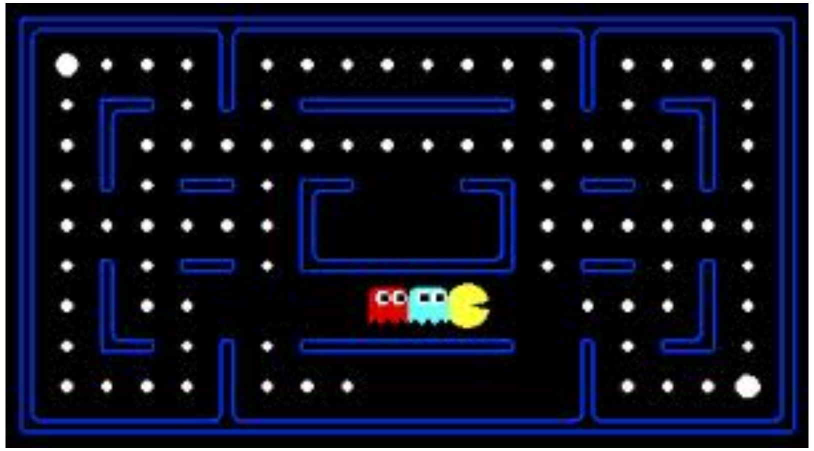
Games: game evaluation



Policies

Deterministic policies: $\pi_p(s) \in \text{Actions}(s)$

action that player p takes in state s

Stochastic policies $\pi_p(s, a) \in [0, 1]$:

probability of player p taking action a in state s

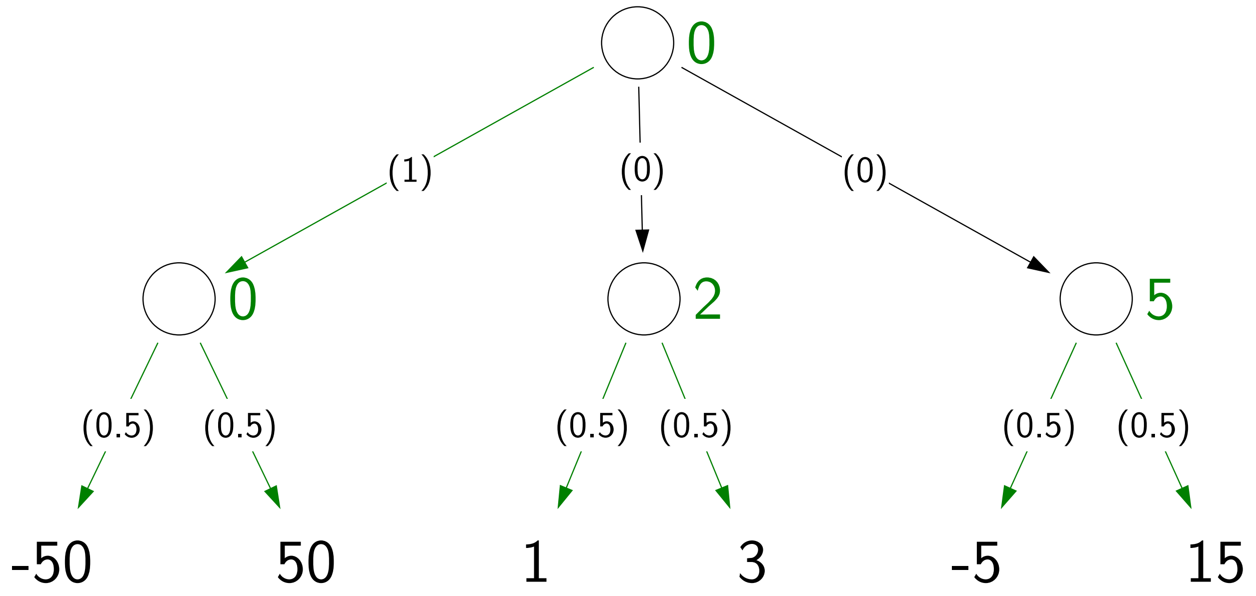
[semi-live solution: `humanPolicy`]

