Reinforcement Learning

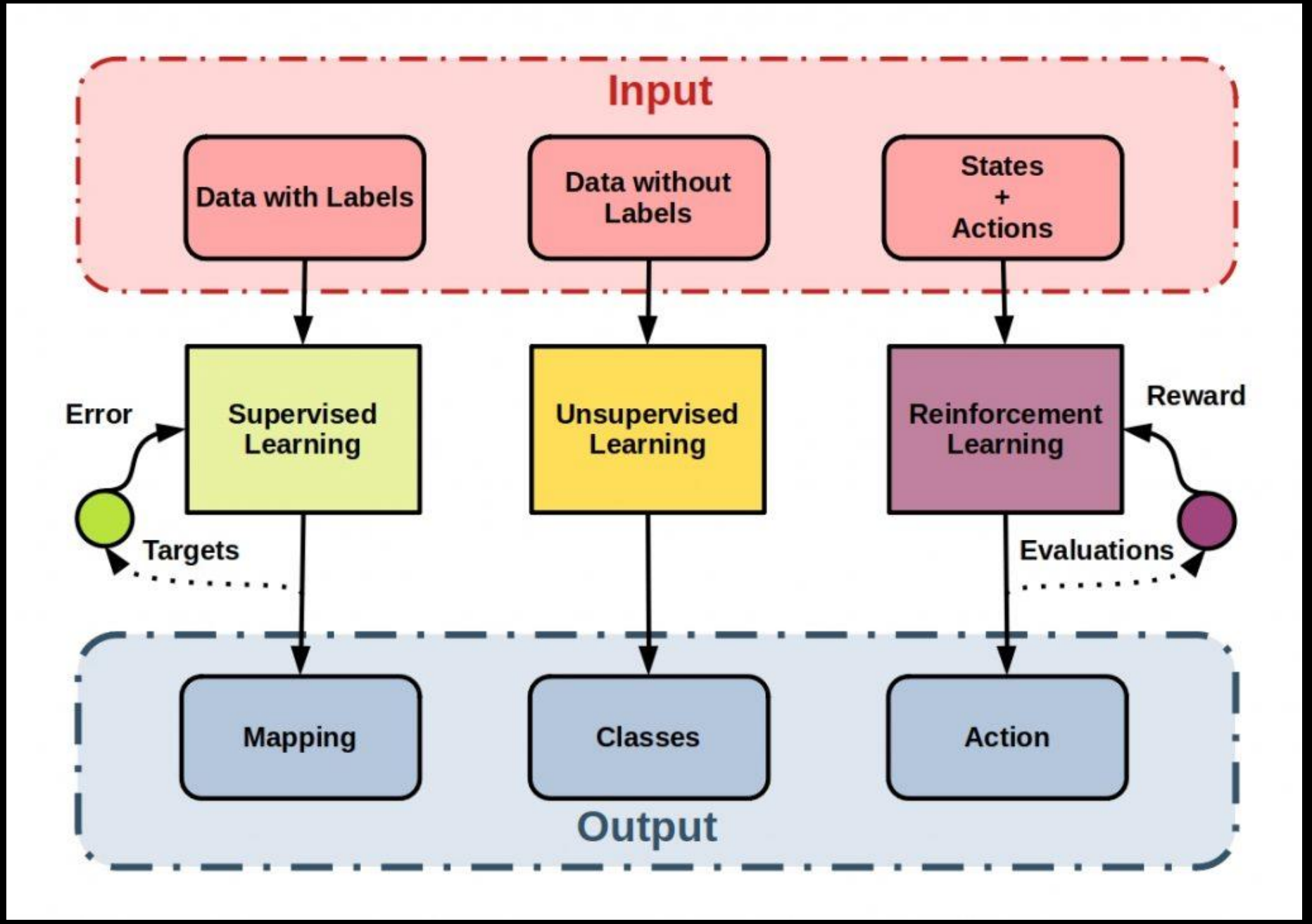
Presented by
Md Mynuddin

Reinforcement Learning?

