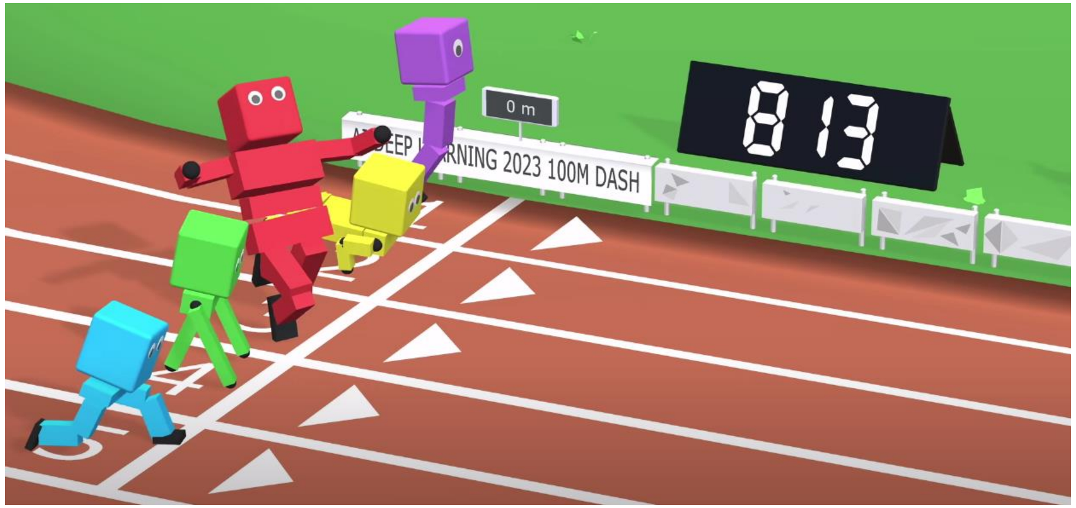
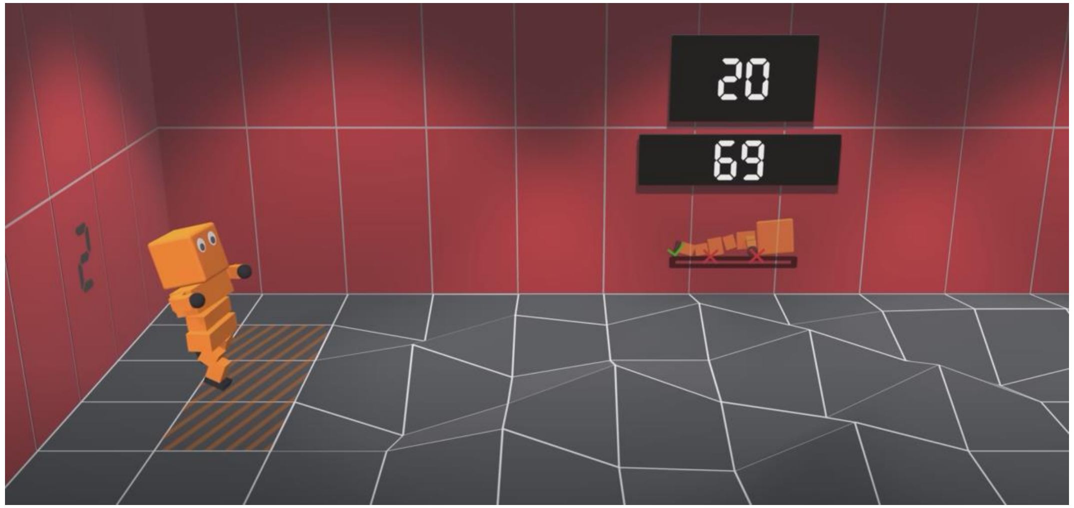
REINFORCEMENT LEARNING

Deep Reinforcement Learning

Reinforcement Learning



AI Olympics - 100m Race: https://www.youtube.com/watch?v=pJPdW8WWAso&ab_channel=AIWarehouse



AI Learns to Walk (deep reinforcement learning): https://www.youtube.com/watch?v=L_4BPjLBF4E&ab_channel=AIWarehouse