Game evaluation example

Example: game evaluation

$\pi_{\text{agent}}(s) = \text{A}$

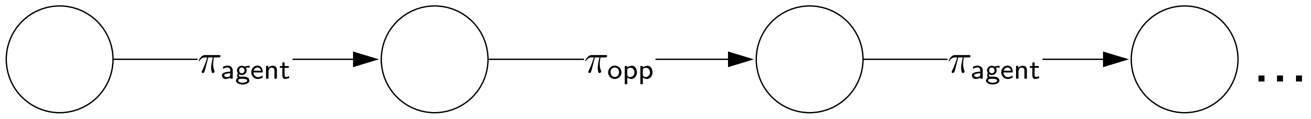
$\pi_{\text{opp}}(s, a) = \frac{1}{2}$ for $a \in \text{Actions}(s)$



$$\boxed{V_{\text{eval}}(s_{\text{start}}) = 0}$$

Game evaluation recurrence

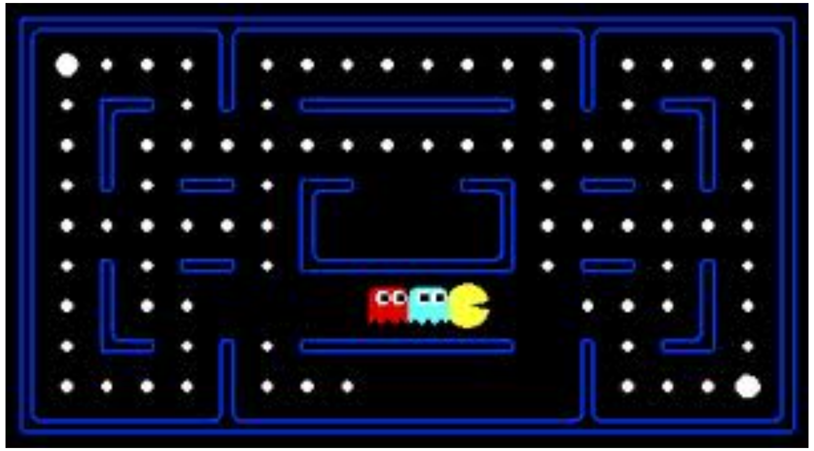
Analogy: recurrence for policy evaluation in MDPs



Value of the game:

$$V_{\text{eval}}(s) = \begin{cases} \text{Utility}(s) & \text{IsEnd}(s) \\ \sum_{a \in \text{Actions}(s)} \pi_{\text{agent}}(s,a) V_{\text{eval}}(\text{Succ}(s,a)) & \text{Player}(s) = \text{agent} \\ \sum_{a \in \text{Actions}(s)} \pi_{\text{opp}}(s,a) V_{\text{eval}}(\text{Succ}(s,a)) & \text{Player}(s) = \text{opp} \end{cases}$$

