Reinforcement Learning Modeling For Human Choice Behavior

About the Team

Nitzan Shahar
(PI)

Gili Katabi
(Value learning in ADHD)

Ido Ben-Artzi
(Outcome-irrelevant learning in OCD)

Maya Marcus
(Exploration in ADHD)

Shir Elchayani
(Adaptive Learning)

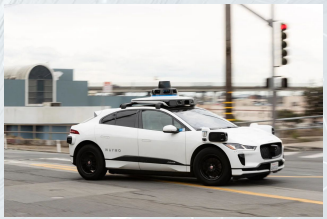
Computational Clinical Psychology Lab

Reinforcement Learning

Learning what actions to take
in an unknown and dynamic environment
with hidden rules

Examples

Autonomous Vehicles:
RL can be used to train self-driving cars to navigate, avoid obstacles, and make decisions in real-time.



Examples

Robotics:
RL can help robots learn tasks such as picking and placing objects, assembly, and even walking.

Examples

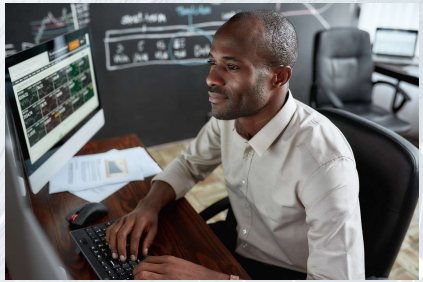
Gaming:
RL models can master complex games like Go, Chess, and video games, often surpassing human performance.

Examples

Healthcare:
RL can be used for personalized treatment planning, optimizing drug dosages, and managing healthcare resources efficiently.

Examples

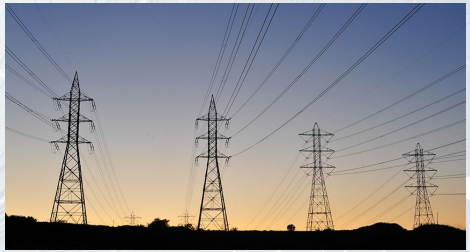
Finance:
RL models can optimize trading strategies, portfolio management, and fraud detection.



Examples

Smart Grids:

RL helps in managing energy distribution, predicting demand, and optimizing the use of renewable energy sources.



Examples

Traffic Management:

RL is used to optimize traffic light timings, reduce congestion, and improve overall traffic flow.

Reinforcement Learning

Learning what actions to take
in an unknown and dynamic environment
with hidden rules

Reinforcement Learning

Can we use RL to better understand the human cognitive system?

Reinforcement Learning

Can we use RL to better understand the human cognitive system?

Reinforcement Learning

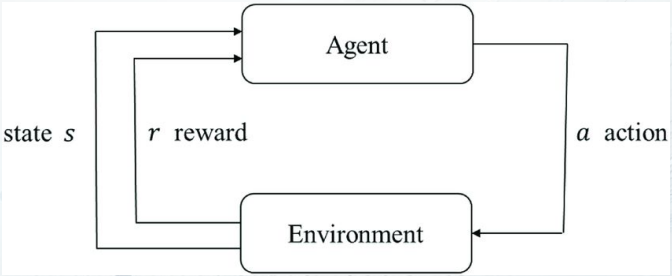
Model-Free

Model-Based

Multi-Agent RL

Deep RL (e.g., DQN)

Learning what actions to take in an unknown, and dynamic, environment with hidden rules



Basic terms in RL

Agent. Entity that interacts with the environment to achieve goals.

State. Representation of the environment at a specific time.

Action. Decision or move made by the agent.

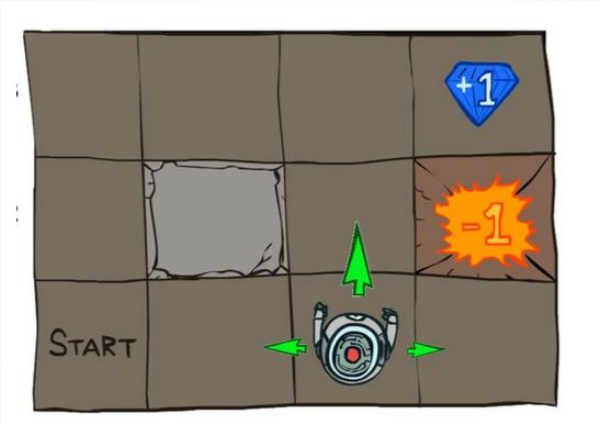
Reward. Feedback received from the environment after an action.

