Reinforcement learning

Firstly we will briefly explain what is reinforcement learning let’s imagine the AI as a child and you want to teach the child how to walk for example, so you don’t tell him you should move your leg in this certain way and move your muscles like this, no instead you make him stand up and let him walk and give him a reward every time he takes a step , as such:  
 Reinforcement learning is a [machine learning](https://www.techtarget.com/searchenterpriseai/definition/machine-learning-ML) training method based on rewarding desired behaviors and punishing undesired ones. In general, a reinforcement learning agent -- the entity being trained -- is able to perceive and interpret its environment, take actions and learn through trial and error.

Reinforcement learning is one of several approaches developers use to train machine learning systems. What makes this approach important is that it empowers an agent, whether it's a feature in a video game or a [robot](https://www.techtarget.com/searchenterpriseai/definition/robot) in an industrial setting, to learn to navigate the complexities of the environment it was created for. Over time, through a feedback system that typically includes rewards and punishments, the agent learns from its environment and optimizes its behaviors.

Here are key components of reinforcement learning:

* Agent: The entity or system that learns and makes decisions. In the context of RL, this is typically a computer program or algorithm.
* Environment: The external system with which the agent interacts. It represents the context or scenario in which the agent operates.
* Observation: partial part of the environment which means what actually the agent sees
* State: A representation of the current situation or configuration of the environment. The agent perceives the state before deciding on an action.
* Action: The decision or move that the agent makes in response to the observed state. Actions can have different consequences and affect the subsequent state.
* Reward: The feedback signal that the agent receives from the environment after taking an action. Rewards indicate the immediate benefit or cost of the action.
* Policy: The strategy or mapping from states to actions that the agent uses to make decisions. The goal of RL is often to learn an optimal policy that maximizes the cumulative rewards over time.



How Does Reinforcement Learning Work?

The Reinforcement Learning problem involves an agent exploring an unknown environment to achieve a goal. RL is based on the hypothesis that all goals can be described by the maximization of expected cumulative reward. The agent must learn to sense and perturb the state of the environment using its actions to derive maximal reward. The formal framework for RL borrows from the problem of optimal control of Markov Decision Processes (MDP).

The main elements of an RL system are:

The agent or the learner

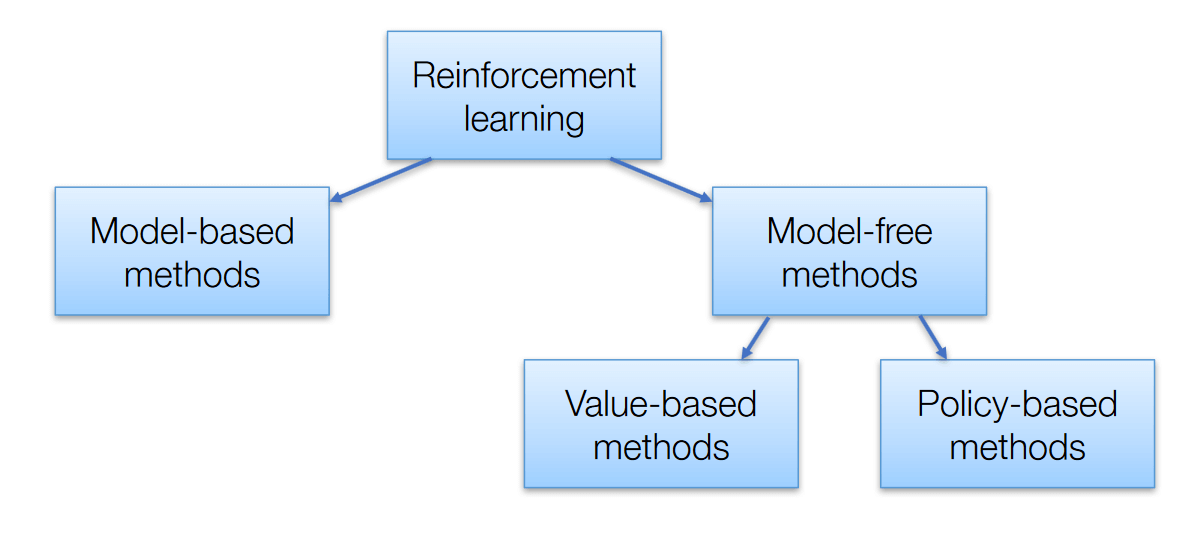
The environment the agent interacts with

