CS7641: Game Theory

Topics covered:

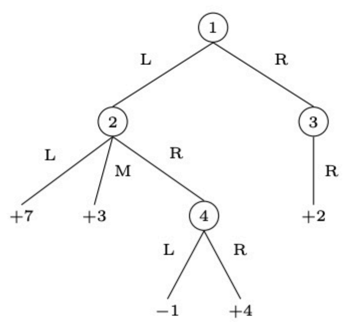
- mathematics of conflict
- single agents → multiple agents
- Economics, politics, sociology biology
- Involvement in AI and ML

- Reinforcement learning: an ML technique that trains software to make decisions to achieve the must optimal results
- zero sum: sum of all rewards is zero and constant
- how we perceive the "economic utility" of the world
- fundamental principle behind game theory is that you are not alone to achieve common and conflicting goals

Games

- MDP has characteristics similar to a basic deterministic game decision tree
- MDP → markov decision process
- MDP says the environment can be described by a particular state, the agent can take in the world based on its state, the model of the world is described by a transition function which describes how the environment responds to the action and actions are rewarded or punished based on their outcome and the resulting state
- MDP has policies which are similar to strategies each

In a simple game, Agent A chooses the row and Agent B chooses the column.



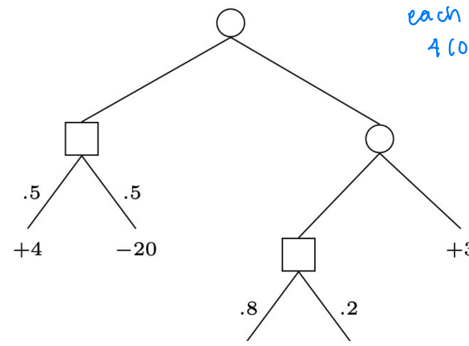
We can actually represent these strategies as a matrix, with each cell being the final reward value for Agent A:

①	④	② L / ③ R	② M / ③ R	② R / ③ R
L	L	7	3	-1
L	R	7	3	4
R	L	2	2	2
R	R	2	2	2

- minmax: two-way process of picking a strategy such that it minimizes tee impact your opponent could have
- in the above game matrix, following the minmax strategy, the "value" of the game is 3 because that is the intersection of the best row for agent A (minimum value in row) and the best column for agent B (row with smallest maximum value) remember optimality: all agents are behaving optimally and thus are trying to maximize their respective reward.
-

> **Theorem 9.1** (Von Neumann Theorem). *In a 2-player, zero-sum, finite game with perfect information, following the minimax strategy is equivalent to following the inverse "maximin" strategy (e.g. Agent A minimizing its reward):*
>
> $$\text{minimax} \equiv \text{maximin}$$
>
> *and there always exists an **optimal** pure strategy for each player.*

Non-determinism and chance



each agent has 50/50 chance
$4(0.5) - 20(0.5) = -8$
(reward on average)

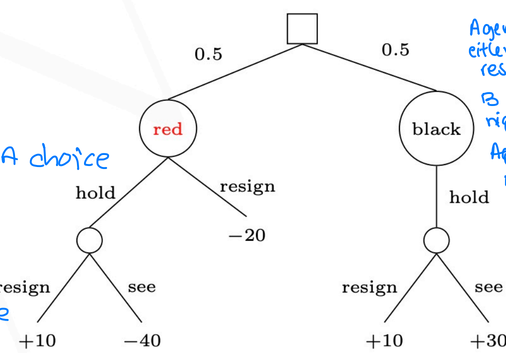
	L	R
L	-8	-8
R	-2	3

Hidden Information

- agent B doesn't know what to do it it doesn't know always what state A is in
- since it is possible for agent B to outperform, the theorem doesn't hold and mnnimax is not equal to maximin
- hidden information means there is no longer a pure consistent strategy that works for both players

Agent A choice

Agent B choice



Agent A would choose either row because both result in −5.
B would choose the right column and 5.
Agent A = −5
Agent B = 5
minimax ≠ maximin

	resign	see
resign	-5	5
hold	10	-5

- A is trying to maximize and B is trying to minimize but both are considering the worst case counter
- A is tying to find maximum minimum and B is trying to find the minimum maximum

FUNDAMENTAL RESULT

In a 2-player, zero-sum deterministic game of perfect information

Minimax ≡ Maximin... and there always exists an optimal pure strategy for each player

- Whether you choose a column or a row first you will get the ' same result (if agent A or B goes first)
- the agent is trying to maximize their reward and is assuming every one else is as well

Mini poker and mixed strategies

- hi