Reinforcement Learning

Reinforcement Learning focuses on teaching agents through trial and error

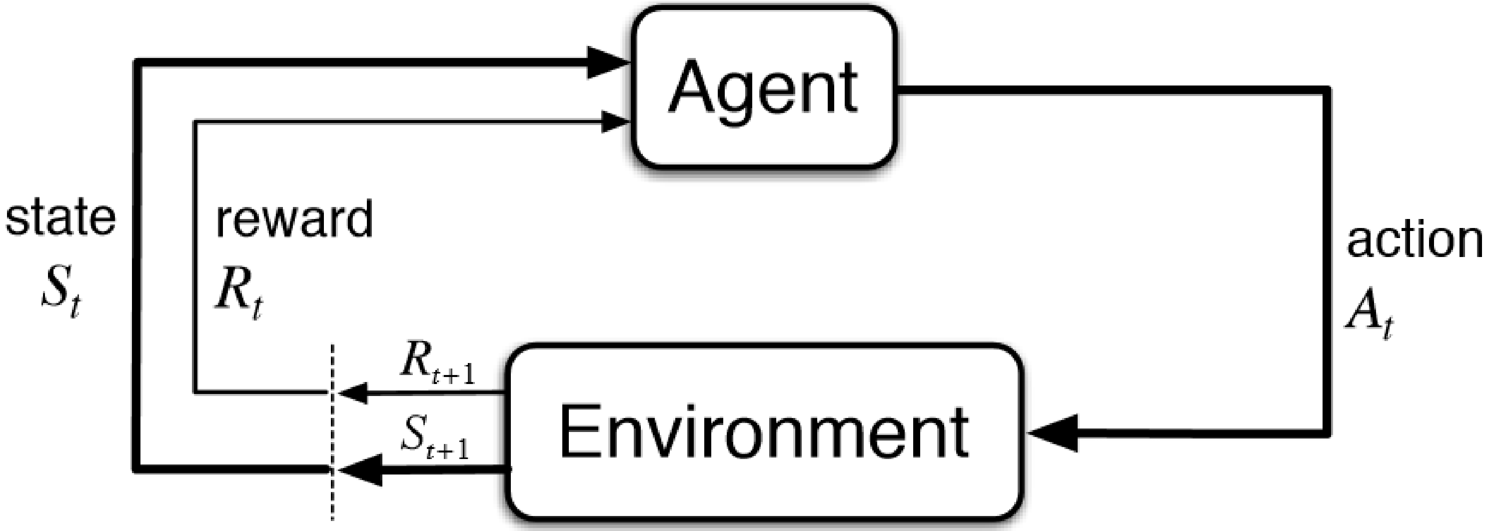


Fig. A simple illustration of the Reinforcement Learning Approach

Reinforcement Learning: Basic Concepts

There are four fundamental concepts that underpin most Reinforcement Learning (RL) projects. They are illustrated as follows:

- **Agent:** The actor operating within the environment, it is usually governed by a policy (a rule that decides what action to take).
- **Environment:** The world in which the agent can operate in.
- **Action:** The agent can do something withing the environment known as an action
- **Reward and observations:** In return, the agent receives a reward and a view of what the environment looks like after acting on it.

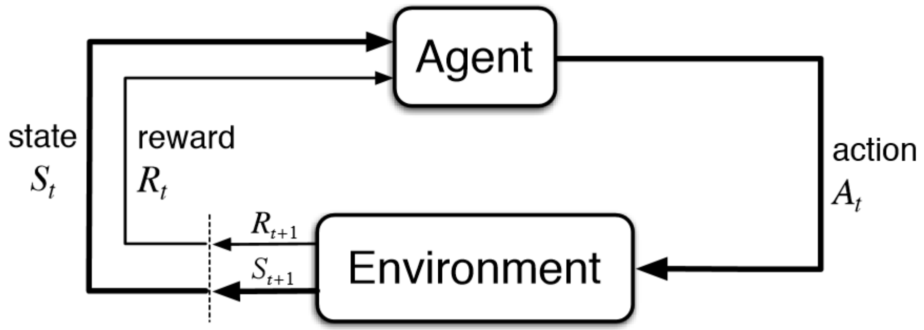
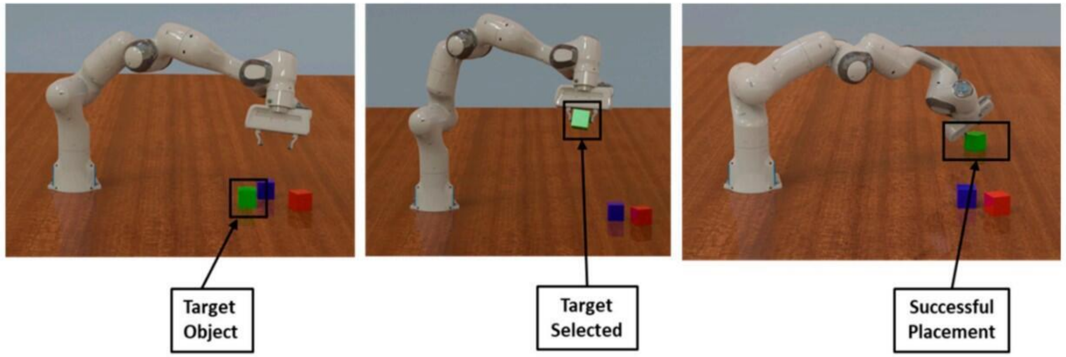


