Foundations of Artificial Intelligence

Alberto Maria Metelli

Exercise Session
10-22-2021

Registrations of the Exercise Sessions on WeBeep

Registration Links also available at

https://albertometelli.github.io/teaching/2021-teaching-fai

New! Some exercises shown during the sessions available at the same page

New! Feedback about Exercise Sessions

https://bit.ly/3C3fs5e

Please fill the form!

Exercise 3.3

Consider the following two-player zero-sum game. The game begins with a pile of seven bricks. On your move, you must split one pile of bricks into two piles. You may not split a pile of bricks into two equal piles. If it is your turn and all the piles of bricks have either one or two bricks, you have lost the game.

1. Formalize the problem
2. Apply the minimax algorithm for finding the best action for the max player at the root.
3. Apply the minimax algorithm with alpha-beta pruning for finding the best action for the max player at the root.

Exercise 3.3 - Problem formalization

$S = \{\{p_1, p_2, \ldots, p_n\} \ : p_i \in \{1, 2, \ldots, 7\} \text{ and } \sum_i p_i = 7\}$

$A = \{(i, p) : i \in \{1, 2, \ldots, 7\} \text{ and } p \in \{1, 2, \ldots, 6\}\}$

(i, p) is **applicable** in $\{p_1, p_2, \ldots, p_n\}$ iff $i \in \{1, 2, \ldots, n\}$ and $p \in \{1, 2, \ldots, \text{floor}((p_i\text{-}1)/2)\}$

result$(\{p_1, p_2, \ldots, p_{i-1}, \mathbf{p_i}, p_{i+1}, \ldots, p_n\}, (i, \mathbf{p})) = \{p_1, p_2, \ldots, p_{i-1}, \mathbf{p}, \mathbf{p_i\text{-}p}, p_{i+1}, \ldots, p_n\}$

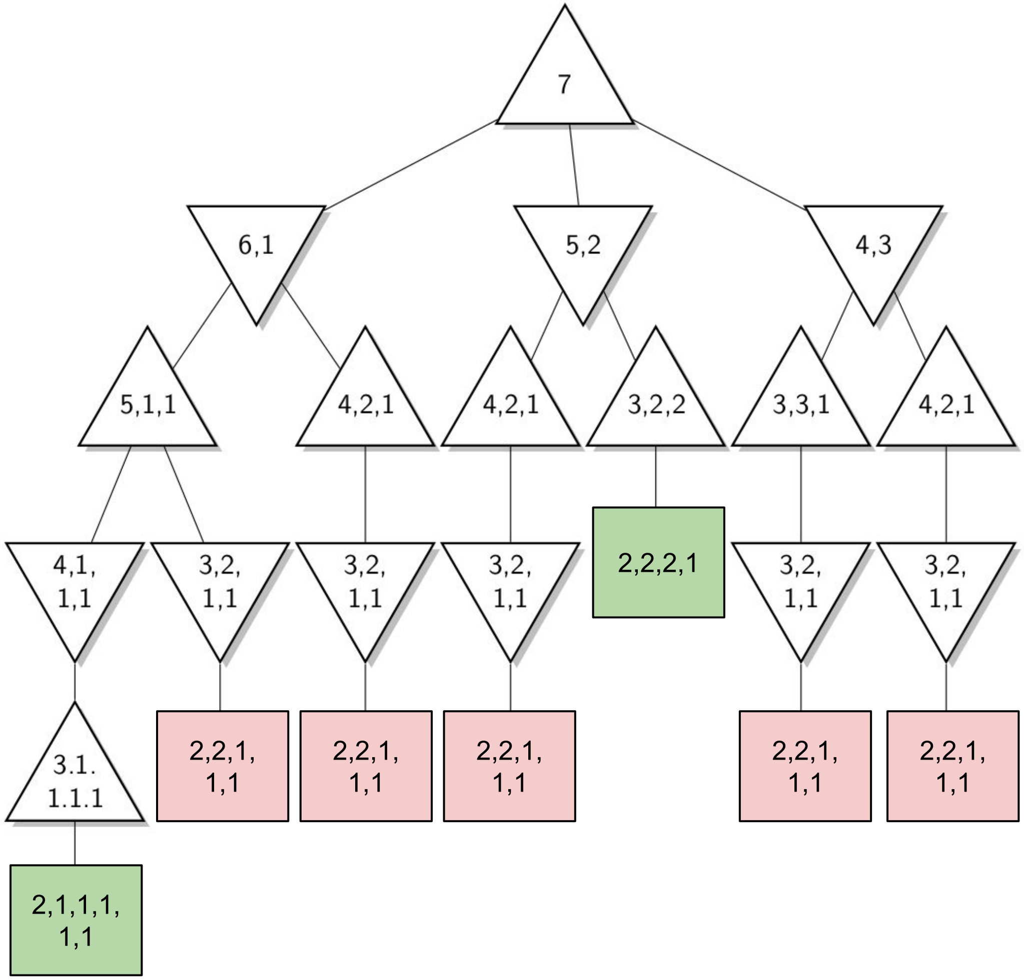
terminal$(\{p_1, p_2, \ldots, p_n\})$ iff $p_i \in \{1, 2\}$ for all $i \in \{1, 2, \ldots, n\}$

