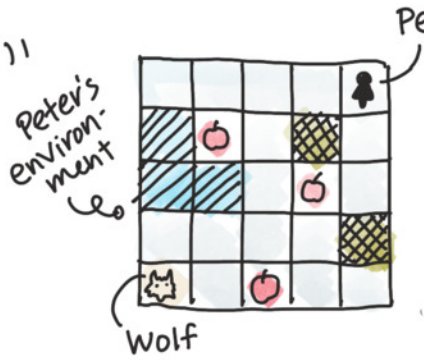
Machine Learning

Reinforcement

Decision-making to take actions in order to maximize cumulative rewards



- Actions: ⇧ ⇩ ⇦ ⇨
- Rewards:
  - +10 apple
  - -10 wolf
  - -5 move

Peter: explores his environment + learns to avoid wolves and collect apples!

Q-Learning ← Algorithm to find the best way to achieve the goal!

Basic: Q-Function

$Q(\text{state}, \text{action}) \rightarrow \text{number}$

current state

At every position at the board, and every action you make, there would be some number

exploring...

Exploration try experimenting on uncharted territory

Exploitation known path

Find a good balance of both!

Bellman Equation ← Dynamic Programming

$$Q(s,a) = r(s,a) + \gamma \max_{a'} Q(s',a')$$

- s: current state
- a: action
- $r(s,a)$: Reward
- γ: discount factor
- $\max_{a'} Q(s',a')$: estimate of optimal future value

@girlie_mac @AzureAdvocates