Reinforcement Learning cont

☆ Key ideas in RL

1) Exploration vs Exploitation

- Exploration = trying new actions to discover better stratigies
- Exploitation = using known action that give high rewards
- Balance is critical (too much Exploration wastes time, too much Exploitation misses better options)

2) Delayed rewards

- Some actions dont give immediate Feedback, but affect long term successes
- Ex: moving pawn in chess might be bad now but will help win later

3) Markov Decision Process (MDP)

- A formal way to describe RL Problems using { states, actions, rewards and transitions

☆ Types of RL

Local optimization in R means choose actions that giv best possibel outcome in immidiet situation (the local step in time) rath that focusing on future.

- RL algorithms genrally fall into 3 families
1) Value-Based → learn how good states/actions are (local optimisation)
2) Policy-Based → learn the stratigy (policy) directly
3) Actor-critic (Hybild) → mix both ideas.

- Q-Learning (value-Based)

- What it dose: learns a Q-function (quality of taking an action in a state
- formula: Q(s,a) = expected total if you take action a in a state s and follow the policy later
- Decision Rule: always pick the action with the highest Q-value
- Example: if i'm at intersection X, turning left has Q=5, going straight has Q=2 → turn left.