Introduction to Reinforcement Learning

Presenter:

Sepideh Nikookar
https://web.njit.edu/~sn627/

Advisor:

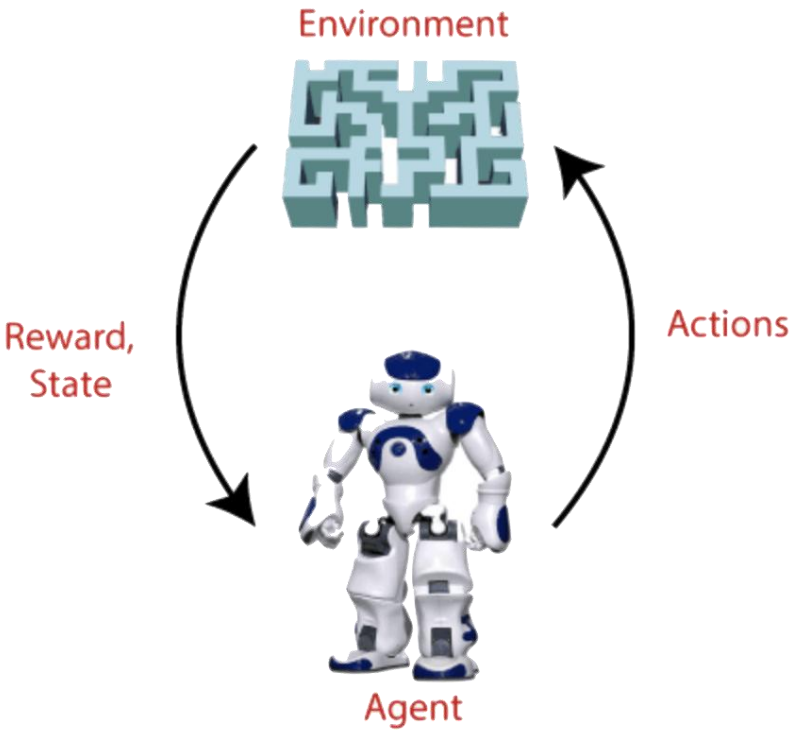
Prof. Senjuti Basu Roy
https://web.njit.edu/~senjutib/

Department of Computer Science
NJIT

Dec 1, 2022

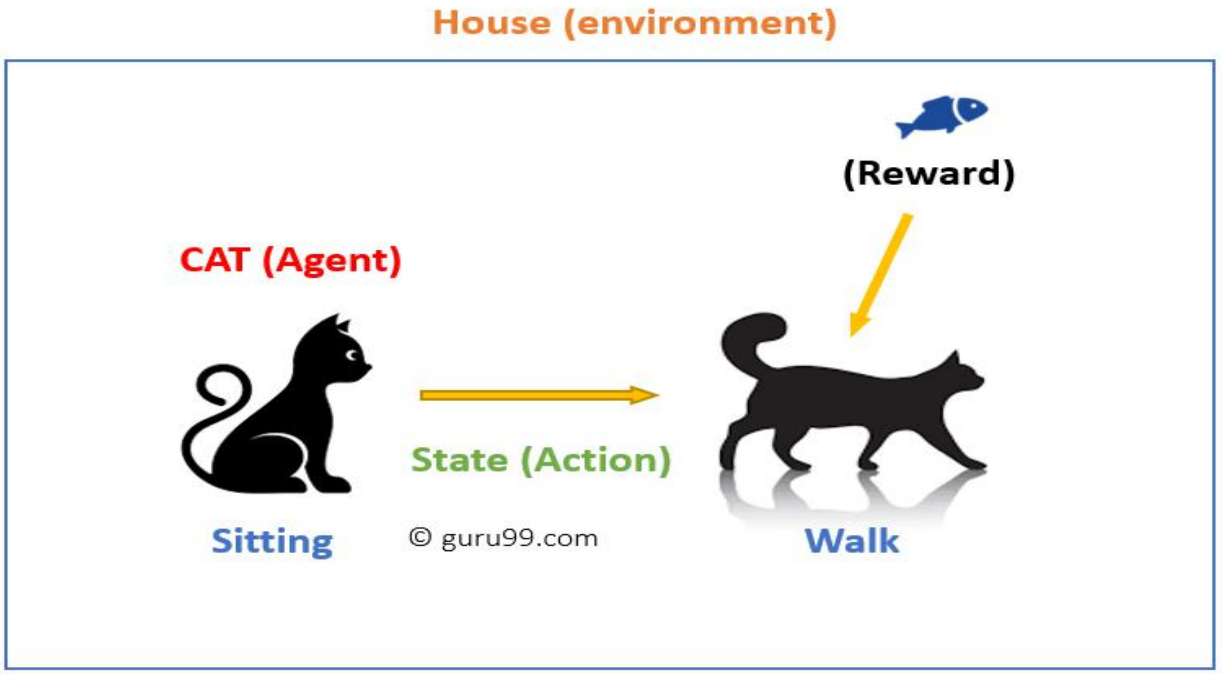
What is Reinforcement Learning?