Fig. A simple illustration of the Reinforcement Learning Approach

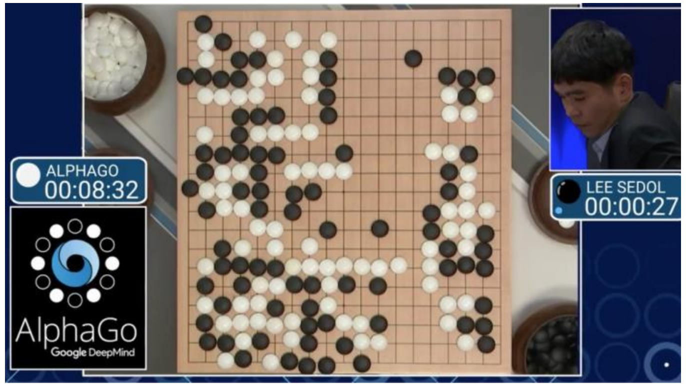
Reinforcement Learning: Applications

- **Game Playing**
  - **AlphaGo** and **AlphaZero**: DeepMind's AlphaGo program demonstrated groundbreaking performance in the ancient game of Go.
  - **OpenAI Five:** OpenAI's team of agents learned to play the complex game of Dota 2 at a professional level.
- **Robotics**
  - **Robotic Arm Control:** RL is used to train robotic arms to perform tasks like picking and placing objects, assembling components, and more.
  - **Autonomous Vehicles:** RL helps in training self-driving cars to navigate complex environments and make decisions on the road.
- **Healthcare**
  - **Drug Discovery:** RL is used to discover and optimize drug compounds by simulating chemical interactions.
  - **Personalized Treatment Plans:** RL models can recommend personalized treatment plans for patients based on their specific medical history and conditions.

Reinforcement Learning: Applications



Pick and Place: Reinforcement Learning for Robotic Control



Go Game: [AlphaGo vs. Lee Sedol]

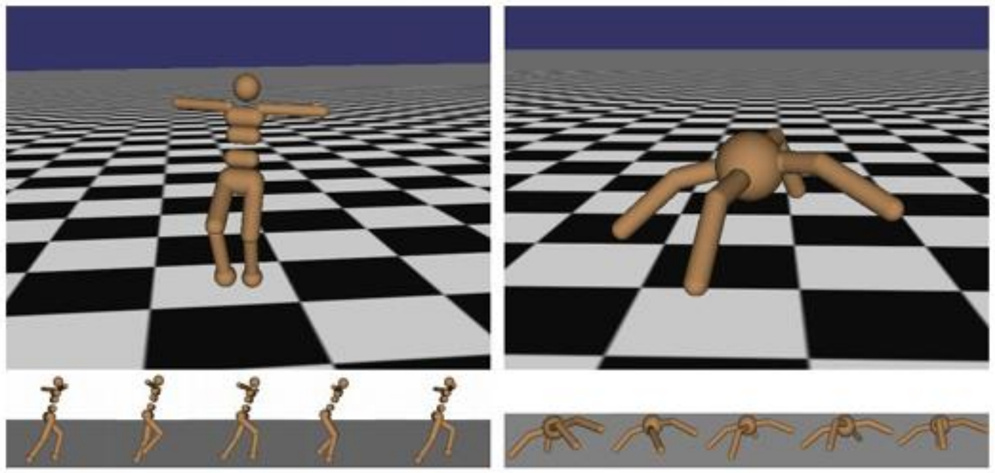
[OpenAI Five Beats World Champion DOTA2 Team]

Reinforcement Learning: Applications

- **Supply chain and logistics**
  - **Inventory Management:** RL can help optimize inventory levels to balance supply and demand, reducing carrying costs while avoiding stockouts.
  - **Routing and Scheduling:** RL can optimize routes for delivery vehicles or schedule tasks in a warehouse.
- **Game design and testing**
  - **Game AI:** RL is used to develop intelligent, adaptive non-player characters (NPCs) in video games.
  - **Automated Game Testing:** RL agents can be used to automatically test and evaluate different aspects of game performance.