Policy. Strategy used by the agent to decide actions based on states.

Value Function. Expected cumulative reward from a state or state-action pair.

Popular environments

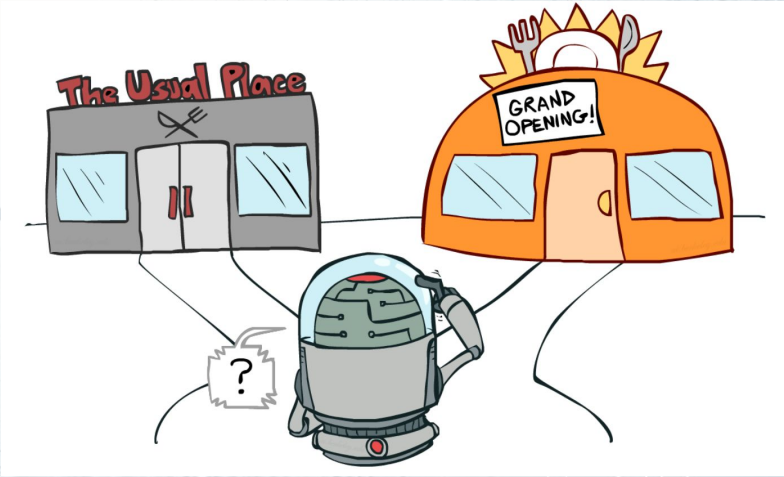
Grid World.
A simple environment where an agent navigates a grid to reach a goal, encountering obstacles and rewards along the way.



Adopted from: https://courses.grainger.illinois.edu/cs440/fa2022/lectures/rl.html

Popular environments

Multi-Armed Bandit.
The agent must choose between multiple actions (arms) with unknown reward distributions to maximize total reward over time.

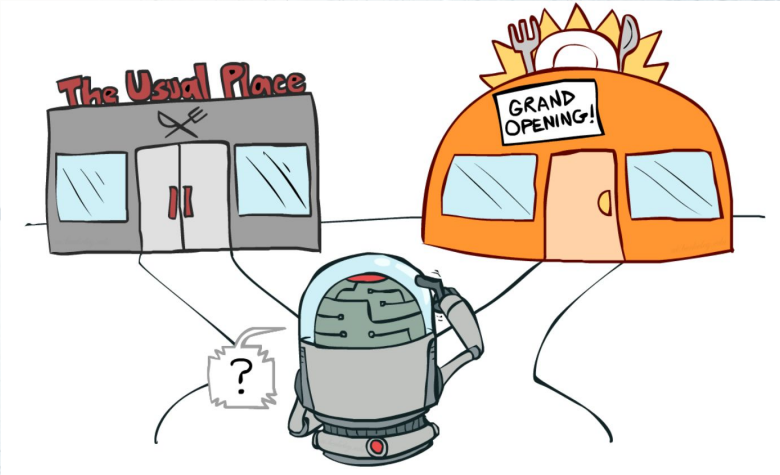


https://images.app.goo.gl/55b3TpTw8ouL5wydA

Popular environments

Multi-Armed Bandit.
The agent must choose between multiple actions (arms) with unknown reward distributions to maximize total reward over time.

- Number of arms
- Rate of change in the true values
- Explore-Exploit
- Contextual bandit (features)
- Regret Minimization
- Prior Knowledge



https://images.app.goo.gl/55b3TpTw8ouL5wydA

The current workshop

Lesson 1: Two-armed bandit task (α, β)

Lesson 2: Sequence learning in a Tree task (α, β, λ)

Lesson 3: Model-based learning in a Two-step task (α, β, λ,ω)