{Loc} = sloc)) - \max_a EU(a)$
- $VPI(S\text{Loc}) = \text{MEU}(S\text{Loc}) - \text{MEU}(\emptyset)$
- $VPI(S\text{Loc}) = \max_{sloc} \text{MEU}(S\text{Loc} = sloc) - \text{MEU}(\emptyset)$
- None of the above

(ii) Gaagle, an established company, wants to sell your startup a device that gives you $S\text{Read}$. Given that you already have Bob's device (that gives you $S\text{Loc}$), what is the maximum amount of money you should pay for Gaagle's device? Suppose you value \$1 at 1 utility.

- $VPI(S\text{Read})$
- $VPI(S\text{Read}) - VPI(S\text{Loc})$
- $VPI(S\text{Read}, S\text{Loc})$
- $VPI(S\text{Read}, S\text{Loc}) - VPI(S\text{Loc})$
- None of the above