Reinforcement Learning: Applications <Robot Locomotion>

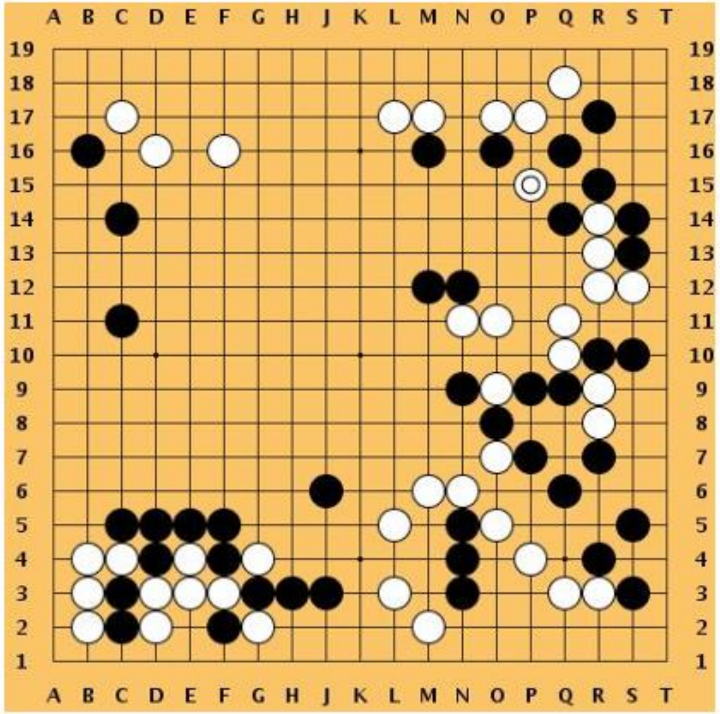
- **Objective:** Make the robot move forward
- **State:** Angle and position of the joints
- **Action:** Torques applied on joints
- **Reward:** 1 at each time step upright + forward movement



Reinforcement Learning: Applications
<Atari Games>

- **Objective:** Complete the game with the highest score
- **State:** Raw pixel inputs of the game state
- **Action:** Game controls e.g. Left, Right, Up, Down
- **Reward:** Score increase/decrease at each time step