The policy that the agent follows to take actions

The reward signal that the agent observes upon taking actions

A useful abstraction of the reward signal is the value function, which faithfully captures the ‘goodness’ of a state. While the reward signal represents the immediate benefit of being in a certain state, the value function captures the cumulative reward that is expected to be collected from that state on, going into the future. The objective of an RL algorithm is to discover the action policy that maximizes the average value that it can extract from every state of the system.

There are 3 approaches to implement reinforcement learning algorithms



Value-Based – The main goal of this method is to maximize a value function. Here, an agent through a policy expects a long-term return of the current states.

Policy-Based – In policy-based, you enable to come up with a strategy that helps to gain maximum rewards in the future through possible actions performed in each state. Two types of policy-based methods are deterministic and stochastic.

Model-Based – In this method, we need to create a virtual model for the agent to help in learning to perform in each specific environment

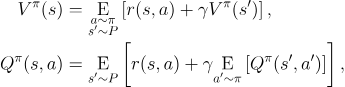
Now comes the importance of the value function

**Value Function:** The value function gives information about how good the situation and action are and how much reward an agent can expect. A reward indicates the **immediate signal for each good and bad action**, whereas a value function specifies **the good state and action for the future**. The value function depends on the reward as, without reward, there could be no value. The goal of estimating values is to achieve more rewards.

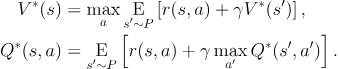
we will use the **Bellman equation**, which is the main concept behind reinforcement learning.  
It is a way of calculating the value functions in dynamic programming or environment that leads to modern reinforcement learning.

The key-elements used in Bellman equations are:

* Action performed by the agent is referred to as "a"
* State occurred by performing the action is "s."
* The reward/feedback obtained for each good and bad action is "R."
* A discount factor is Gamma "γ."
* The value of your starting point is the reward you expect to get from being there, plus the value of wherever you land next.
* The Bellman equations for the on-policy value functions are

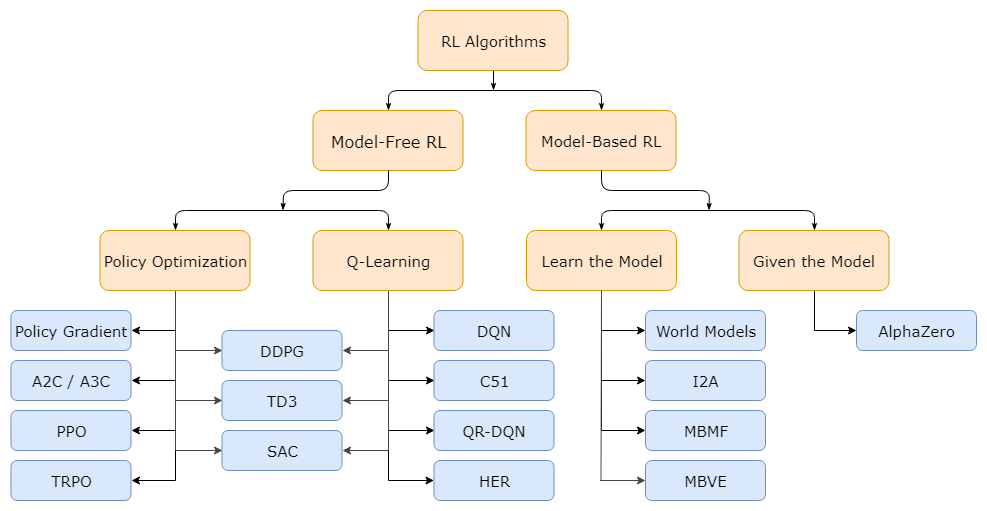


* where  is shorthand for , indicating that the next state  is sampled from the environment’s transition rules;  is shorthand for ; and  is shorthand for .
* The Bellman equations for the optimal value functions are



* The crucial difference between the Bellman equations for the on-policy value functions and the optimal value functions, is the absence or presence of the  over actions. Its inclusion reflects the fact that whenever the agent gets to choose its action, in order to act optimally, it has to pick whichever action leads to the highest value.

