Reinforcement Learning (RL) is the field of machine learning in which an agent (i.e. the software being trained) learns to take actions to maximise some cumulative reward.

Some classical examples of RL are driverless cars, game playing agents (Chess, Go, etc.), mechanical robots in factories/warehouses etc.

**RL is a learning problem where an agent is trying to learn from its environment by taking an action and understanding the consequences; in order to achieve an objective**

**Automobile:**

[Tesla's](https://www.autotrader.com/tesla-cars.jsp)[Autopilot](https://www.tesla.com/autopilot) , Alphabet's [Waymo](https://waymo.com/), [Ford's](https://corporate.ford.com/innovation/autonomous-2021.html) self-driving car are trying for Autonomus cars.

**Robotics:**

In Robotics - finding defects in objects, carrying an object from one place to other ex: [Fanuc](https://www.technologyreview.com/s/601045/this-factory-robot-learns-a-new-job-overnight/) has deployed a robot that uses RL to pick a device from one box and put it in a container.

**Finance:**

RL is turning out to be a robust tool for evaluating trading strategies. Many companies are leveraging the "Q-Learning" algorithm of RL with the simple objective of maximising the "rewards" i.e. profits.

**Law of Effect:** Edward L. Thorndike who talked about **learning by trial and error**.

“Responses that produce a satisfying effect in a particular situation become more likely to occur again in that situation, and responses that produce a discomforting effect become less likely to occur again in that situation.”

Examples of RL :

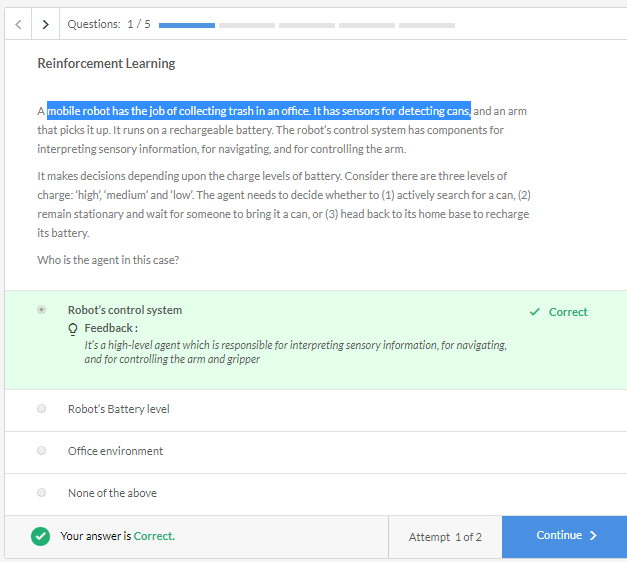
1. Solving a Maze problem (video games)
2. Managing investment portfolio
3. Deciding among the pickup requests in cab –scervice scenario
4. Process control System.

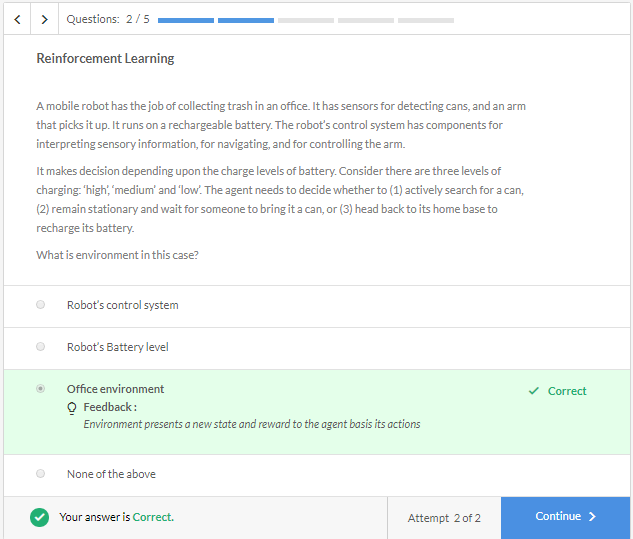
Agent is trying to solve the problem in the environment.

