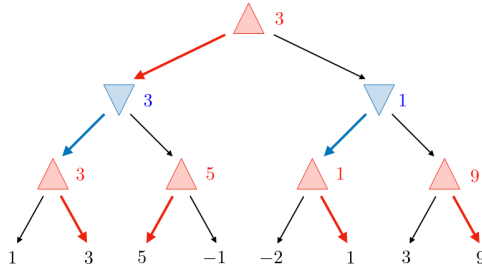


□ **Minimax** – Le but des politiques minimax est de trouver une politique optimale contre un adversaire que l'on assume effectuer toutes les pires actions, i.e. toutes celles qui minimisent l'utilité de l'agent. La valeur correspondante est calculée par :

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

Remarque : on peut déduire π_{\max} et π_{\min} à partir de la valeur minimax V_{minimax} .



□ **Propriétés de minimax** – En notant V la fonction de valeur, il y a 3 propriétés sur minimax qu'il faut avoir à l'esprit :

- *Propriété 1* : si l'agent changeait sa politique en un quelconque π_{agent} , alors il ne s'en sortirait pas mieux.

$$\forall \pi_{\text{agent}}, \quad V(\pi_{\max}, \pi_{\min}) \geq V(\pi_{\text{agent}}, \pi_{\min})$$

- *Propriété 2* : si son adversaire change sa politique de π_{\min} à π_{opp} , alors il ne s'en sortira pas mieux.

$$\forall \pi_{\text{opp}}, \quad V(\pi_{\max}, \pi_{\min}) \leq V(\pi_{\max}, \pi_{\text{opp}})$$

- *Propriété 3* : si l'on sait que son adversaire ne joue pas les pires actions possibles, alors la politique minimax peut ne pas être optimale pour l'agent.

$$\forall \pi, \quad V(\pi_{\max}, \pi) \leq V(\pi_{\text{exptmax}}, \pi)$$

À la fin, on a la relation suivante :

$$V(\pi_{\text{exptmax}}, \pi_{\min}) \leq V(\pi_{\max}, \pi_{\min}) \leq V(\pi_{\max}, \pi) \leq V(\pi_{\text{exptmax}}, \pi)$$