utility$(\{p_1, p_2, \ldots, p_n\}) = +1$ iff the current player is MIN

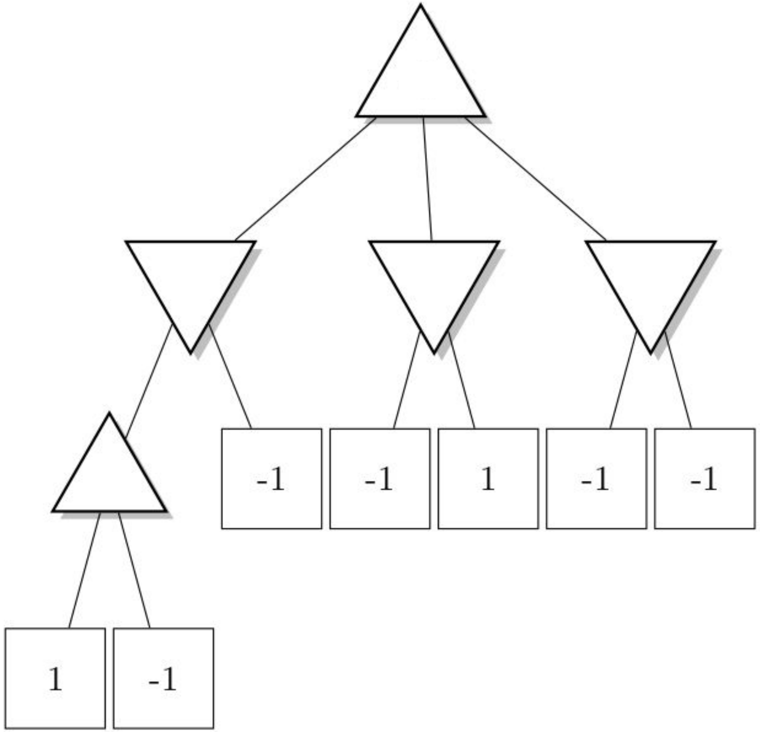
Exercise 3.3 - Game tree

+1

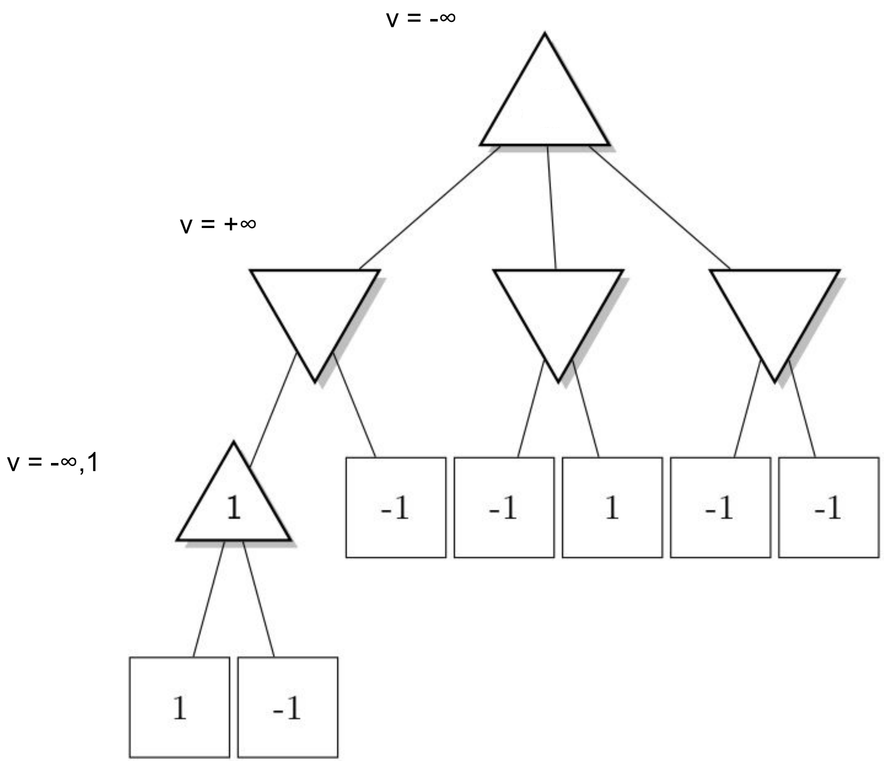
-1



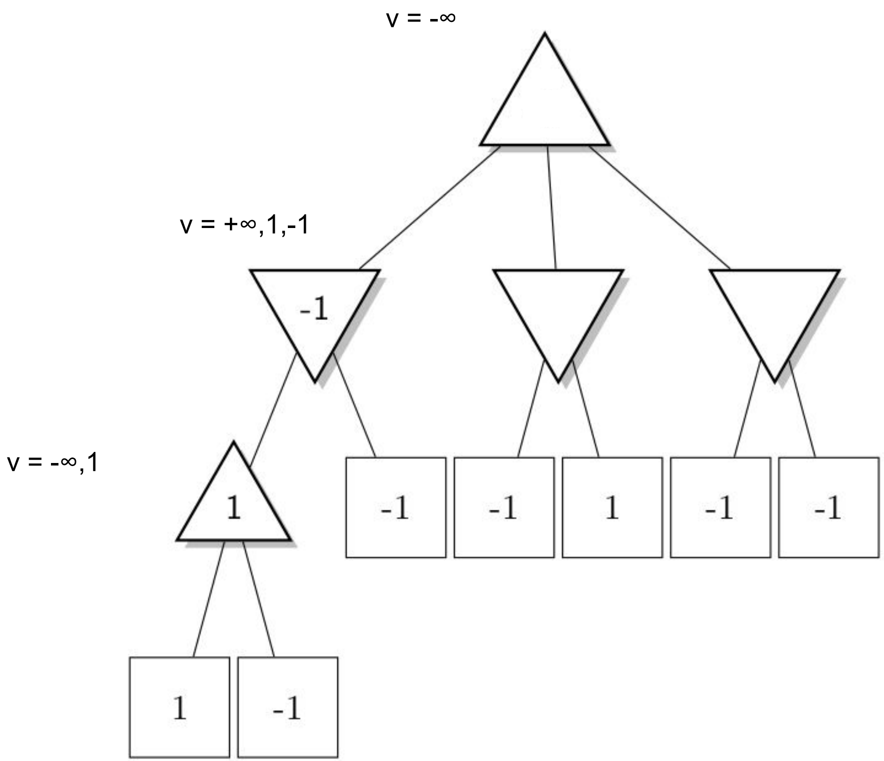
MINIMAX Algorithm



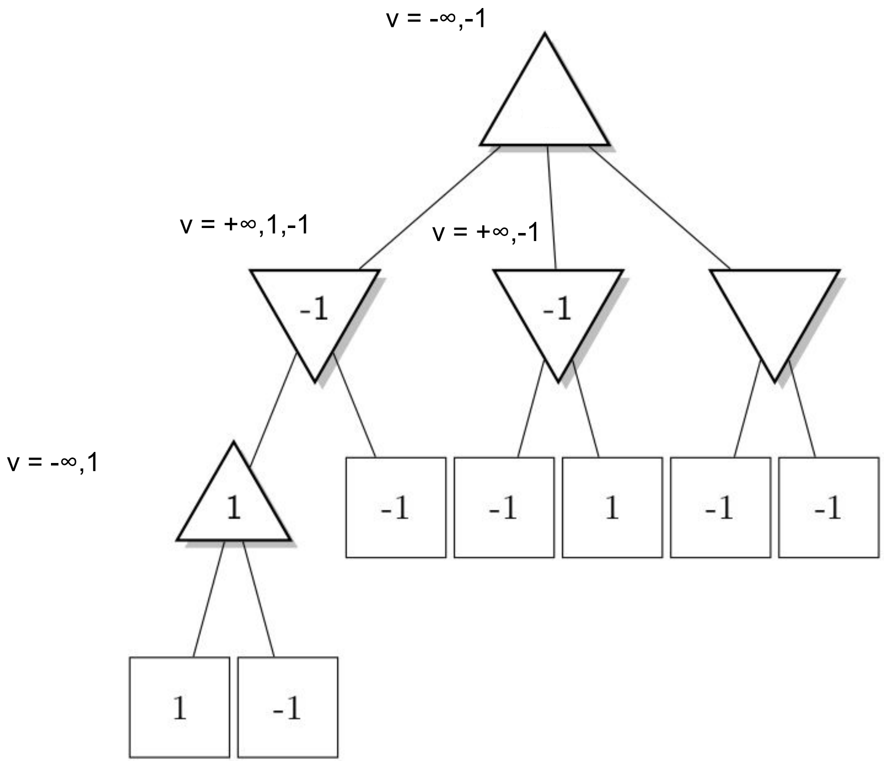