Reinforcement Learning: Applications
<GO GAME>



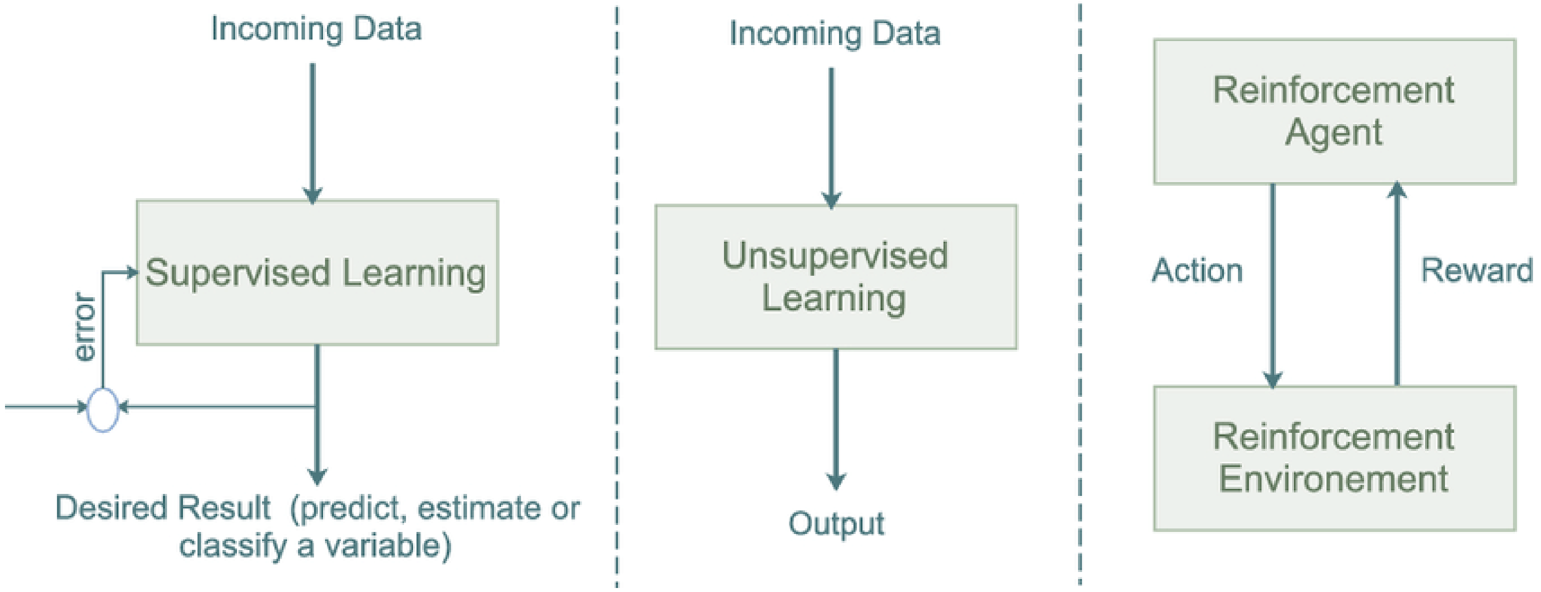
Objective: Win the game!
State: Position of all pieces
Action: Where to put the next piece down
Reward: 1 if win at the end of the game, 0 otherwise

Reinforcement Learning: Limitations and Considerations

Reinforcement Learning (RL) comes with several limitations and challenges that can make it a complex paradigm to apply in certain situations. Here are some of the key limitations:

- **Sample Inefficiency:** RL algorithms often require a large number of interactions with the environment to learn effective policies. This can be impractical or costly in real-world scenarios.
- **Exploration vs. Exploitation Trade-off:** Finding the right balance between exploring new actions and exploiting known good actions is a challenging problem, especially in environments with sparse rewards.
- **High-Dimensional State and Action Spaces:** Many real-world problems have high-dimensional state and action spaces, which can make learning and generalization difficult for traditional RL algorithms.
- **Multi-Agent Interactions:** In environments with multiple agents, the behavior of one agent may impact the learning of others, leading to complex and sometimes adversarial interactions.

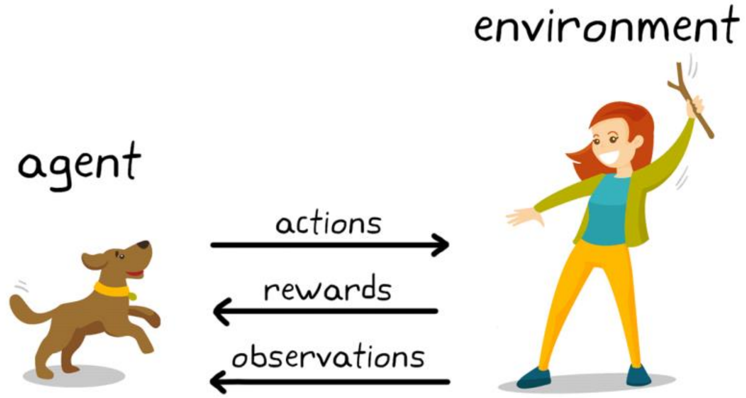
Reinforcement Learning: Comparison with Supervised and Unsupervised Learning



