

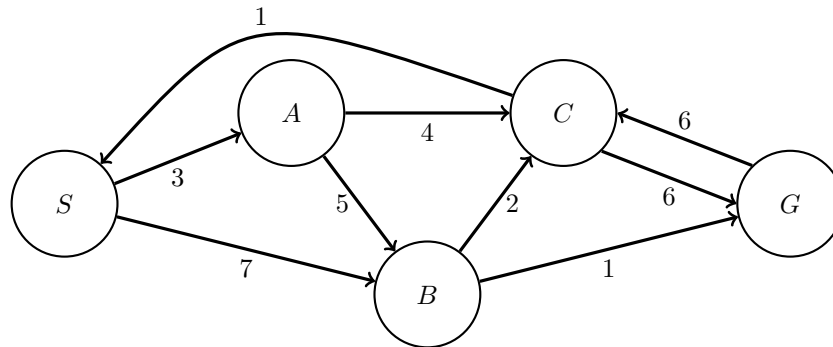
Learning Objectives

- Represent various problems (e.g., search, min/max, rectified linear) as LP/IP
- Explore the connections between different topics

IMPORTANT: Please return the completed activity to the front at the end of class. Write your name and Andrew ID on the top of the first page. You will get 1 point per full page completed (4 total).

Q1. Cost-Based Search as IP

In this problem we are going to explore how to formulate search problems as Integer Programming problems. For the sake of simplicity we will be considering the weighted directed graph below (S is the start, and G is the goal):



We can derive a representation for any path¹ in a graph by considering variables to represent edges in the graph each with domain $\{0,1\}$; value 0 if the edge is not in the path and 1 if the edge is in the path. Specifically, define a binary variable $x_{X \rightarrow Y}$ if there is an edge $X \rightarrow Y$ in the graph.

For example, the path $S \rightarrow A \rightarrow C \rightarrow G$ can also be seen as the set of edges $\{S \rightarrow A, A \rightarrow C, C \rightarrow G\}$. We can then define $x_{S \rightarrow A}$ to be the indicator variable for whether the edge $S \rightarrow A$ is on the path, $x_{S \rightarrow B}$ to be the indicator variable for whether the edge $S \rightarrow B$ is on the path, and so on. Using this binary representation, $S \rightarrow A \rightarrow C \rightarrow G$ can be represented as:

$$(x_{S \rightarrow A} = 1 \quad x_{S \rightarrow B} = 0 \quad x_{A \rightarrow B} = 0 \quad x_{A \rightarrow C} = 1 \quad x_{B \rightarrow C} = 0 \quad x_{B \rightarrow G} = 0 \quad x_{C \rightarrow S} = 0 \quad x_{C \rightarrow G} = 1 \quad x_{G \rightarrow C} = 0)$$

If we fix the order of the indicator variables in the order above, the path can be represented as a 9-tuple: (1, 0, 0, 1, 0, 0, 0, 1, 0)

(a) Answer the following questions below about the representation.

(i) Write the 9-tuple binary representation for the path $S \rightarrow A \rightarrow B \rightarrow C \rightarrow G$

Answer:
 (1, 0, 1, 0, 1, 0, 0, 1, 0)

(ii) Write the 9-tuple binary representation for the path $A \rightarrow C \rightarrow S \rightarrow B$

Answer:
 (0, 1, 0, 1, 0, 0, 1, 0, 0)

¹Note that a “path” is different from a “walk”; for our purpose each node can only appear in the path at most once.

- (iii) Write the path that corresponds to
- $(0, 0, 1, 0, 1, 0, 0, 0, 0)$

Answer: $A \rightarrow B \rightarrow C$

- (b) Note with our binary representation, it is possible to make “paths” that do not start from
- S
- and end at
- G
- , and some of the 9–tuples are not even valid paths at all. Consider the constraints on valid paths in this representation.

From the binary representation itself, it is clear that we need to constrain the indicator variables to be an integer between 0 and 1 inclusive. To find valid paths from the start to the goal, we need to start by imposing some additional constraints on the 9–tuples.

- (i) Write the constraint(s) (in
- inequality form**
-) that a valid path in the specific graph above must start at
- S
- . (Hint: What must be true about the values or sums of values of the indicator variables involving
- S
- defined above? Remember that constraints must be linear sums of variables.)