MINIMAX Algorithm



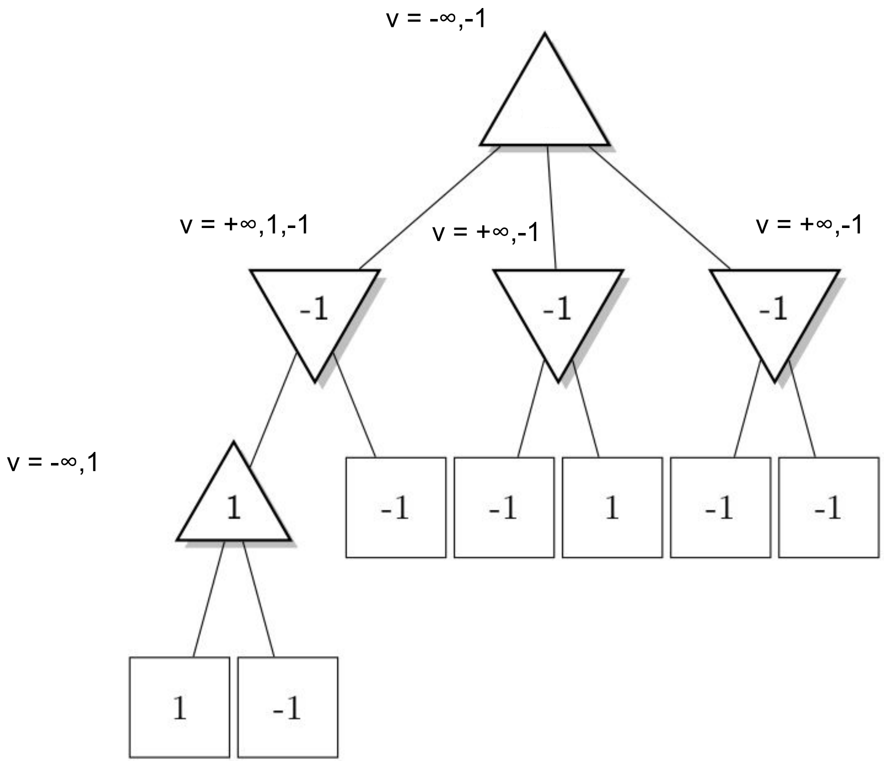
MINIMAX Algorithm



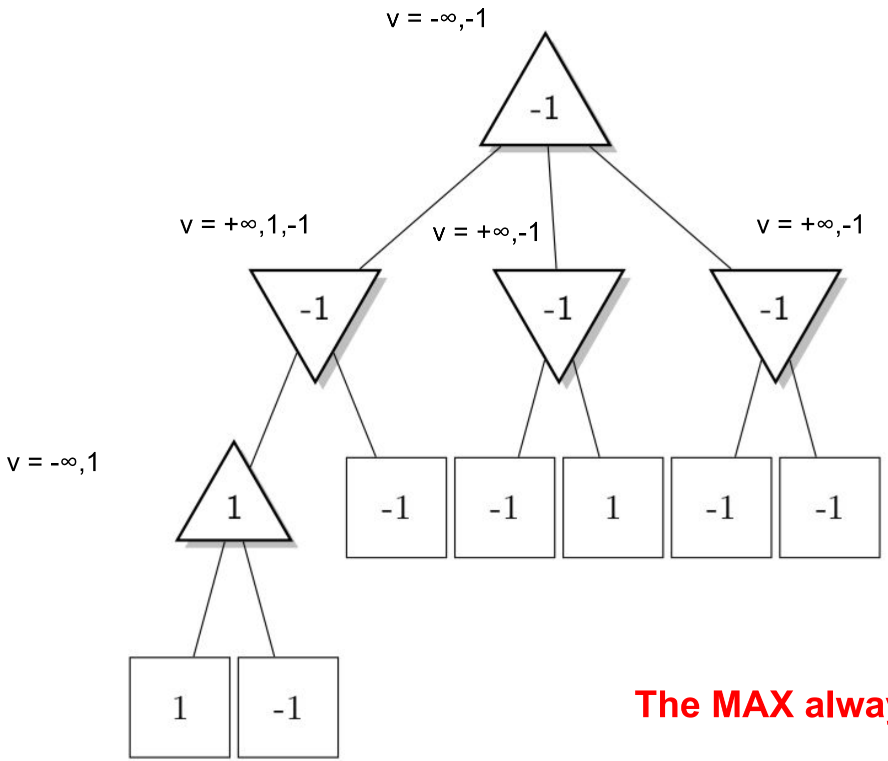
MINIMAX Algorithm



MINIMAX Algorithm

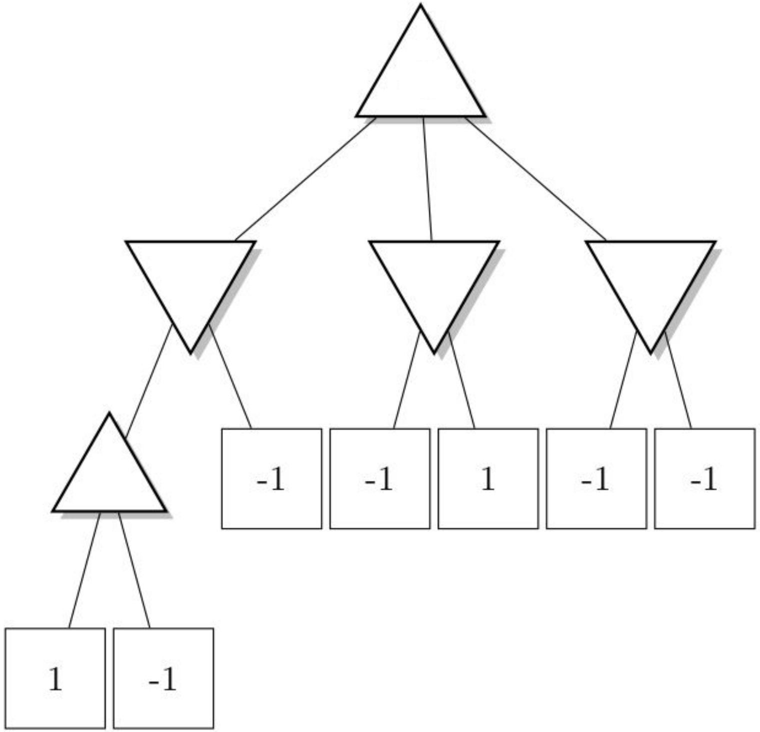


MINIMAX Algorithm

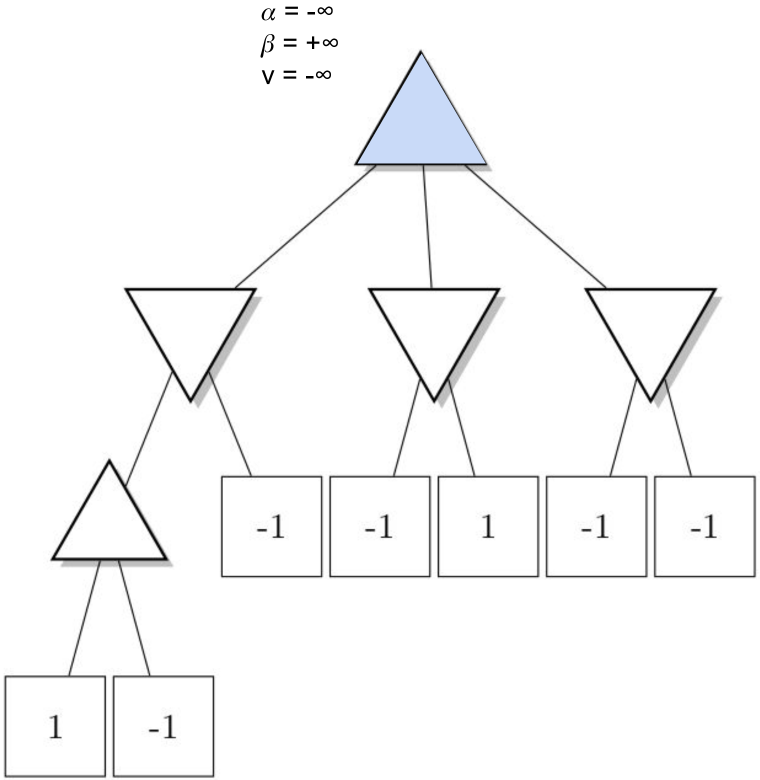


The MAX always loses!