Reinforcement Learning: Main Steps

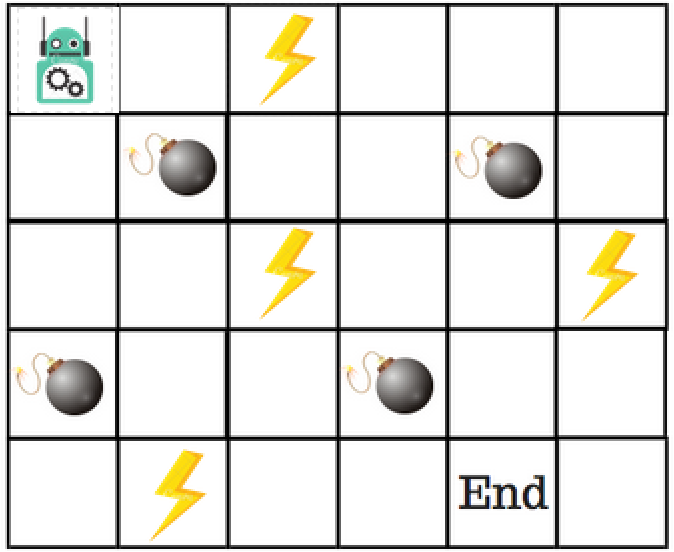
We can break down reinforcement learning into five simple steps:

1. The agent is at state zero in an environment.
2. It will take an action based on a specific strategy.
3. It will receive a reward or punishment based on that action.
4. By learning from previous moves and optimizing the strategy.
5. The process will repeat until an optimal strategy is found.



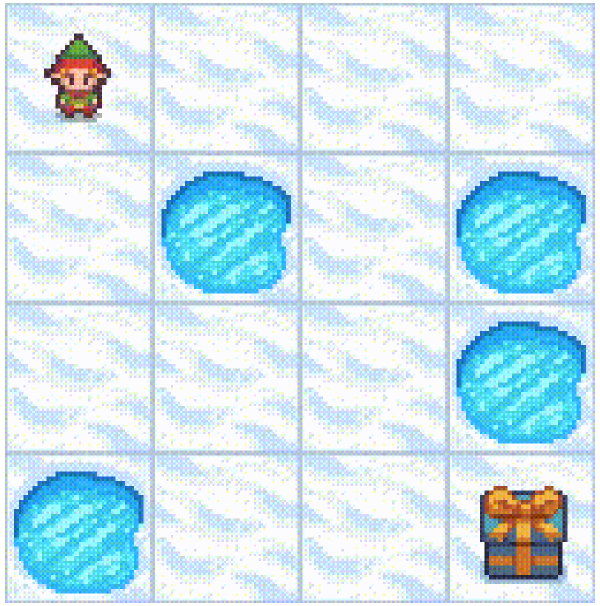
Reinforcement Learning: Q-Learning

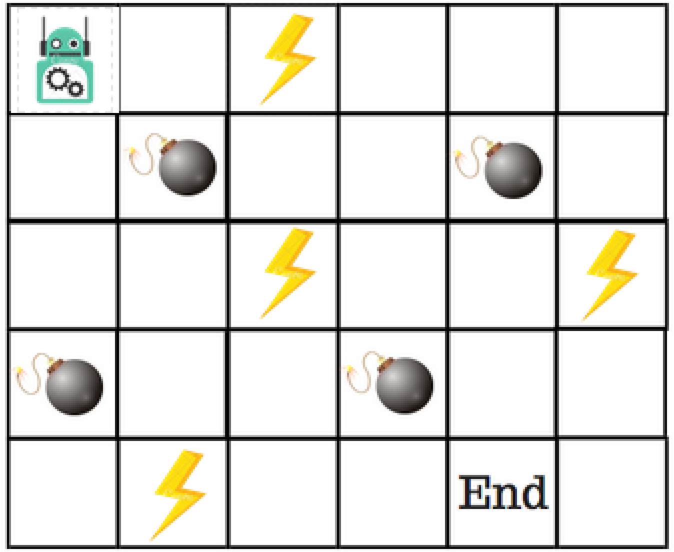
Definition: Q-Learning is a model-free reinforcement learning algorithm that allows an agent to learn a policy, or make decisions, without having a model of the environment. It's suitable for situations where the dynamics of the environment are not known in advance. The “Q” stands for quality. Quality represents how valuable the action is in maximizing future rewards.



Reinforcement Learning: How does Q-Learning Algorithm work?

We will learn in detail how Q-learning works by using the example of a frozen lake. In this environment, the agent must cross the frozen lake from the start to the goal, without falling into the holes. The best strategy is to reach goals by taking the shortest path.

