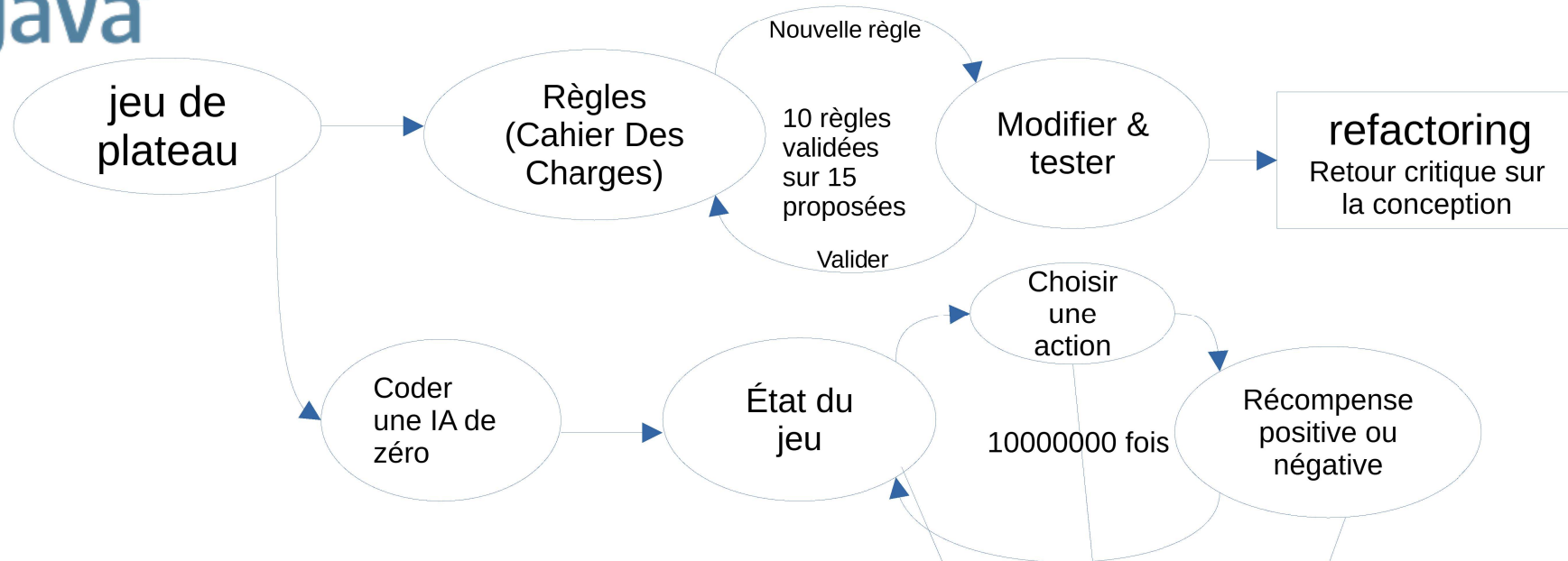


Prendre des risques

Savoir préférer la qualité à la quantité



Q-learning

$$Q[s, a] := (1 - \alpha)Q[s, a] + \alpha \left(r + \gamma \max_{a'} Q[s', a'] \right)$$

Code état	action jouer	résultat	score IA	score Adversaire
0	10	10	1	0
2	40	-10	2	3
4	20	0	4	4