Answer:

$$\begin{aligned} x_{S \rightarrow A} + x_{S \rightarrow B} &\leq 1 \\ -x_{S \rightarrow A} - x_{S \rightarrow B} &\leq -1 \\ x_{C \rightarrow S} &\leq 0 \\ -x_{C \rightarrow S} &\leq 0 \end{aligned}$$

- (ii) Similarly, write the constraint(s) that a valid path must end at
- G
- (in inequality form).

Answer:

$$\begin{aligned} x_{C \rightarrow G} + x_{B \rightarrow G} &\leq 1 \\ -x_{C \rightarrow G} - x_{B \rightarrow G} &\leq -1 \\ x_{G \rightarrow C} &\leq 0 \\ -x_{G \rightarrow C} &\leq 0 \end{aligned}$$

- (iii) Write a 9–tuple (binary representation) that satisfies the constraints in (i) and (ii) but does
- not**
- represent a valid path from
- S
- to
- G

Answer:Multiple answers, one possibility is $(1, 0, 0, 0, 0, 0, 0, 1, 0)$

- (c) Let us define a non-terminal node as a node that is neither the defined start node (
- S
-) nor the goal node (
- G
-). Note that the constraints in part (b) do
- not**
- guarantee that all other nodes in the path are non-terminal.

- (i) We want to ensure that the path only passes through each non-terminal node (e.g.,
- B
-) at most once. Write the constraint(s) that node
- B
- does not appear more than once on the path (inequality form).

Answer:

$$\begin{aligned} &x_{S \rightarrow B} + x_{A \rightarrow B} \leq 1 \quad \text{and correct Cii} \\ \text{OR} & \quad x_{B \rightarrow C} + x_{B \rightarrow G} \leq 1 \quad \text{and correct Cii} \\ \text{OR} & \quad \text{both} \end{aligned}$$

- (ii) Finally, to ensure B is non-terminal, we have to make sure that if there is an edge to B in a path, then there should also be an edge from B to some other node. If there is an edge to B but not from B , then the path has a dead-end at B . Similarly, if there is an edge from B to another node but no edge to B , then B is a starting node but not S . Write down the corresponding constraint(s) about B 's non-terminal position (inequality form).

Answer:

$$\begin{aligned}x_{S \rightarrow B} + x_{A \rightarrow B} - x_{B \rightarrow C} - x_{B \rightarrow G} &\leq 0 \\ -x_{S \rightarrow B} - x_{A \rightarrow B} + x_{B \rightarrow C} + x_{B \rightarrow G} &\leq 0\end{aligned}$$

- (d) We have all the constraints to ensure a valid path (actually we don't, but for now assume we do). Write the objective function for the IP search problem representing the graph above. (Hint: the goal of solving this IP problem is to minimize the objective function that corresponds to the search objective.)

Answer:

$$3x_{S \rightarrow A} + 7x_{S \rightarrow B} + 5x_{A \rightarrow B} + 4x_{A \rightarrow C} + 2x_{B \rightarrow C} + 1x_{B \rightarrow G} + 1x_{C \rightarrow S} + 6x_{C \rightarrow G} + 6x_{G \rightarrow C}$$

- (e) Let us take a closer look at the representation. As hinted before, the current constraints **DO NOT** ensure a valid path for graphs in general.

- (i) Come up with a directed graph such that there exists a tuple that passes all the previously mentioned constraints but does not represent a valid path on the graph. You should be able to construct such an example with less than 10 edges. The weights of the edges are not important for this question, you may omit them.

Draw your graph, and give the counterexample tuple. As a reminder, the length of your tuple should be the same as the number of edges in your graph. Clearly label the correspondence between the elements in your tuple and the edges on the graph

Answer:



The invalid representation is $(1, 1)$. Basically anything with a loop outside the path is a valid counter-example.

- (ii) Despite the existence of these invalid tuples, they will never be returned by the IP above (as defined by the previous constraints and objective function) as long as every edge has a positive cost. Briefly explain why this is the case.

Answer:

Because having an extra cycle will increase the total cost as defined in the objective function. Therefore it cannot be the optimal solution for the IP.