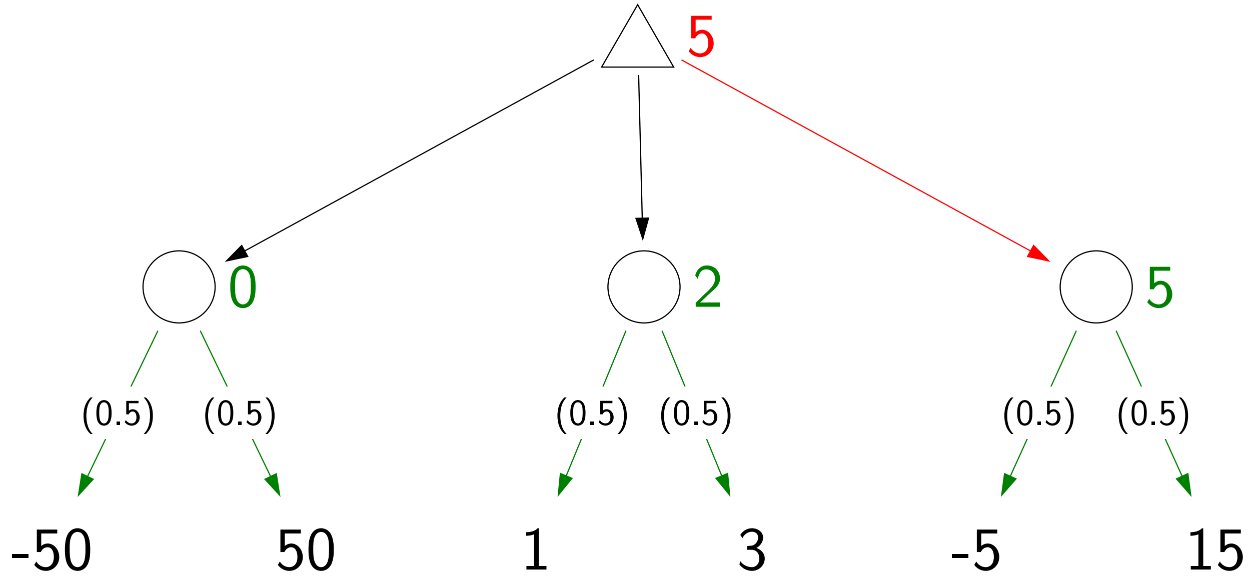
Games: expectimax



Expectimax example

Example: expectimax

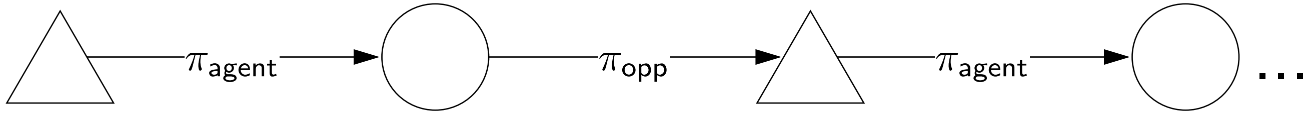
$\pi_{\text{opp}}(s, a) = \frac{1}{2}$ for $a \in \text{Actions}(s)$



$$\boxed{V_{\text{exptmax}}(s_{\text{start}}) = 5}$$

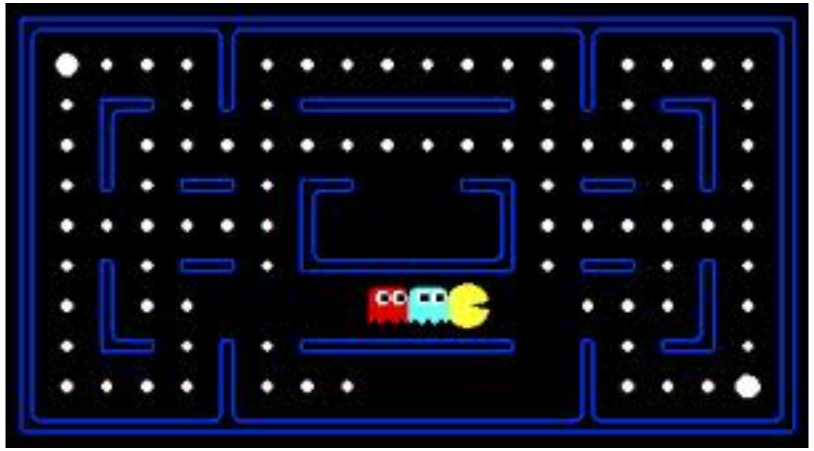
Expectimax recurrence

Analogy: recurrence for value iteration in MDPs



$$V_{\text{exptmax}}(s) = \begin{cases} \text{Utility}(s) & \text{IsEnd}(s) \\ \max_{a \in \text{Actions}(s)} V_{\text{exptmax}}(\text{Succ}(s,a)) & \text{Player}(s) = \text{agent} \\ \sum_{a \in \text{Actions}(s)} \pi_{\text{opp}}(s,a) V_{\text{exptmax}}(\text{Succ}(s,a)) & \text{Player}(s) = \text{opp} \end{cases}$$