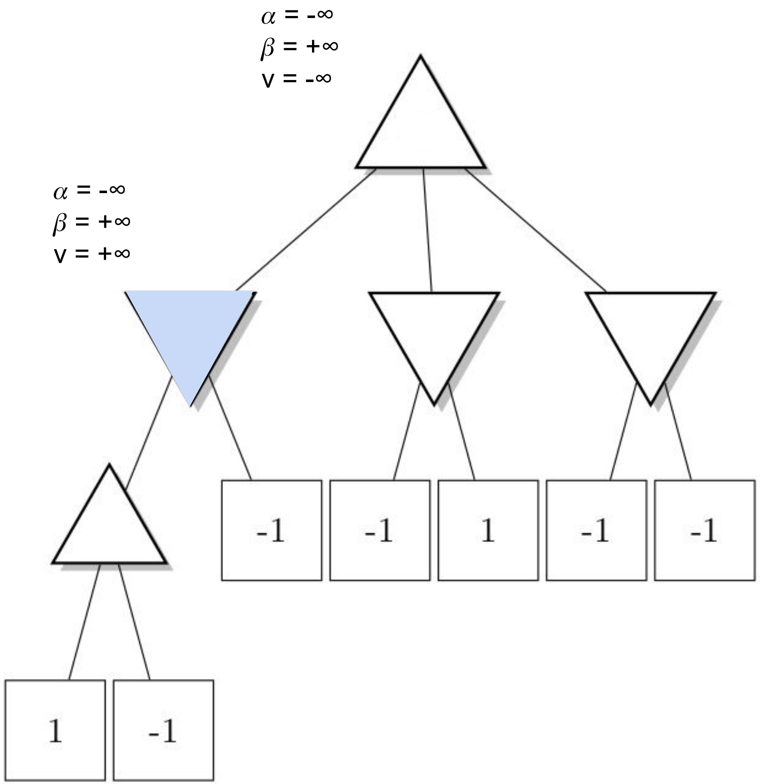
α-β Pruning



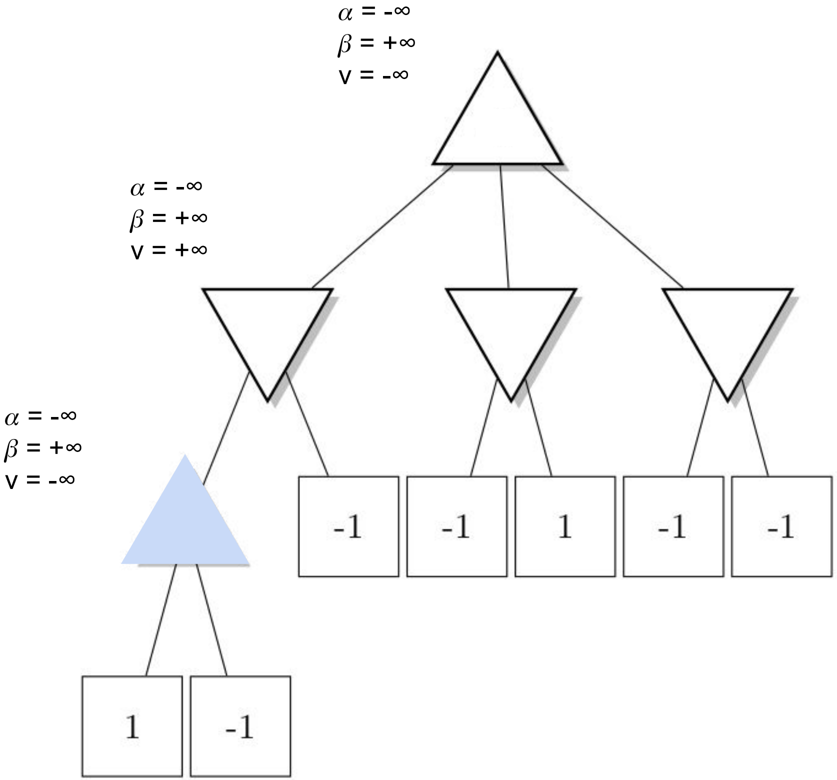
α-β Pruning



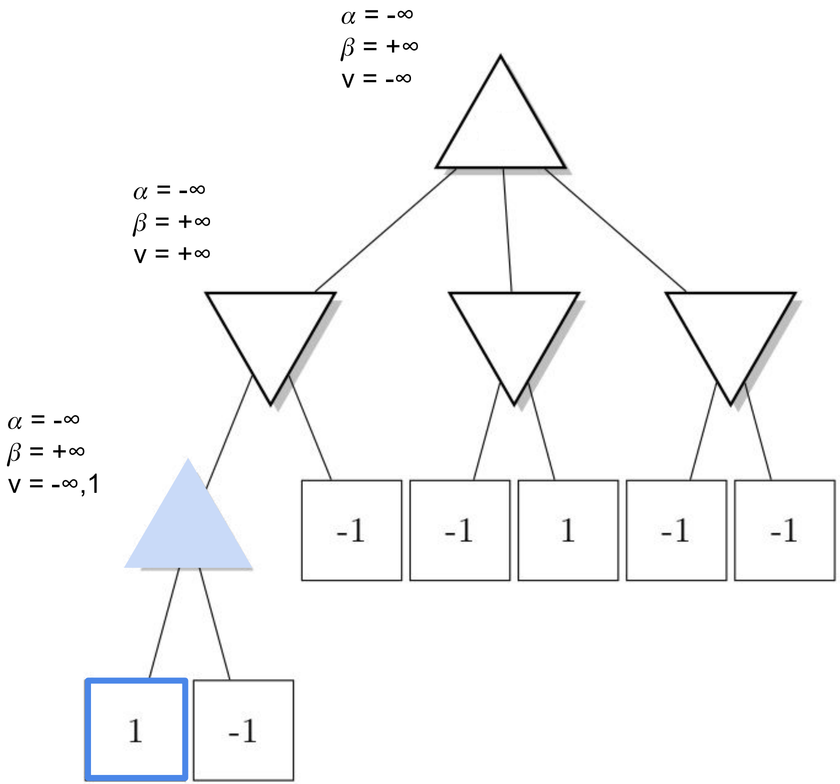
α-β Pruning



α-β Pruning

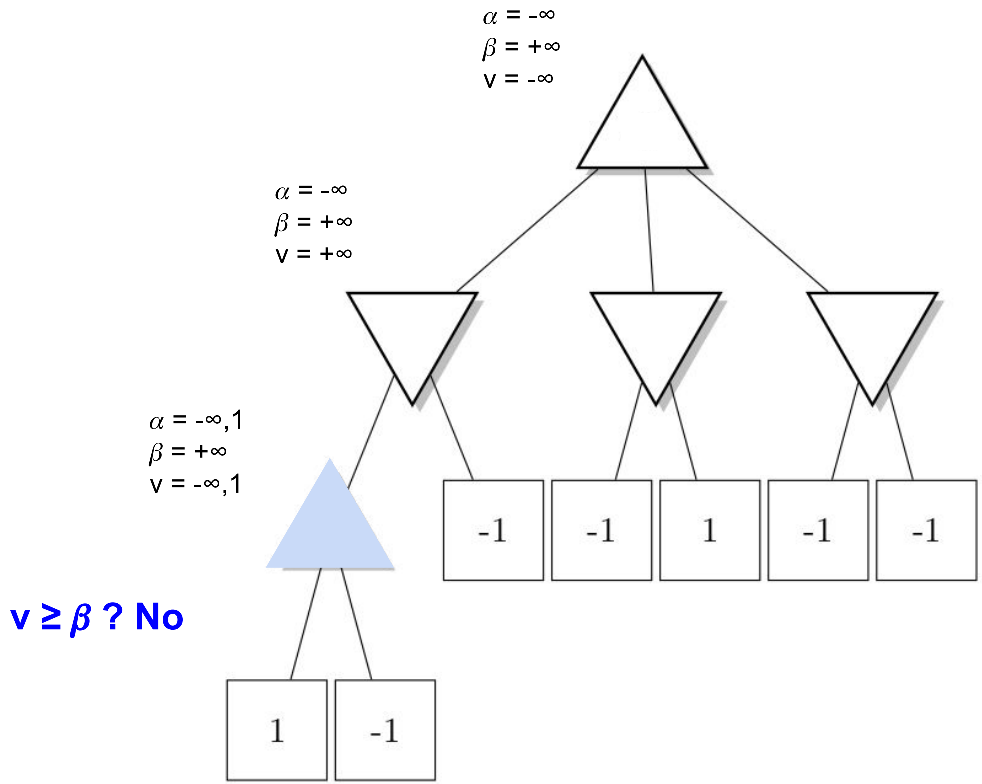


α-β Pruning

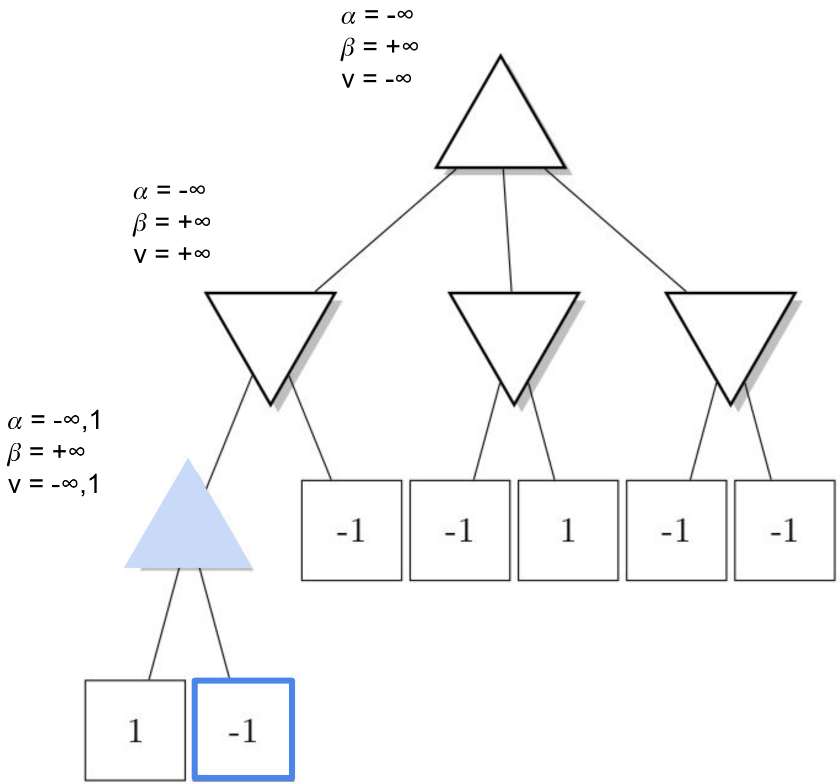


v ≥ β ?