- (f) Define a new search algorithm to be: formulate the search as an IP problem, then run an IP solver (which is optimal and complete) and return the solution as a path. Is this new search algorithm:
- (i) Complete? Yes No
- (ii) Optimal? Yes No
- (g) What will happen if we use an LP solver instead of the IP solver (still constraining the binary representations to be in range $[0, 1]$)? Will the returned solution (value assignments of the binary variables) and/or the object value be different? If they might differ, give the high-level description of such an example and explain why they might differ. Otherwise, briefly explain why they will be the same.

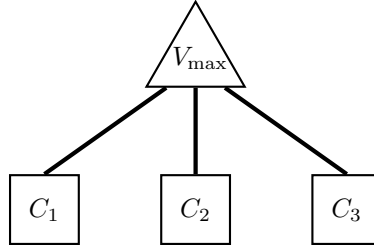
Answer:

The returned value assignment might be different if there are multiple paths with the optimal cost. But the objective value will be the same (the least cost from start to goal) as the IP version is already optimal.

Q2. Adversarial Search as IP (Bonus Question)

In this problem² we are going to explore how to do a similar representation for the adversarial search, specifically the **Minimax**. As the Minimax tree consists of different layers of maximizing and minimizing nodes, an intuitive way to get started is to represent each individual node and put them together.

- (a) As shown in the figure below, let us start with a maximizing node with three children. For now, assume that all the children are just constant values (not other nodes).



If we define a variable V_{\max} to represent the true value of the maxie node (without any pruning), then we can define the following constraints to represent that the value of the node is greater or equal to all of its children.

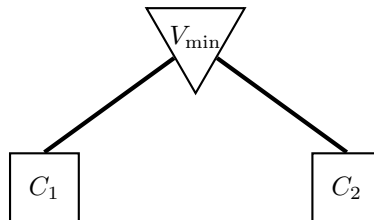
$$C_1 - V_{\max} \leq 0 \quad C_2 - V_{\max} \leq 0 \quad C_3 - V_{\max} \leq 0$$

These constraints alone do not guarantee that V_{\max} will end up being the maximum value of its children. However, we can enforce that with the help of the objective function. Define an objective function such that with this objective function, the value of V_{\max} will be equal to the maximum value of its children when we solve the IP. (Hint: in the standard form of LP/IP we are *minimizing* the objective function).

Answer:

$$V_{\max}$$

- (b) Now similarly write the constraints and objective function for a minimizing node, where we define a variable V_{\min} to represent the value of the minnie node.



- (i) Write the constraints in inequality form (Hint: there should be two constraints).

Answer:

$$V_{\min} - C_1 \leq 0 \quad V_{\min} - C_2 \leq 0$$

- (ii) Write the objective function to be minimized.

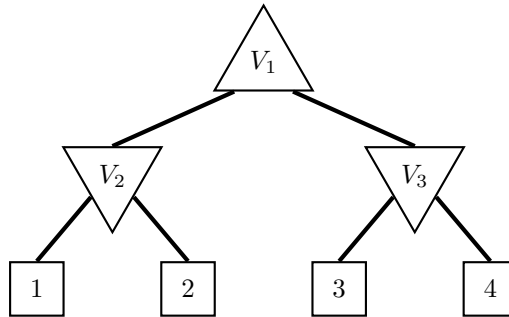
Answer:

$$-V_{\min}$$

²Although technically still within scope, this is a **VERY HARD** question (part (a-c) should be fine, the rest needs some good intuition). You are not expected to handle questions at this level of difficulty on your own in any part of this course. The background of this question is the thought process of a TA trying to design a problem on Minimax as IP but later realizing it is too hard to be anywhere in the course. It should give you some inspiration on how to approach a problem. For those that are interested, you are encouraged to think about other smarter ways to convert minimax into an IP.

- (c) Now that we have two separate IPs for maxie and minnie nodes respectively, we can now put them together to solve the more complex Minimax problem.

Specifically, we start with the following Minimax tree:



- (i) First, write the values of the three internal nodes if we run minimax without pruning.

Answer:

$$V_1 = 3 \quad V_2 = 1 \quad V_3 = 3$$

- (ii) Putting all constraints in part (a) and (b) together, we have the following constraints (you should complete the omitted constraints when working on the rest of this question).

$$-V_1 + V_2 \leq 0 \quad -V_1 + V_3 \leq 0 \quad \text{and the constraints in your part (b)(i)}$$

Since we have two types of objective functions (part (a) and (b)(ii)), and now we want to optimize both of them, the simplest way to achieve this is to add them together.