Games: minimax

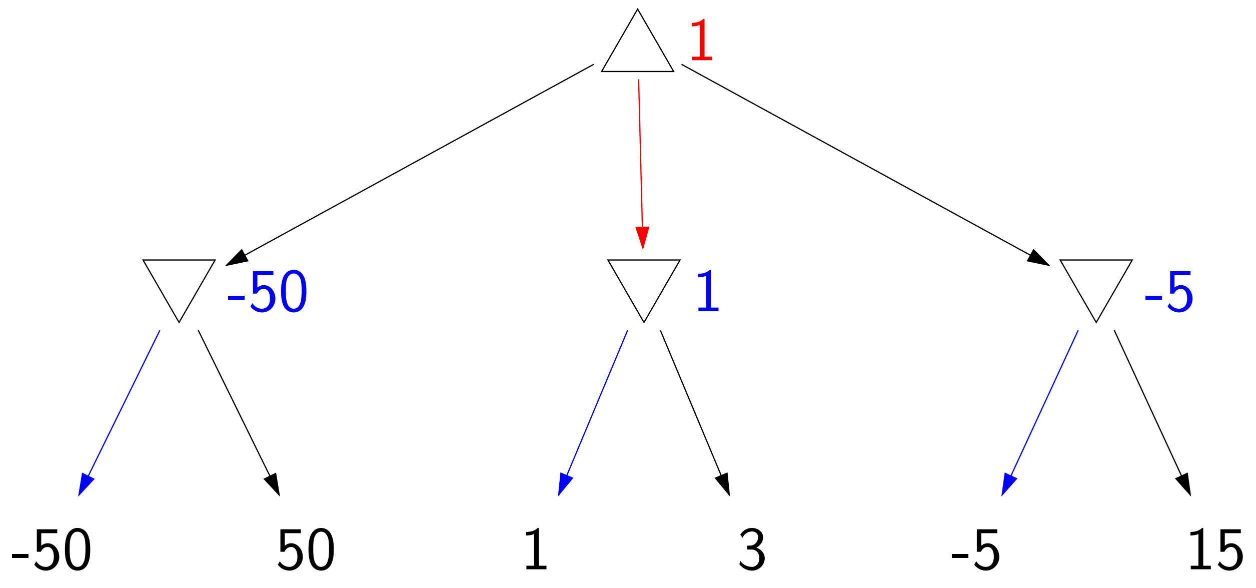


Problem: don't know opponent's policy

Approach: assume the worst case

Minimax example

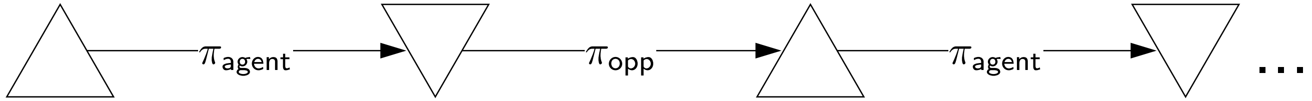
Example: minimax



$$\boxed{V_{\text{minmax}}(s_{\text{start}}) = 1}$$

Minimax recurrence

No analogy in MDPs:

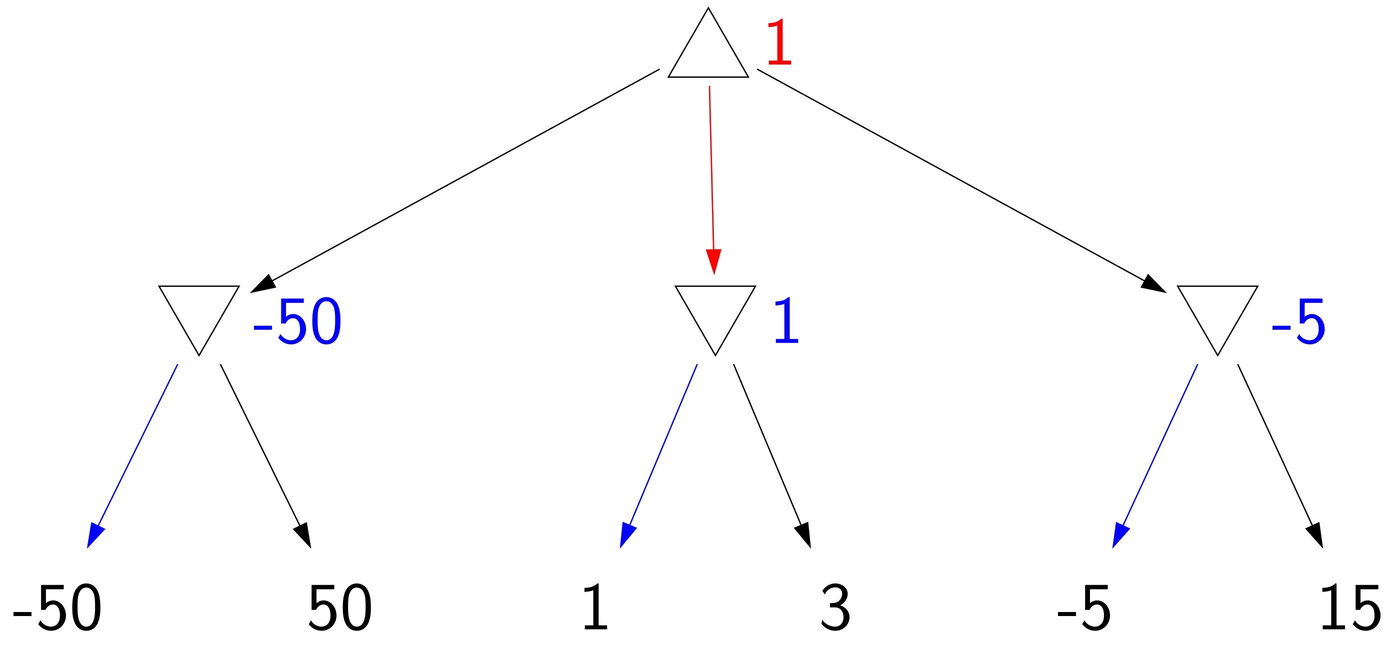


$$V_{\text{minmax}}(s) = \begin{cases} \text{Utility}(s) & \text{IsEnd}(s) \\ \max_{a \in \text{Actions}(s)} V_{\text{minmax}}(\text{Succ}(s,a)) & \text{Player}(s) = \text{agent} \\ \min_{a \in \text{Actions}(s)} V_{\text{minmax}}(\text{Succ}(s,a)) & \text{Player}(s) = \text{opp} \end{cases}$$

Extracting minimax policies

$$\pi_{\max}(s) = \arg\max_{a \in \text{Actions}(s)} V_{\text{minmax}}(\text{Succ}(s, a))$$

$$\pi_{\min}(s) = \arg\min_{a \in \text{Actions}(s)} V_{\text{minmax}}(\text{Succ}(s, a))$$



The halving game

Problem: halving game

Start with a number N.

Players take turns either decrementing N or replacing it with $\lfloor \frac{N}{2} \rfloor$.

The player that is left with 0 wins.

[semi-live solution: `minimaxPolicy`]

Face off

Recurrences produces policies:

$$V_{\text{exptmax}} \Rightarrow \pi_{\text{exptmax(7)}}, \pi_7 \text{ (some opponent)}$$
$$V_{\text{minmax}} \Rightarrow \pi_{\max}, \pi_{\min}$$

Play policies against each other:

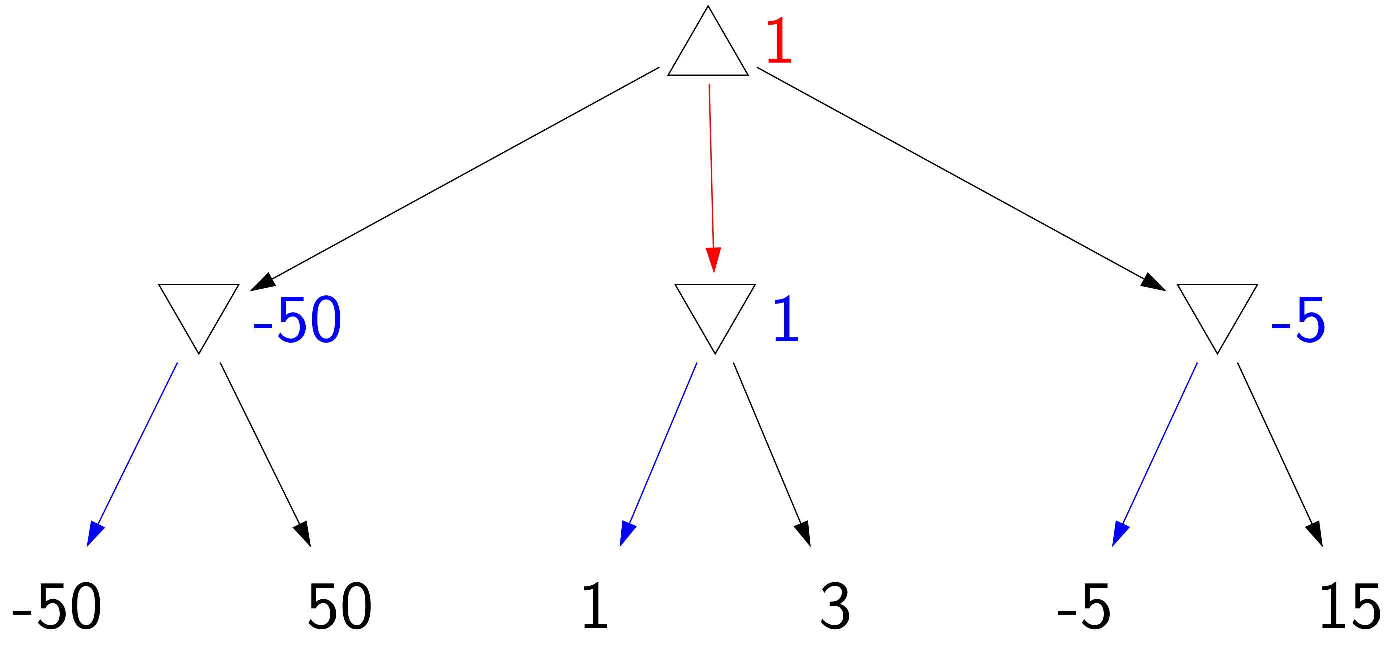
	$\pi_{\min}$	π_7
$\pi_{\max}$	$V(\pi_{\max}, \pi_{\min})$	$V(\pi_{\max}, \pi_7)$
$\pi_{\text{exptmax(7)}}$	$V(\pi_{\text{exptmax(7)}}, \pi_{\min})$	$V(\pi_{\text{exptmax(7)}}, \pi_7)$

What's the relationship between these values?