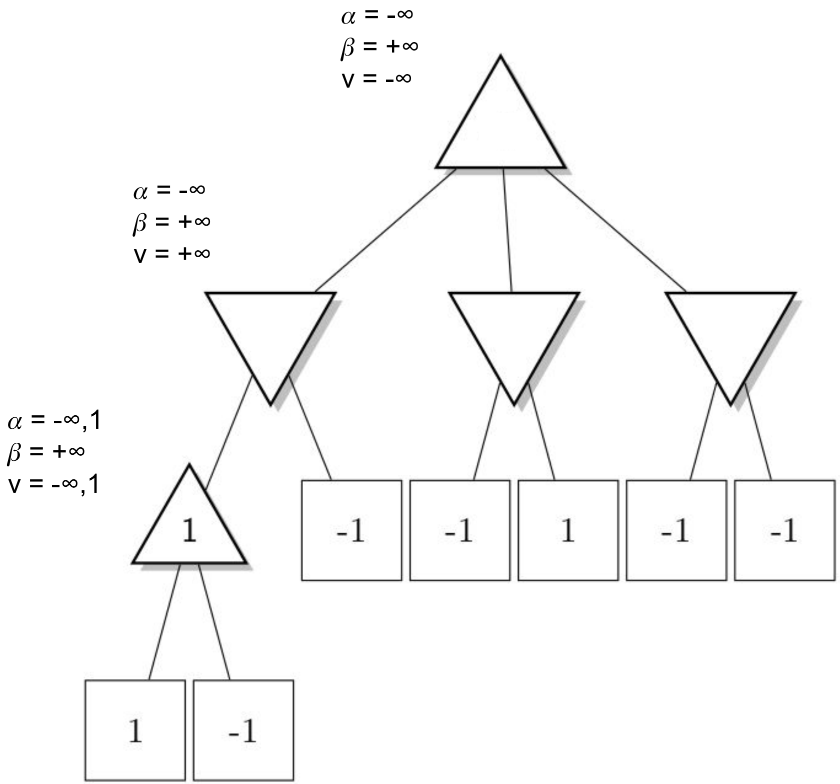
α-β Pruning



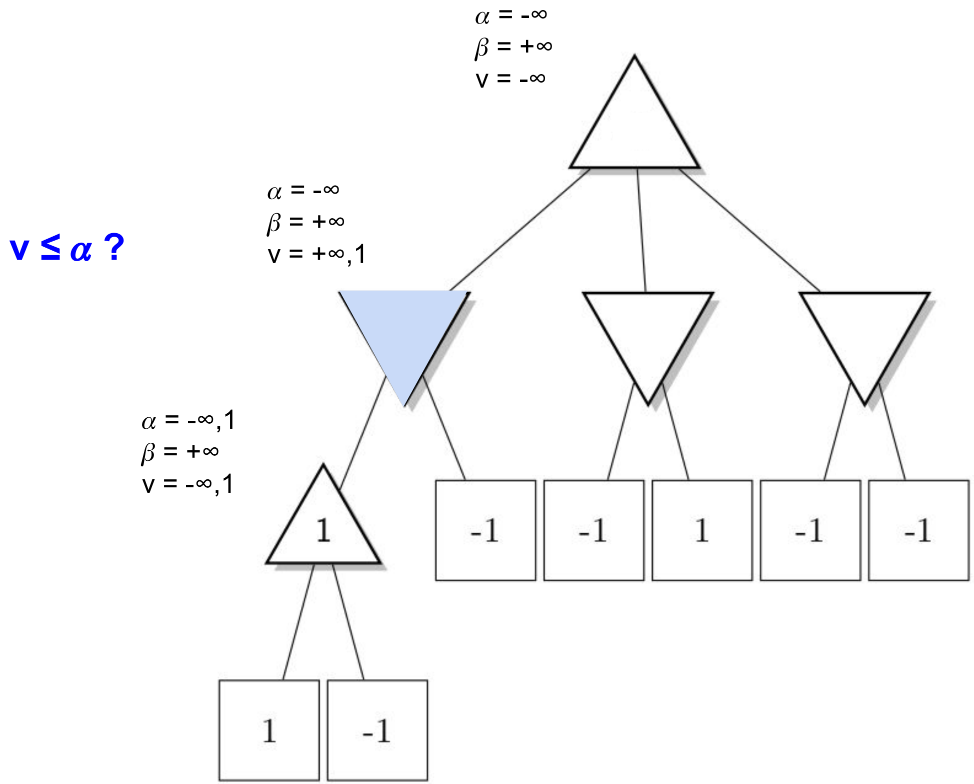
α-β Pruning



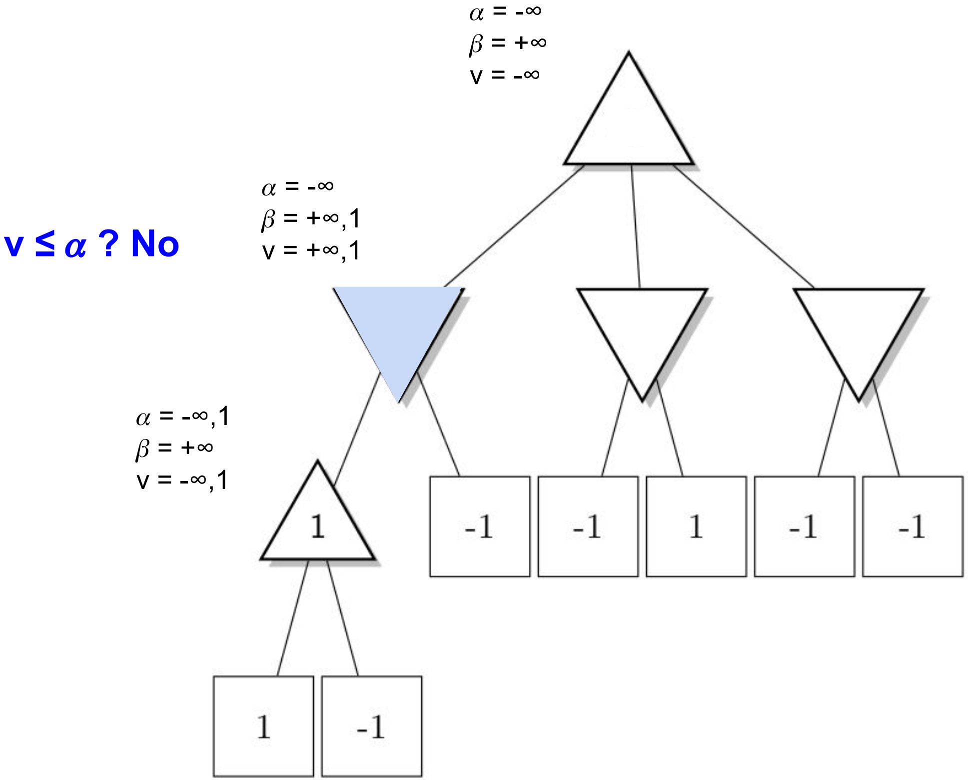
α-β Pruning