- Reinforcement learning is a type of machine learning method where an intelligent agent interacts with the environment and learns to act within that.
- For each good action, the agent gets positive feedback, and for each bad action, the agent gets negative feedback or penalty.
- RL solves a specific type of problem where decision making is sequential, and the goal is long-term.
- The agent continues doing following three things:
  - T**ake action,**
  - **change state/remain in the same state**
  - **get feedback**



By doing these actions, he learns and explores the environment.

Reinforcement Learning Example



- An example of a state could be your cat sitting, and you use a specific word for cat to walk.
- Your cat reacts by performing an action transition from one "state" (sitting) to another "state" (walking).
- The reaction of your cat is an action, and the policy is a method of selecting an action given a state in expectation of better outcomes.
- After the transition, cat may get a reward or penalty in return.

Reinforcement Learning Applications

Supervised vs. Unsupervised vs. Reinforcement Learning

Criteria	Supervised ML	Unsupervised ML	Reinforcement ML
Definition	Learns by using labelled data	Trained using unlabelled data without any guidance.	Works on interacting with the environment
Type of problems	Regression and classification	Association and Clustering	Exploitation or Exploration
Algorithms	Linear Regression, Logistic Regression, SVM, KNN etc.	K – Means, C – Means, Apriori	Q – Learning, SARSA
Aim	Calculate outcomes	Discover underlying patterns	Learn a series of action
Application	Risk Evaluation, Forecast Sales	Recommendation System, Anomaly Detection	Self Driving Cars, Gaming, Healthcare

Term Used in Reinforcement Learning

Agent An entity that can perceive/explore the environment and act upon it.

Environment A situation in which an agent is present or surrounded by.

Action Actions are the moves taken by an agent within the environment.

State State is a situation returned by the environment after each action taken by the agent.

Reward A feedback returned to the agent from the environment to evaluate the action of the agent.

Policy Policy is a strategy applied by the agent for the next action based on the current state.

Value It is expected long-term retuned with the discount factor and opposite to the short-term reward.

Elements of Reinforcement Learning

There are four main elements of Reinforcement Learning, which are given below

1. Policy:

A policy can be defined as a way how an agent behaves at a given time. It maps the perceived states of the environment to the actions taken on those states.

The policy-based approach has mainly two types of policy:

- **Deterministic:** The same action is produced by the policy (π) at any state.
- **Stochastic:** In this policy, probability determines the produced action.

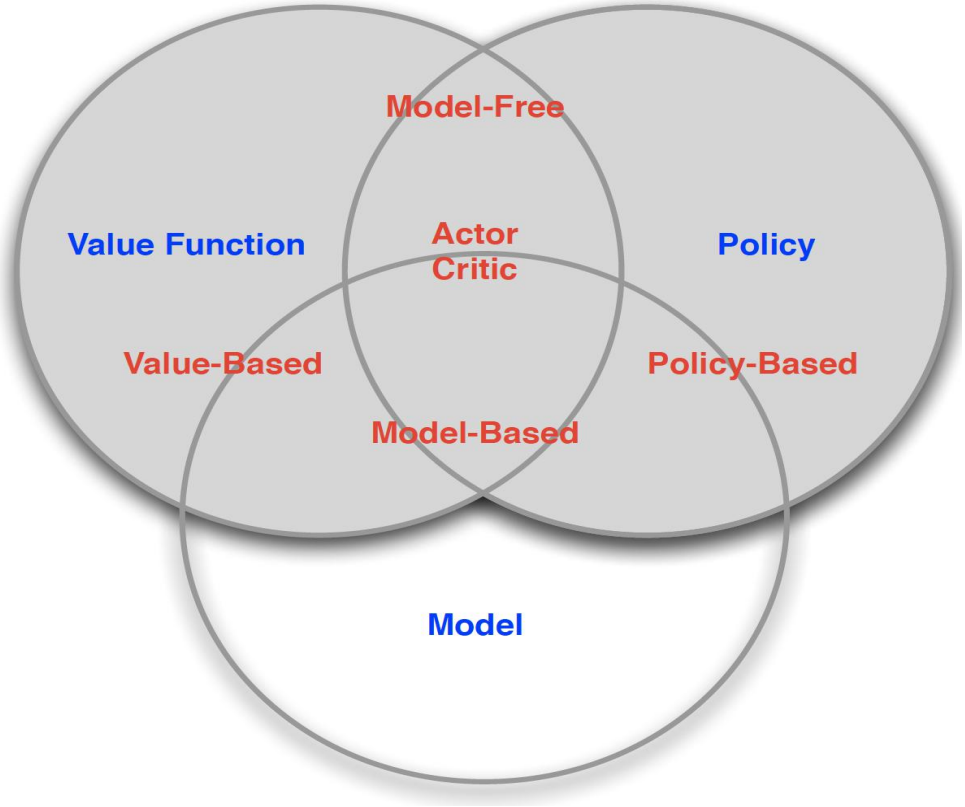
2. Reward Signal: The goal of RL is defined by the reward signal. Reward signals are given according to the good and bad actions taken by the learning agent. The main objective is to maximize the total number of rewards for good actions.