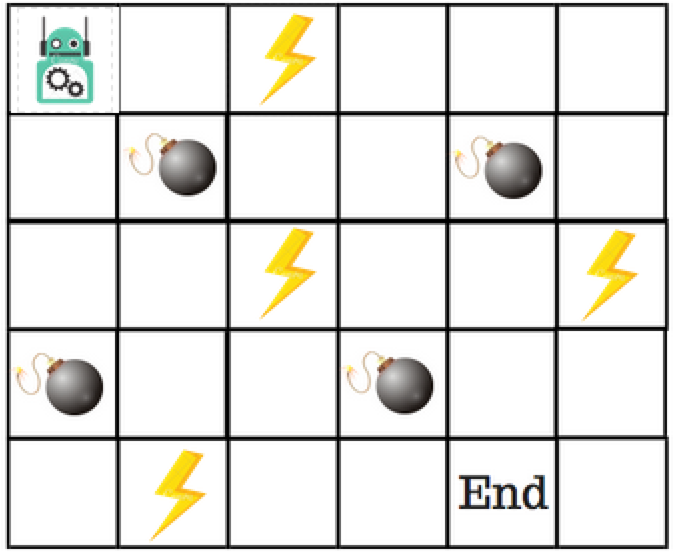




Q-Learning Algorithm: Q-Table

Q-Table is just a fancy name for a simple lookup table where we calculate the maximum expected future rewards for action at each state. Basically, this table will guide us to the best action at each state. Q-table is a data structure of sets of actions and states, and we use the Q-learning algorithm to update the values in the table. (In the Q-Table, the columns are the actions and the rows are the states.)

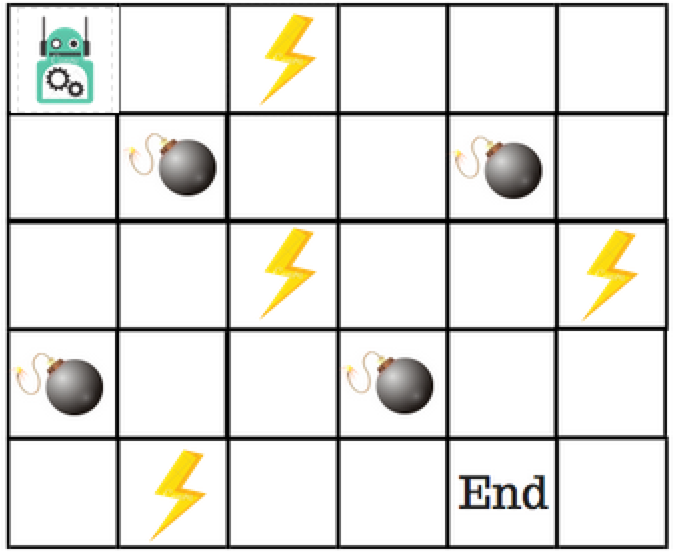
Actions :	↑	→	↓	←
Start				
Nothing / Blank				
Power				
Mines				
END				



Q-Learning Algorithm: Q-Table

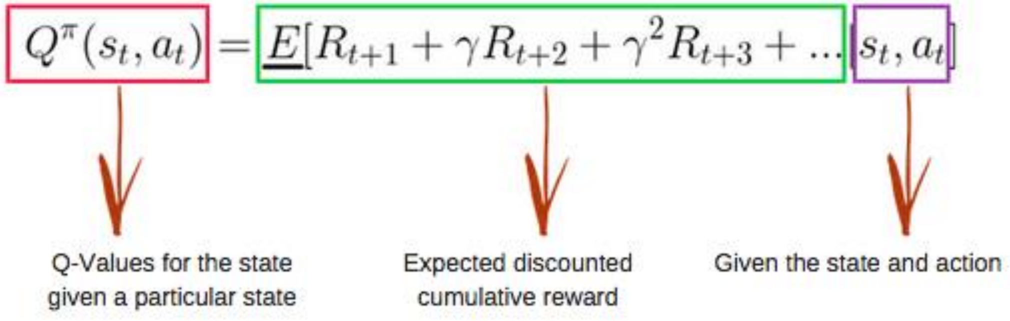
Each Q-table score will be the maximum expected future reward that the robot will get if it takes that action at that state. This is an iterative process, as we need to improve the Q-Table at each iteration.