After knowing about the Bellman Equation and the optimal Q value now we will understand what are the types of algorithms that can be implemented in the RL concept

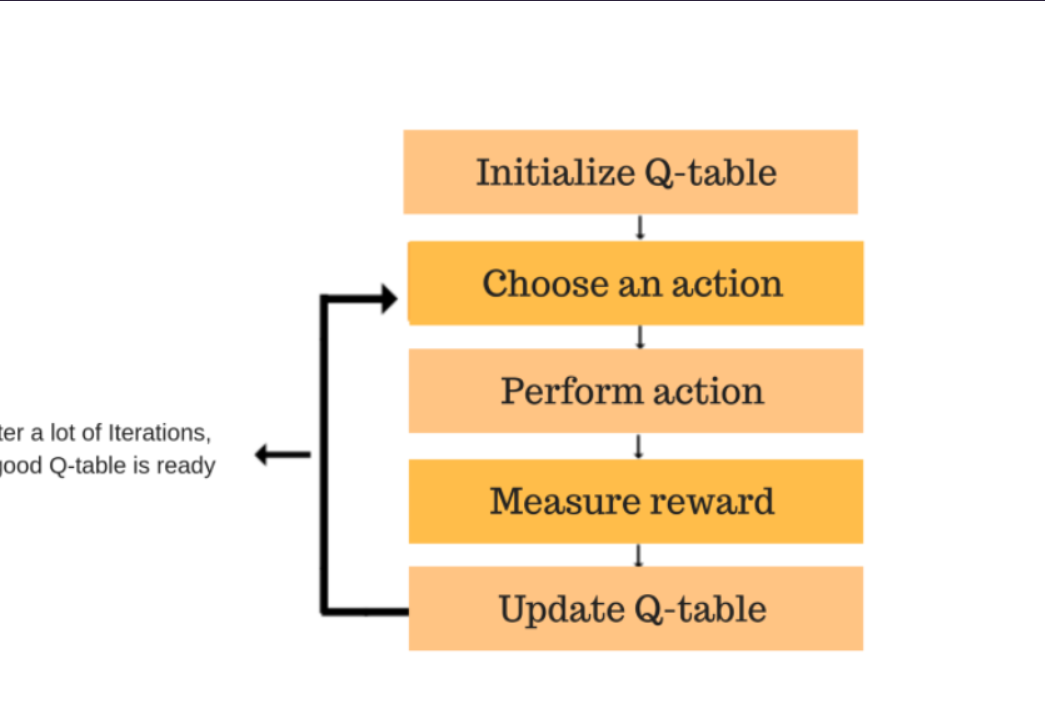


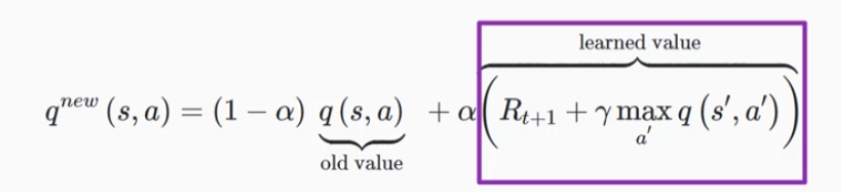
We will focus on the Q-learning because it’s the most common algorithm

Q-Learning. Methods in this family learn an approximator for the optimal action-value function, . Typically they use an objective function based on the Bellman equation. This optimization is almost always performed off-policy, which means that each update can use data collected at any point during training, regardless of how the agent was choosing to explore the environment when the data was obtained. The corresponding policy is obtained via the connection between and : the actions taken by the Q-learning agent are given by