3. Value Function: Gives information about how good the situation and action are and how much reward an agent can expect.

4. Model: Mimics the behavior of the environment.

Reinforcement Learning Categories

- **Value Based**
  - No Policy
  - Value Function
- **Policy Based**
  - Policy
  - No Value Function
- **Actor Critic**
  - Policy
  - Value Function
- **Model Free**
  - Policy and/or Value Function
  - No Model
- **Model Based**
  - Policy and/or Value Function
  - Model



State Representation

We can represent the agent state using the **Markov State** that contains all the required information from the history. The State s_t is Markov state if it follows the given condition:

$$P[s_{t+1} \mid s_t] = \mathrm{P}[\mathrm{s}_{t+1} \mid s_1,\ \mathrm{s}_2,\ \ldots,\ \mathrm{s}_\mathrm{t}]$$

The Markov state follows the **Markov property**, which says that the future is independent of the past and can only be defined with the present.

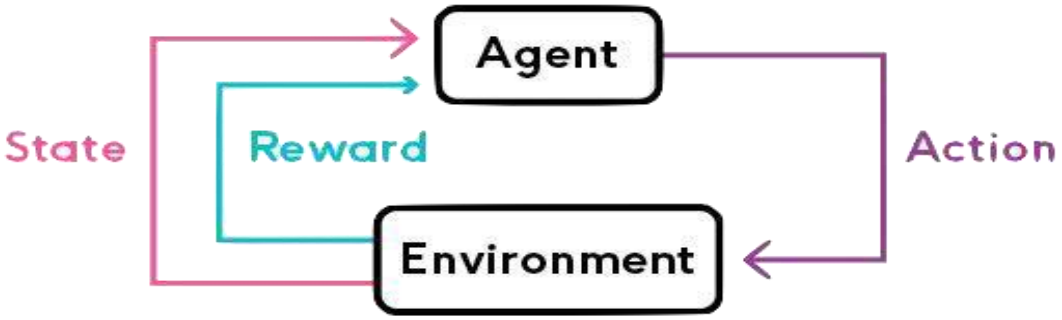
The RL works on fully observable environments, where the agent can observe the environment and act for the new state. The complete process is known as Markov Decision process.

Markov Decision Process

- Markov Decision Process or MDP, is used to **formalize the Reinforcement Learning problems**.

  MDP contains a tuple of four elements $(S,\ A, P_a,\ R_a)$:

  - A set of finite States S
  - A set of finite Actions A
  - Rewards received after transitioning from state s to state s', due to action a.
  - Probability P_a.



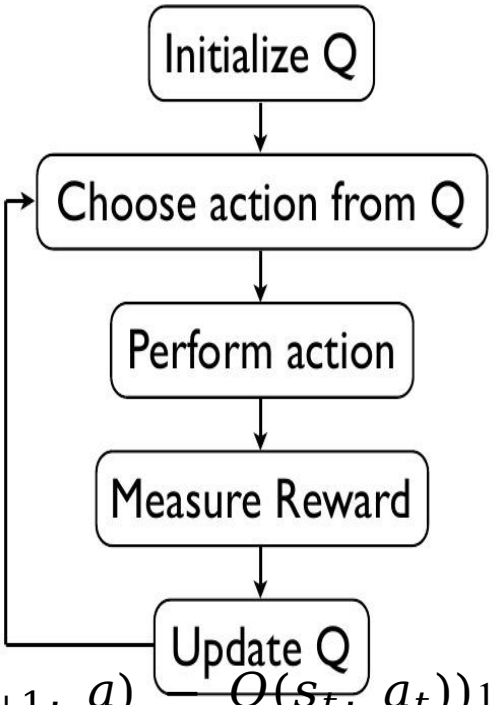
Reinforcement Learning Algorithms

- RL algorithms are mainly used in AI applications and gaming applications. The main used algorithm is:

Q-Learning:

Q-learning is a popular model-free Reinforcement Learning algorithm based on the Bellman equation $V(s) = \max_a(R(s,a) + \gamma V(s'))$.

- The main objective of Q-learning is to learn the policy which can inform the agent that what actions should be taken for maximizing the reward under what circumstances.
- It is an **off-policy RL** that attempts to find the best action to take at a current state.



$$Q_{new}(s_t,\ a_t) = Q(s_t,\ a_t) + \alpha \times (r_t + \gamma \times max_a\ Q(s_{t+1},\ a) - Q(s_t,\ a_t))$$

Please open your Jupyter Notebook for the hands-on experience .

Thank You!