Write the objective function for the Minimax tree above using variables V_1, V_2, V_3 (Hint: we are minimizing this objective, and your expression should include all three variables).

Answer:

$$V_1 - V_2 - V_3$$

- (iii) The value assignment in part (c)(i) passes all the constraints. Write the value of the objective function (defined in part (c)(ii)) given this value assignment.

Answer:

$$-1$$

- (iv) Unfortunately, this naive method of converting minimax to IP does not work. To illustrate this, find another set of value assignments to V_1, V_2 , and V_3 such that:

- V_1 is assigned to a different value
- The value assignment passes all the constraints in part (c)(ii)
- The objective function returns the same value for this value assignment as in part (c)(iii)

Write the new assignment in the form of " $V_1 = \text{some constant}$ $V_2 = \text{some constant}$ $V_3 = \text{some constant}$ "

Answer:

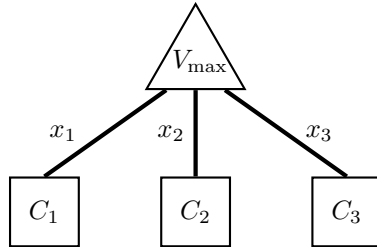
$$\text{Multiple answers, one possibility is } V_1 = 2 \quad V_2 = 1 \quad V_3 = 2$$

Because this new assignment has the same objective value as the correct one in part (c)(i), it is possible for the IP to return it as the final solution, showing that the current definition of the IP cannot be used for solving minimax.

You can try to play around with editing constraints and/or the objective function. However, this variable representation (defining one variable for each node to represent their values) is not expressive enough to solve the minimax problem.

- (d) To fix this problem, we should revisit the representation for the maximizing node. Previously we used the constraints to specify that its value is no less than its children, and used the objective function to represent that its value is the actual maximum value among its children. Now we should try doing both things with **constraints only**.

Revisiting the cost-based search as an IP problem, one way to accomplish this is to define binary variables to represent each action, such that we can enforce that the value of the parent is the same as one of its children. So as shown in the figure below, in addition to defining V_{\max} to represent the value of the maximizing node, we also define three binary variables x_1, x_2, x_3 to represent whether the corresponding action is chosen (in this case assume C_1, C_2 , and C_3 are constants).



We keep the following constraints the same to enforce V_{\max} must be greater than or equal to its children:

$$C_1 - V_{\max} \leq 0 \quad C_2 - V_{\max} \leq 0 \quad C_3 - V_{\max} \leq 0$$

But we also need extra constraints to enforce that V_{\max} must take the value of one of its children.

- (i) Write the constraint(s) to enforce **exactly one** action is taken at the maximizing node

Answer:

$$\begin{aligned} x_1 + x_2 + x_3 &\leq 1 \\ -x_1 - x_2 - x_3 &\leq -1 \end{aligned}$$

- (ii) Write the constraint(s) to enforce the value of V_{\max} is the value of the chosen child

Answer:

$$\begin{aligned} V_{\max} - C_1 \cdot x_1 - C_2 \cdot x_2 - C_3 \cdot x_3 &\leq 0 \\ -V_{\max} + C_1 \cdot x_1 + C_2 \cdot x_2 + C_3 \cdot x_3 &\leq 0 \end{aligned}$$

- (iii) Explain in 1-2 sentences why this formulation does not work if the children are values of other nodes (variables) instead of constants (so instead of constants C_1, C_2, C_3 we used variables V_1, V_2, V_3).

Answer:

It won't work because in that case $V_i \cdot x_i$ will be a multiplication between variables, which is not allowed in LP/IP

- (e) Now it's time to think about a sub-problem, which is how to represent $x \cdot V$ (where both x and V are variables) using *linear* constraints.

This would not be possible in general, but lucky for us, there are additional constraints (listed below) that we can utilize in the minimax problem so that we *can* present the multiplication using linear constraints.

- x is binary
- V is bounded (i.e., there is a maximum possible value and a minimum possible value for it)

- (i) Note that V represents the value of a node in the minimax tree. Briefly explain why it is bounded.