Minimax property 1

Proposition: best against minimax opponent

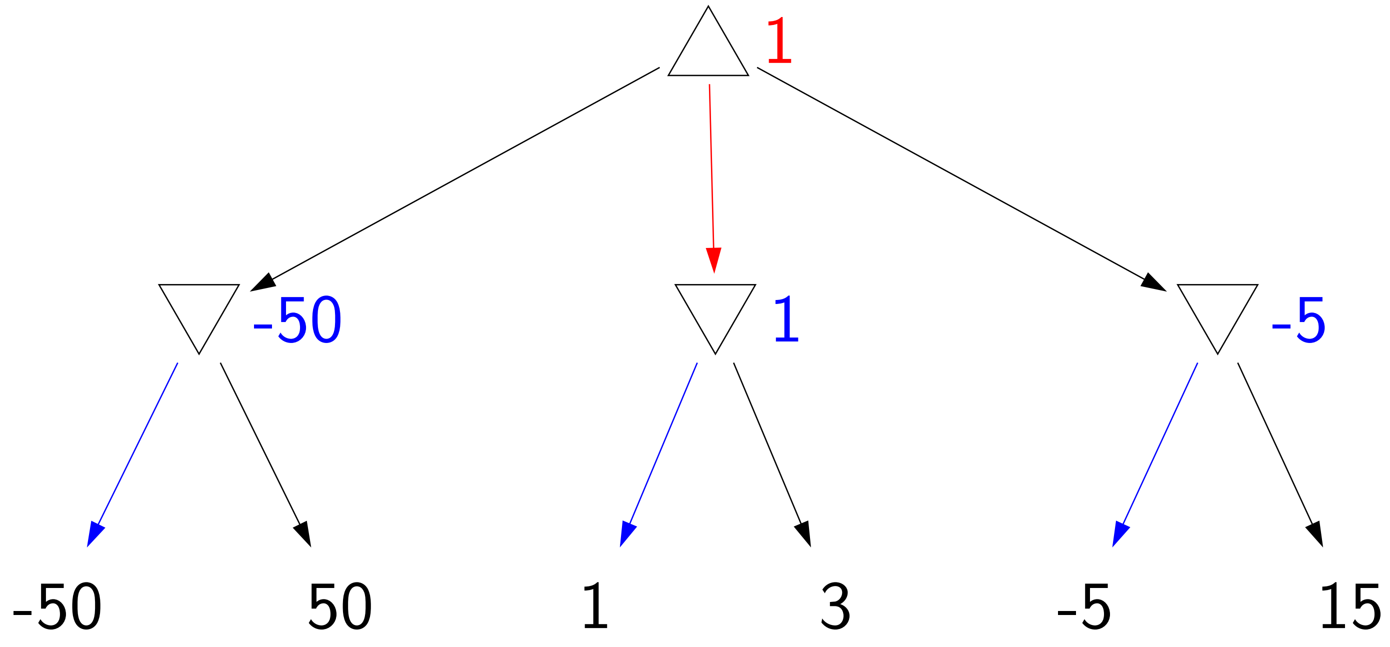
$V(\pi_{\max}, \pi_{\min}) \geq V(\pi_{\text{agent}}, \pi_{\min})$ for all π_{agent}



Minimax property 2

Proposition: lower bound against any opponent

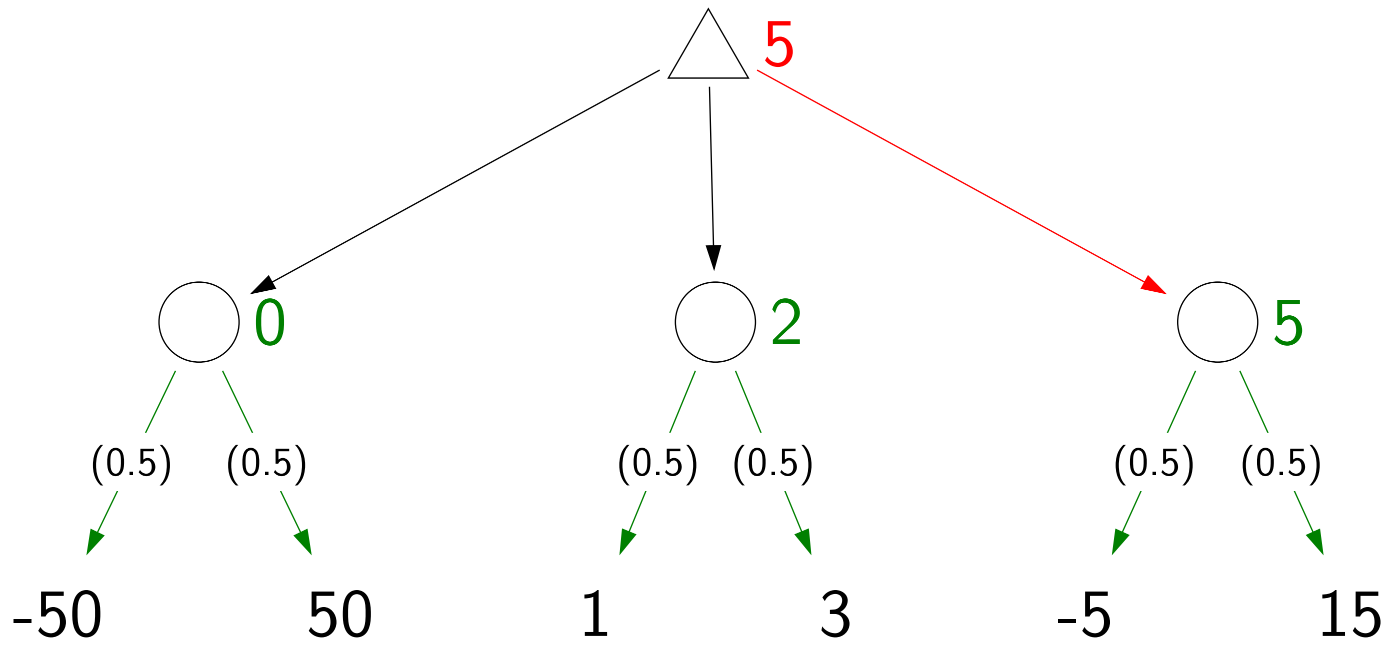
$V(\pi_{\max}, \pi_{\min}) \leq V(\pi_{\max}, \pi_{\text{opp}})$ for all π_{opp}