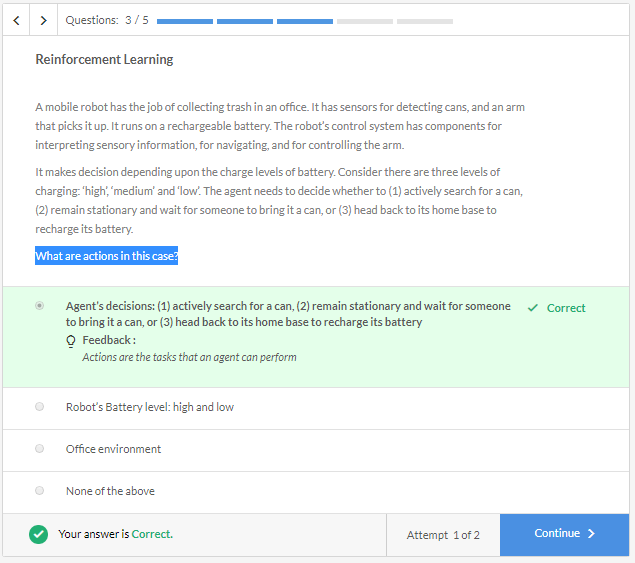
Agent: The **agent** is any robot that is trying to learn the task

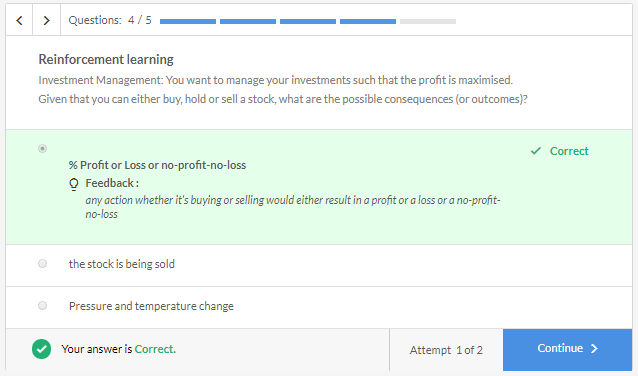
While the **environment** is the world around it that gives it the feedback.

Agent can observe the Environment and take actions . But the Consequence of the Action is not in control of Agent. It can only observe and take action next time . but cant change the consequences of the Already taken action.









**RL is a learning problem where an agent is trying to learn from its environment by taking an action and understanding the consequences; in order to achieve an objective.**

An engineer is learning to design a humanoid robot to stand and his objective is to make it stand. He can do that by turning some joints.

Here Agent is : Engineer

Environment : Robot

Actions : Turning some Joints.

Objective : Robot to Stand.

consequences : could be robot falling or standing up.

**Agent** has only control of the Actions it takes. **Environment** tells whether the Action taken by Agent is good or bad by a **Reward**. Reward only tells how well the task is performed. It doesn’t guarantee that this is the best action.

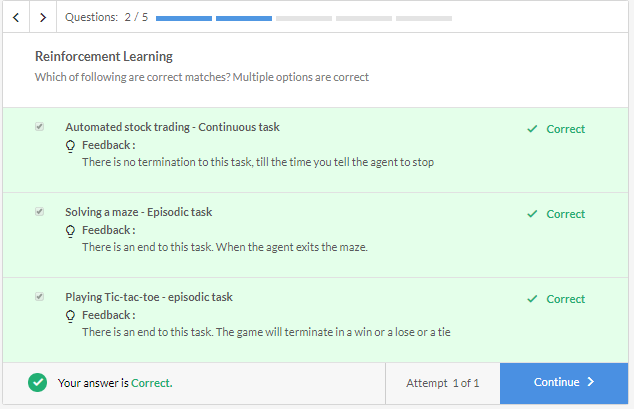
Reward is a Weak signal which indicates whether the Agent is taking decisions in right decision or not.

Unlike supervised learning, which classifies each observation as 'right' or 'wrong'; reward in reinforcement learning is just a number indicating how well you are performing the action.

**Two types of tasks:**

* **Continuous** - tasks that do not have a definite end - e.g. learning to walk, controlling a chemical plant, driving a car
* **Episodic tasks** - tasks that have a definite end - e.g. most games (videos games, Chess, Ludo) etc. are episodic since at the end of the game the agent either wins or loses.

Questions:

1. Rewards are under the control of the agent, i.e., after taking an action, it can modify the rewards obtained - False
2. 
3. 