Answer:

It is bounded by the minimum and maximum value of the leaves

- (ii) Hence, we can just assume the value of *all* nodes are bounded in $[0, 1]$ for the rest of this problem. Now let us introduce a new variable V' and enforce $V' = x \cdot V$ using linear constraints. First, we observe that the following are true

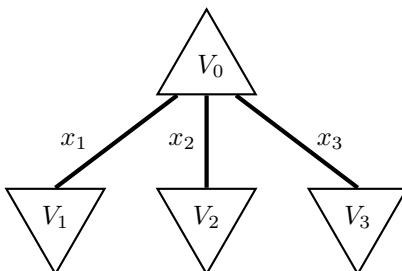
$$V' \leq x \quad \text{and} \quad V' \leq V$$

Write the objective function on V' that enforces $V' = x \cdot V$ given the constraints above. (Hint: this should be very similar to your answer in part (a))

Answer:

$$-V'$$

- (f) Returning from the detour, now we have everything we need to represent a maximizing node with other nodes as children. In the following graph, V_0, V_1, V_2, V_3 are the variables representing the values of each node, and x_1, x_2, x_3 are the binary variables representing whether each action is the chosen option for the node.



- (i) Write **all** the constraints relevant to the V_0 node (you should reuse your constraints from parts (d) and (e)). You may introduce extra variables (e.g., V' in part (e)(ii)).

Answer:

$$\begin{aligned}
 & -V_0 + V_1 \leq 0 \quad -V_0 + V_2 \leq 0 \quad -V_0 + V_3 \leq 0 \\
 & V'_1 \leq x_1 \quad V'_2 \leq x_2 \quad V'_3 \leq x_3 \quad V'_1 \leq V_1 \quad V'_2 \leq V_2 \quad V'_3 \leq V_3 \\
 & V_0 - V'_1 - V'_2 - V'_3 \leq 0 \quad -V_0 + V'_1 + V'_2 + V'_3 \leq 0 \\
 & x_1 + x_2 + x_3 \leq 1 \quad -x_1 - x_2 - x_3 \leq -1 \\
 & x_1, x_2, x_3 \leq 1 \quad -x_1, -x_2, -x_3 \leq 0 \quad x_i \in \mathbb{Z}
 \end{aligned}$$

The first row ensures V_0 is no less than its children. The second row ensures $V'_i = x_i \cdot V_i$. The third row ensures the value of the maximizing node is actually the value of the chosen child. The fourth row ensures exactly one action is chosen. The last row ensures binary variables

- (ii) Write the objective function for this maximizing node.

Answer:

$$-V'_1 - V'_2 - V'_3 \text{ or } -V_0$$

- (iii) Identify which constraint(s) you would change if all the nodes are flipped (that is, the root is now a minimizing node while the children are maximizing), and explain what you would change them to. Also, explain whether you need to change the objective function, and if so, what it should be changed to.

Answer:

We only need to flip the constraints in the first row (e.g., $V_0 - V_1 \leq 0$). This is because instead of enforcing the value of the root is no less than the children, now we want to enforce the value of the children is no greater than the children.

We do not change the rest (including the objective function) because they enforce that the value of the root node is the same as the value of the chosen child.

- (g) Finally, it is time to put everything together again. As we did in part (c), we can use the following steps to combine the individual IPs of each node:

- Stack all the constraints
- Sum all the objective functions to be the new objective function

However, in part (c), we have seen that adding up all the objective functions did not work. Explain, on a high level, why the updated representation, especially the objective function, works when combining different nodes together but not the original naive representation.

Answer:

In the first representation, we are trying to optimize both the minimizing nodes and the maximizing nodes directly. They have conflicting objectives, dependent on each other, and are unbounded by the constraints.

For the new representation, the objective function is used to enforce “the value of the parent node is the same as one of its children”, which is the same for both types of node.

- (h) Briefly explain why the integer constraints on the action variables are necessary. Describe what might happen if we lift those integer constraints.

Answer:

It does not work for LP because LP cannot enforce that the value of the parent node is the value of one of its children. In the example in part (f)(i), exactly one of the V'_i equals V_i ($i \in \{1, 2, 3\}$), the rest are enforced to be 0. If we lift the integer constraint, more than one V'_i might be non-zero, so the final value of V_0 might be greater than any of its children’s values.