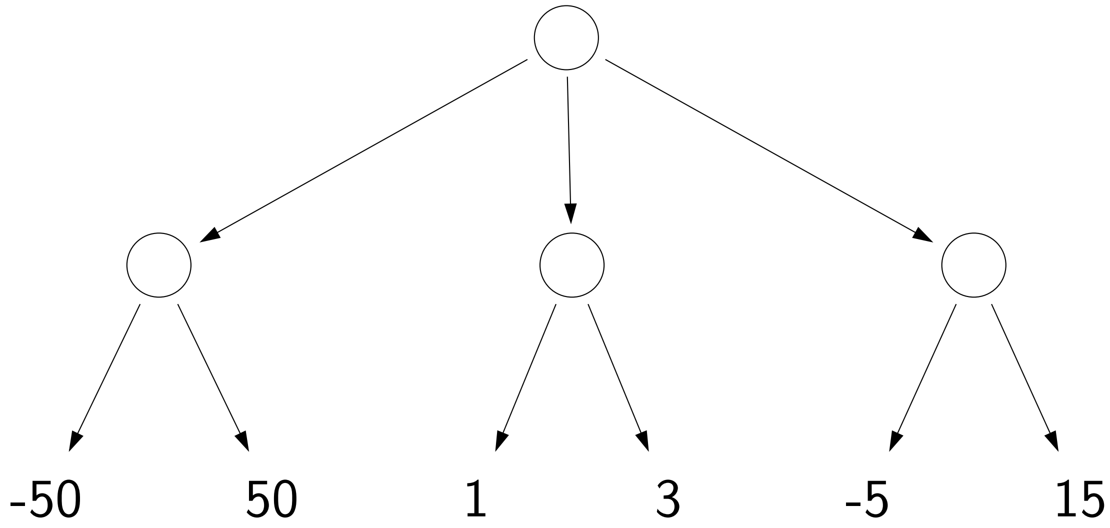
Minimax property 3

Proposition: not optimal if opponent is known

$V(\pi_{\max}, \pi_7) \leq V(\pi_{\text{exptmax(7)}}, \pi_7)$ for opponent π_7



Relationship between game values



	$\pi_{\min}$		π_7
$\pi_{\max}$	$V(\pi_{\max}, \pi_{\min})$ 1	$\leq$	$V(\pi_{\max}, \pi_7)$ 2
	$\geq$		$\leq$
$\pi_{\text{exptmax(7)}}$	$V(\pi_{\text{exptmax(7)}}, \pi_{\min})$ -5		$V(\pi_{\text{exptmax(7)}}, \pi_7)$ 5