α-β Pruning

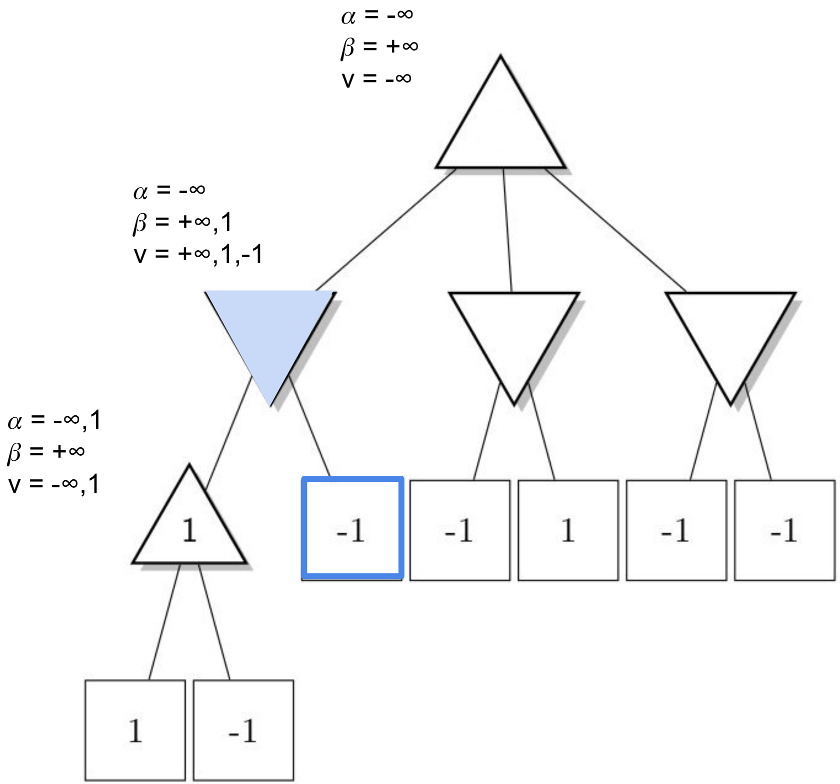


α-β Pruning

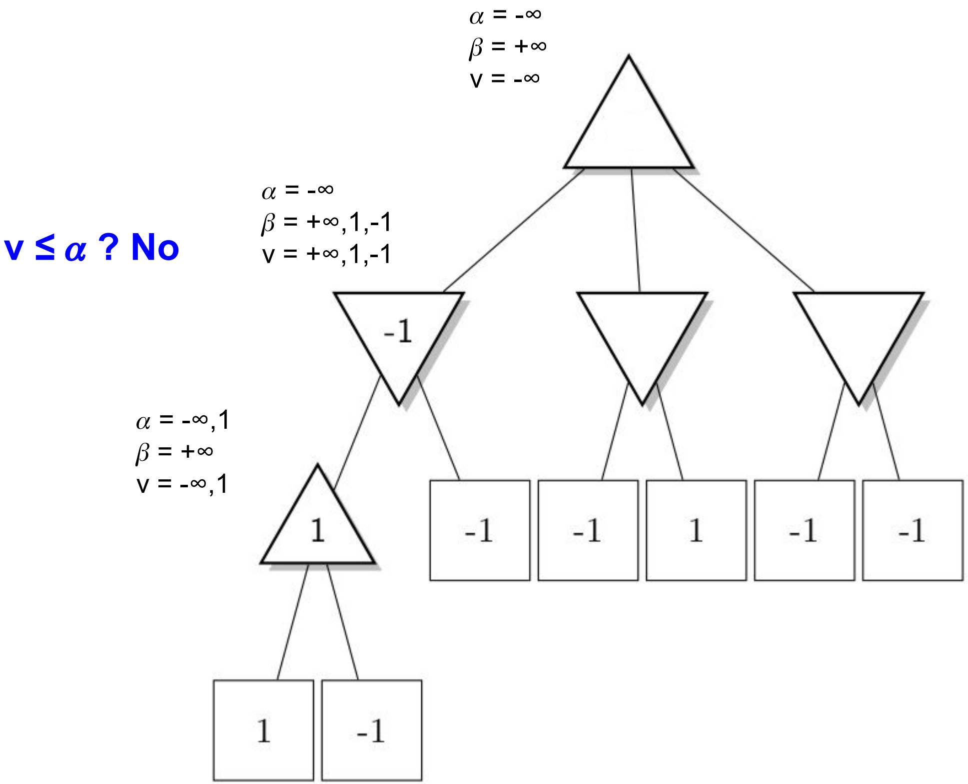


α-β Pruning

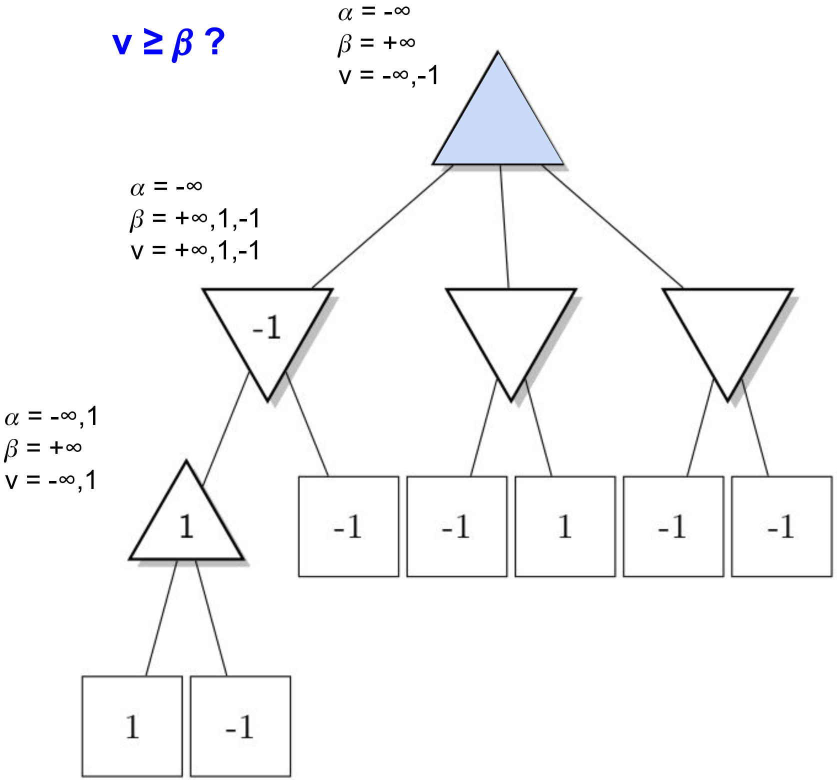
v ≤ α ?