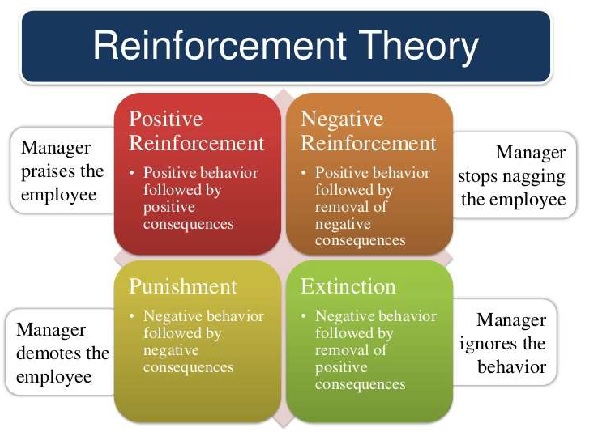
Now this is the way to implement the Q-learning method

****

****

**Types of Reinforcement Learning**

There are two types :



**1. Positive Reinforcement**

Positive reinforcement is defined as when an event, occurs due to specific behavior, increases the strength and frequency of the behavior. It has a positive impact on behavior.

**Advantages**

– Maximizes the performance of an action

– Sustain change for a longer period

**Disadvantage**

– Excess reinforcement can lead to an overload of states which would minimize the results.

**2. Negative Reinforcement**

Negative Reinforcement is represented as the strengthening of a behavior. In other ways, when a negative condition is barred or avoided, it tries to stop this action in the future.

Advantages

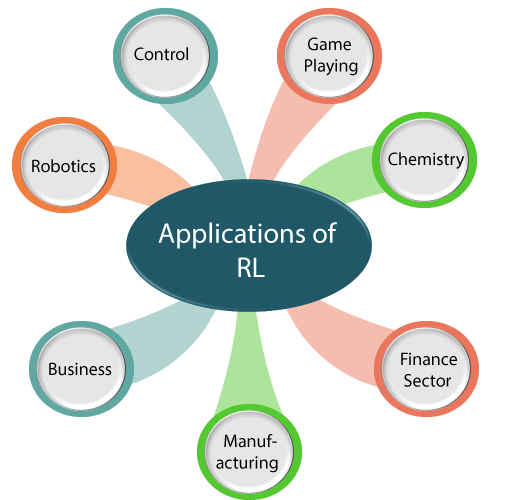
– Maximized behavior

– Provide a decent to minimum standard of performance

Disadvantage

– It just limits itself enough to meet up a minimum behavior

Reinforcement Learning Applications



1. **Robotics:**
   1. RL is used in **Robot navigation, Robo-soccer, walking, juggling**, etc.
2. **Control:**
   1. RL can be used for **adaptive control** such as Factory processes, admission control in telecommunication, and Helicopter pilot is an example of reinforcement learning.
3. **Game Playing:**
   1. RL can be used in **Game playing** such as tic-tac-toe, chess, etc.
4. **Chemistry:**
   1. RL can be used for optimizing the chemical reactions.
5. **Business:**
   1. RL is now used for business strategy planning.
6. **Manufacturing:**
   1. In various automobile manufacturing companies, the robots use deep reinforcement learning to pick goods and put them in some containers.
7. **Finance Sector:**
   1. The RL is currently used in the finance sector for evaluating trading strategies.

## Conclusion:

From the above discussion, we can say that Reinforcement Learning is one of the most interesting and useful parts of Machine learning. In RL, the agent explores the environment by exploring it without any human intervention. It is the main learning algorithm that is used in Artificial Intelligence. But there are some cases where it should not be used, such as if you have enough data to solve the problem, then other ML algorithms can be used more efficiently. The main issue with the RL algorithm is that some of the parameters may affect the speed of the learning, such as delayed feedback.