Actions :	↑	→	↓	←
Start				
Nothing / Blank				
Power				
Mines				
END				



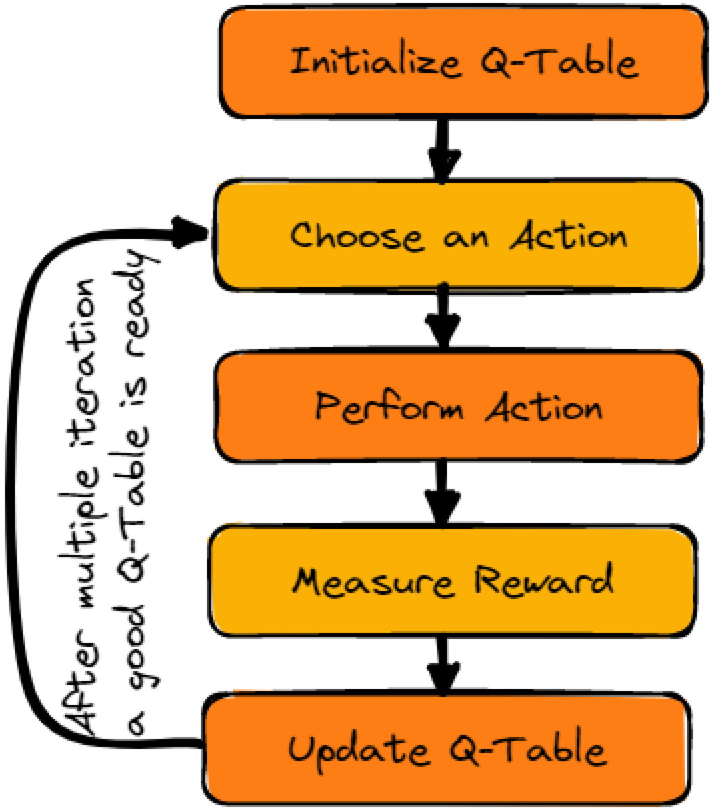
Q-Learning Algorithm: Q-Function

The Q-function uses the Bellman equation and takes state(s) and action(a) as input. The equation simplifies the state values and state-action value calculation.

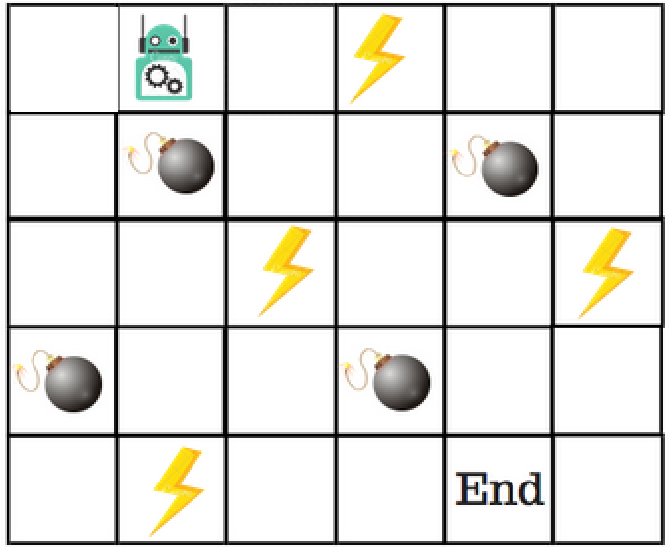


Using the above function, we get the values of Q for the cells in the table.
When we start, all the values in the Q-table are zeros.

Q-Learning Algorithm: Q-Learning Algorithm



Reinforcement Learning: DEMO



Actions :	↑	→	↓	←
Start	0	0	0	0
Nothing / Blank	0	0	0	0
Power	0	0	0	0
Mines	0	0	0	0
END	0	0	0	0

Demo: [Session 6 - Reinforcement Learning] centralelille

ufr3s faculté d'ingénierie et management de la santé

Université de Lille

REFERENCES

kaggle

Medium

YouTube

coursera

Stanford University

Massachusetts Institute of Technology

datacamp

3Blue1Brown
@3blue1brown 5.49M subscribers 135 videos

أكاديمية الفهرية Al-Fihriya Academy
@al-fihriyaacademy6448 · 25.1K subscribers · 81 videos

StatQuest with Josh Starmer
@statquest 1.01M subscribers 264 videos

m2dsupsdlclass
Master Datascience Université Paris Saclay Deep Learning
11 followers Paris

centralelille

ufr3s faculté d'ingénierie et management de la santé

Université de Lille

Your Feedback Matters!

https://forms.gle/T3co5xZK6pWTBCm67

THANK YOU

Idriss JAIRI

Email: Idriss.jairi@univ-lille.fr

LinkedIn: linkedin.com/in/jairiidriss/