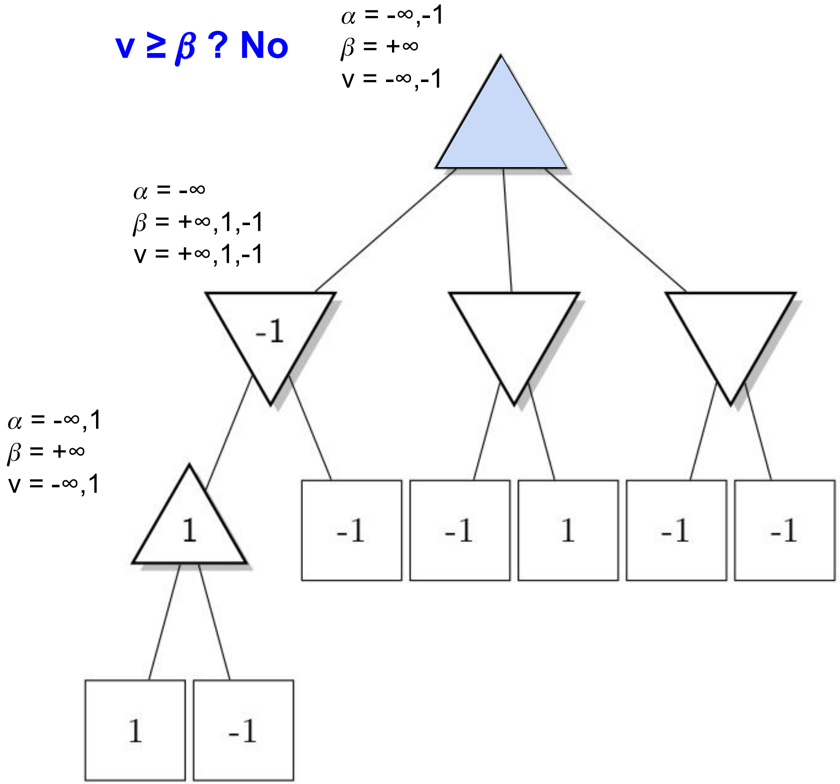
α-β Pruning



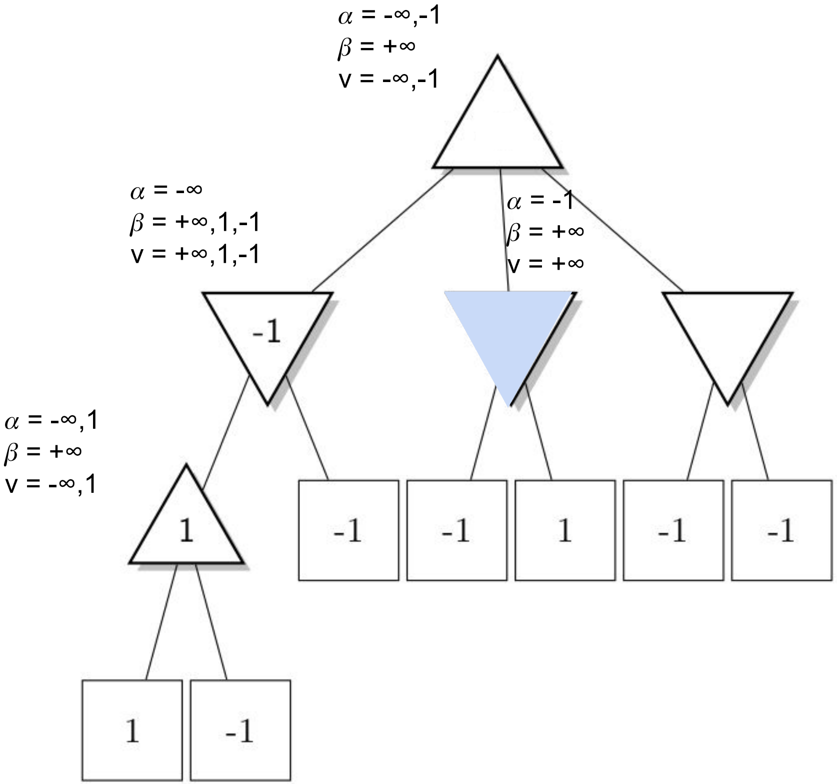
α-β Pruning



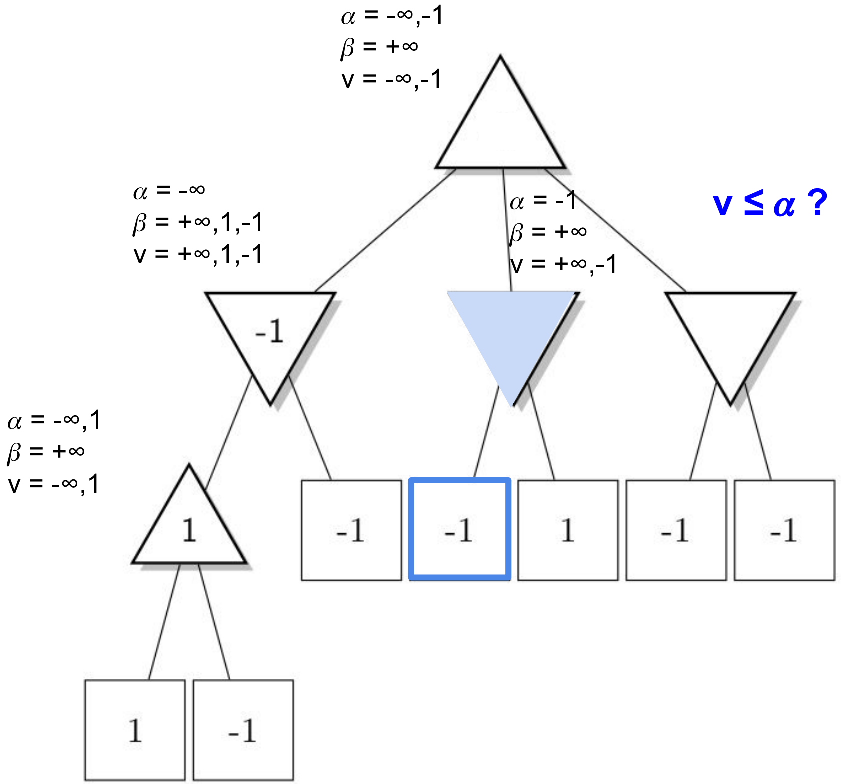
α-β Pruning



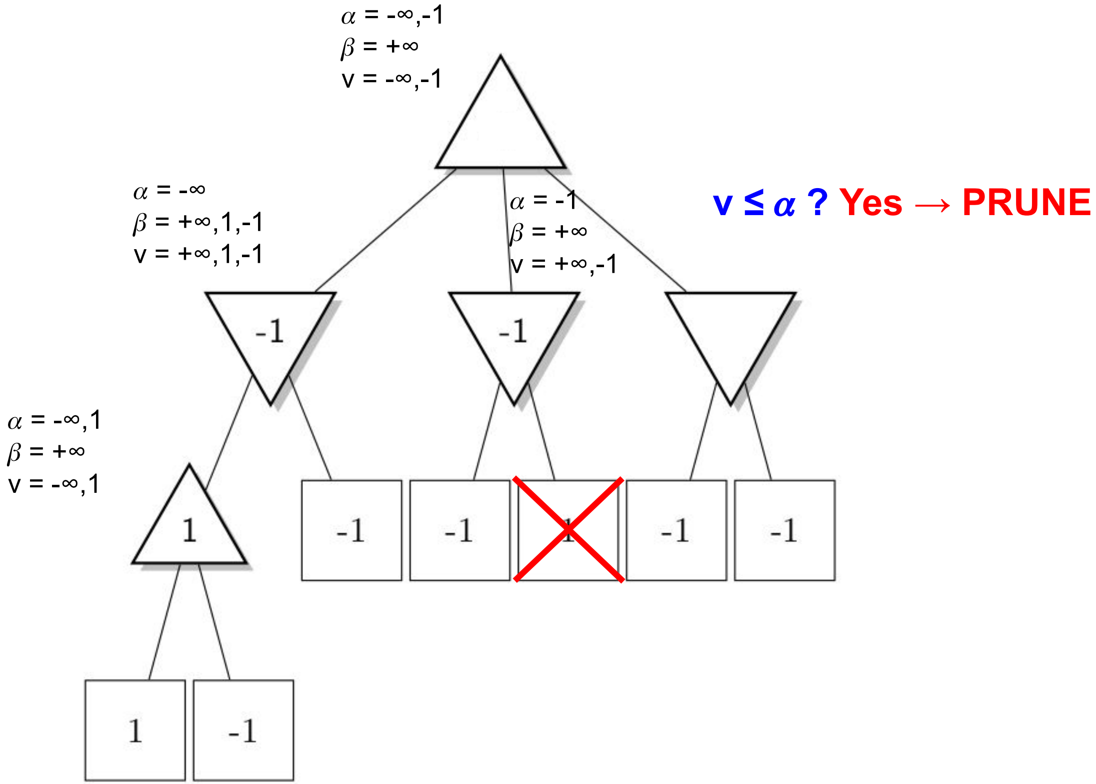
α-β Pruning



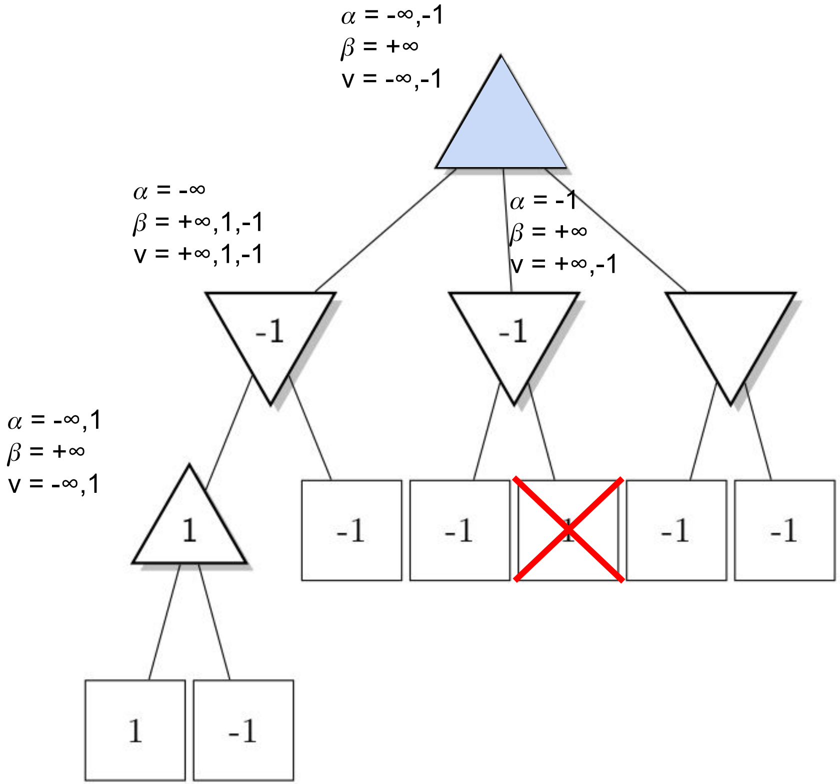
α-β Pruning



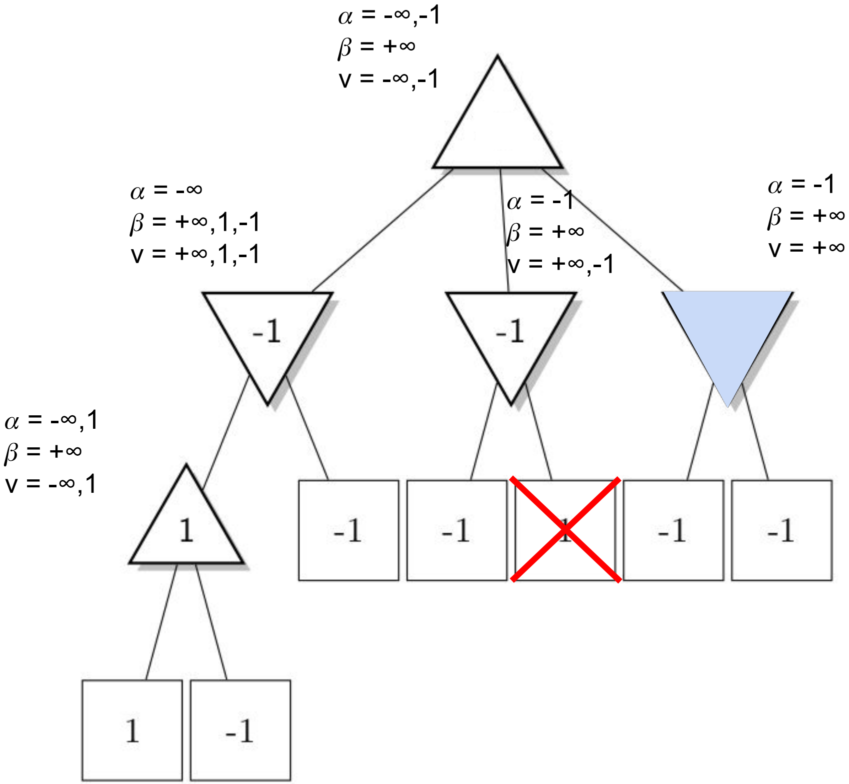
α-β Pruning



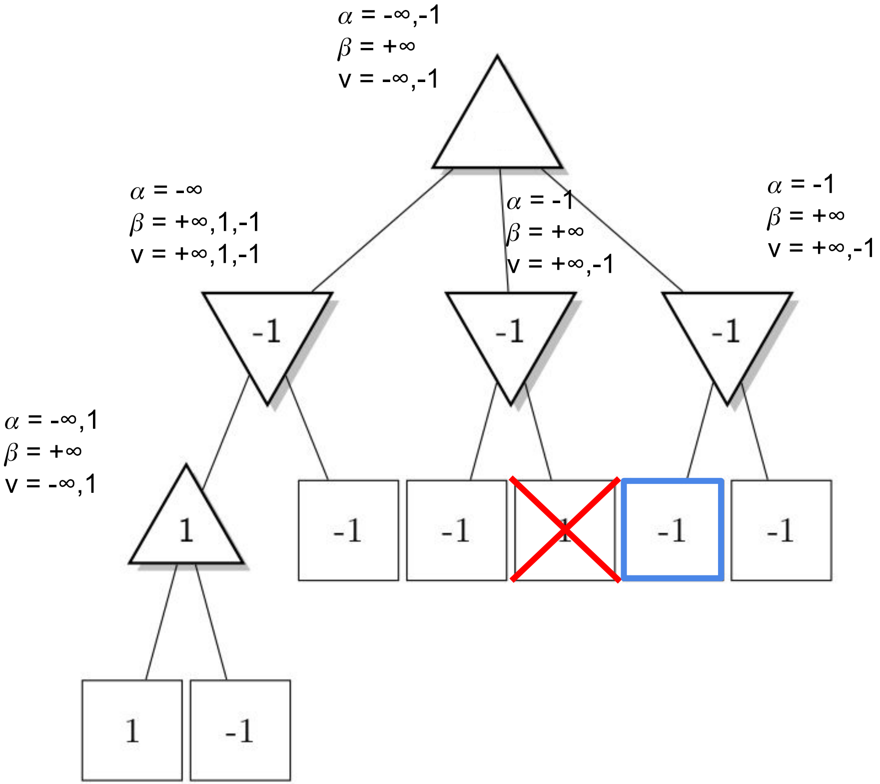
α-β Pruning



α-β Pruning

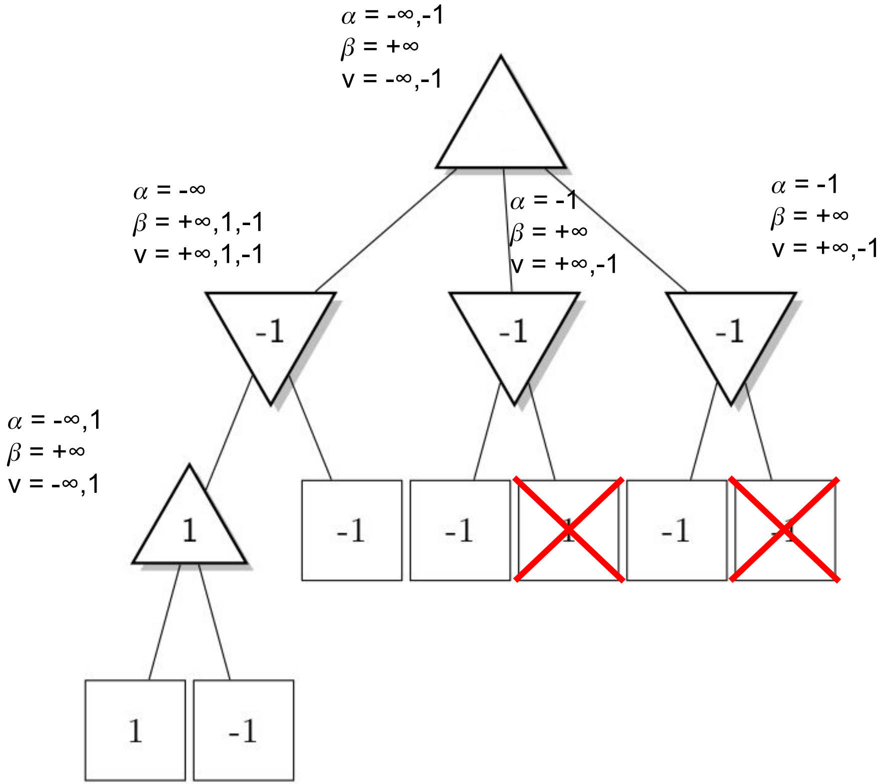


α-β Pruning



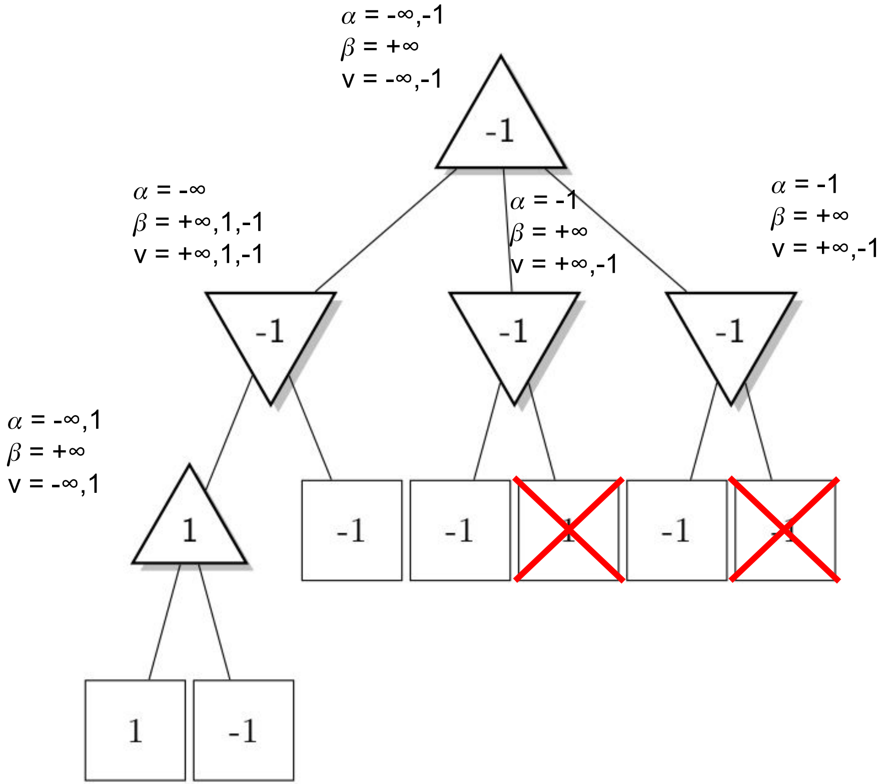
v ≤ α ?

α-β Pruning



v ≤ α ? Yes → PRUNE

α-β Pruning



Reinforcement Learning Exercise

Consider the following sequential decision making-problem. An agent in a 3×3 grid can move in the four directions or stay still, provided that it does not crush against a border. Whenever performing a valid action, the agent reaches deterministically to the corresponding cell. The interaction starts in the lower left cell (blue) and the upper right cell (green) is a terminal state. The immediate reward is represented in the following grid:

0	0	2
-1	-10	0
0	-1	0

Reinforcement Learning Exercise

1. Formalize the problem as a Markov decision process (MDP);

2. For which values of the discount factor γ ∈ [0, 1] the optimal policy consists in staying in the initial state forever?

3. Simulate the execution of Q-learning, starting with a Q-table initialized with the immediate reward, supposing to have observed the following trajectories:

$$(0,0) \xrightarrow{\rightarrow} (1,0) \xrightarrow{\uparrow} (1,1) \xrightarrow{\rightarrow} (2,1) \xrightarrow{\uparrow} (2,2)$$

$$(2,1) \xrightarrow{\downarrow} (2,0) \xrightarrow{\uparrow} (2,1)$$

$$(1,1) \xrightarrow{\downarrow} (1,0) \xrightarrow{\rightarrow} (2,0)$$

$$(0,0) \xrightarrow{\rightarrow} (1,0) \xrightarrow{\uparrow} (1,1)$$

Use discount factor γ = 0.9 and learning rate α = 1.

4. Say which is the greedy policy once completed the updates of the Q-table.

Remember to answer to the feedback form!

https://bit.ly/3C3fs5e