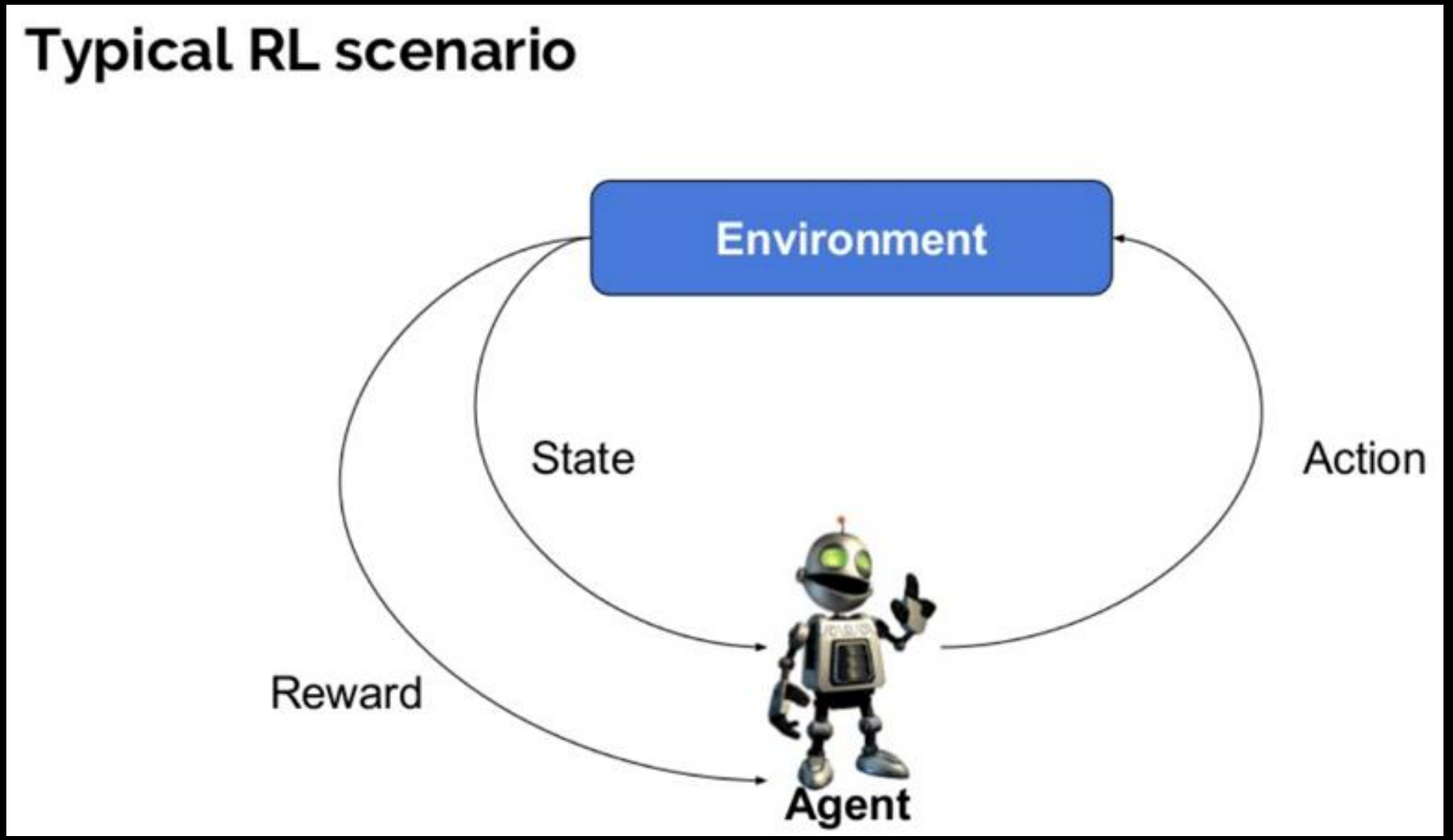
SUPERVISED LEARNING

UNSUPERVISED LEARNING

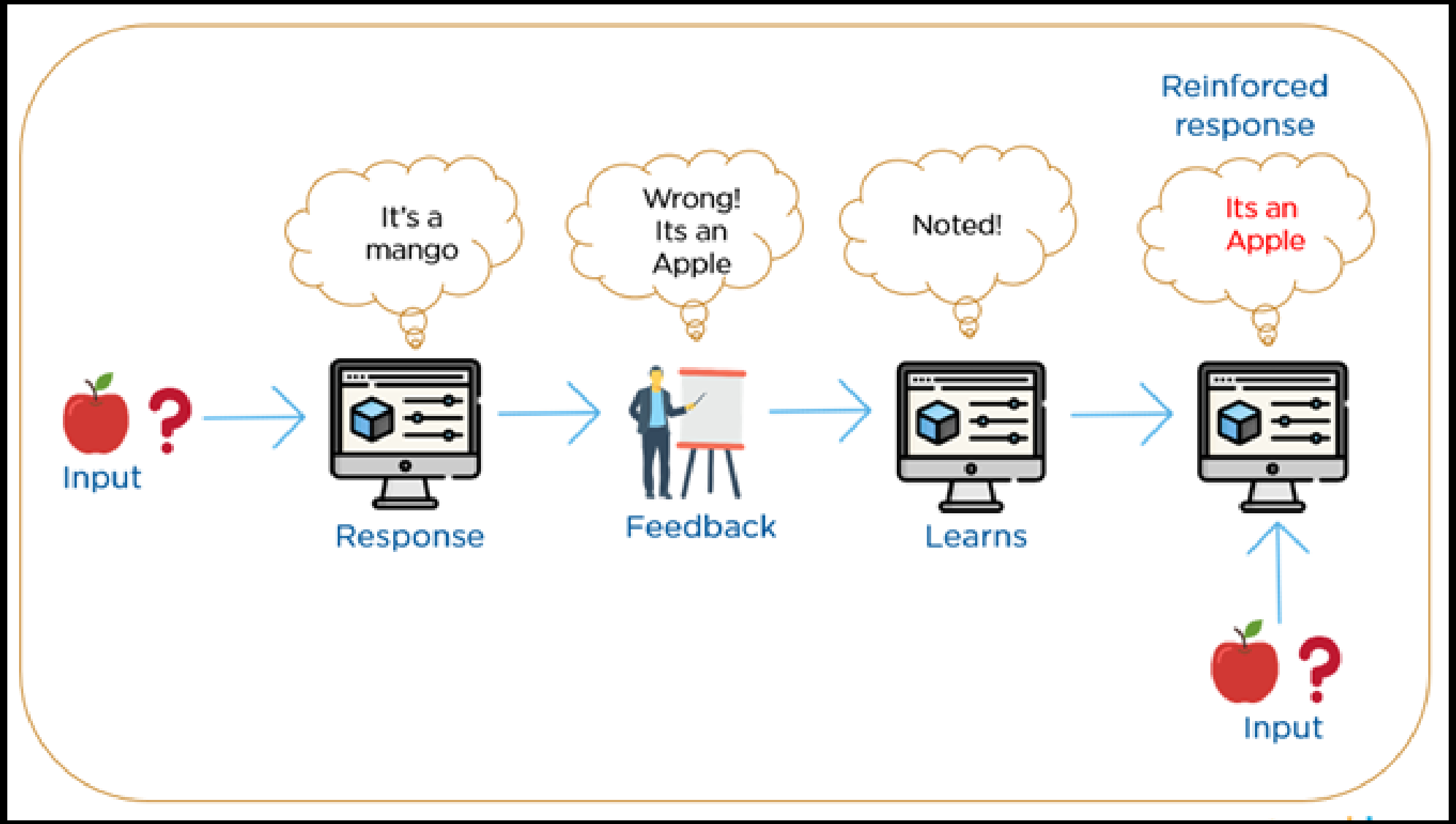
REINFORCEMENT LEARNING


Input
Data with Labels
Data without Labels
States + Actions
Error
Supervised Learning
Unsupervised Learning
Reinforcement Learning
Reward
Targets
Evaluations
Mapping
Classes
Action
Output

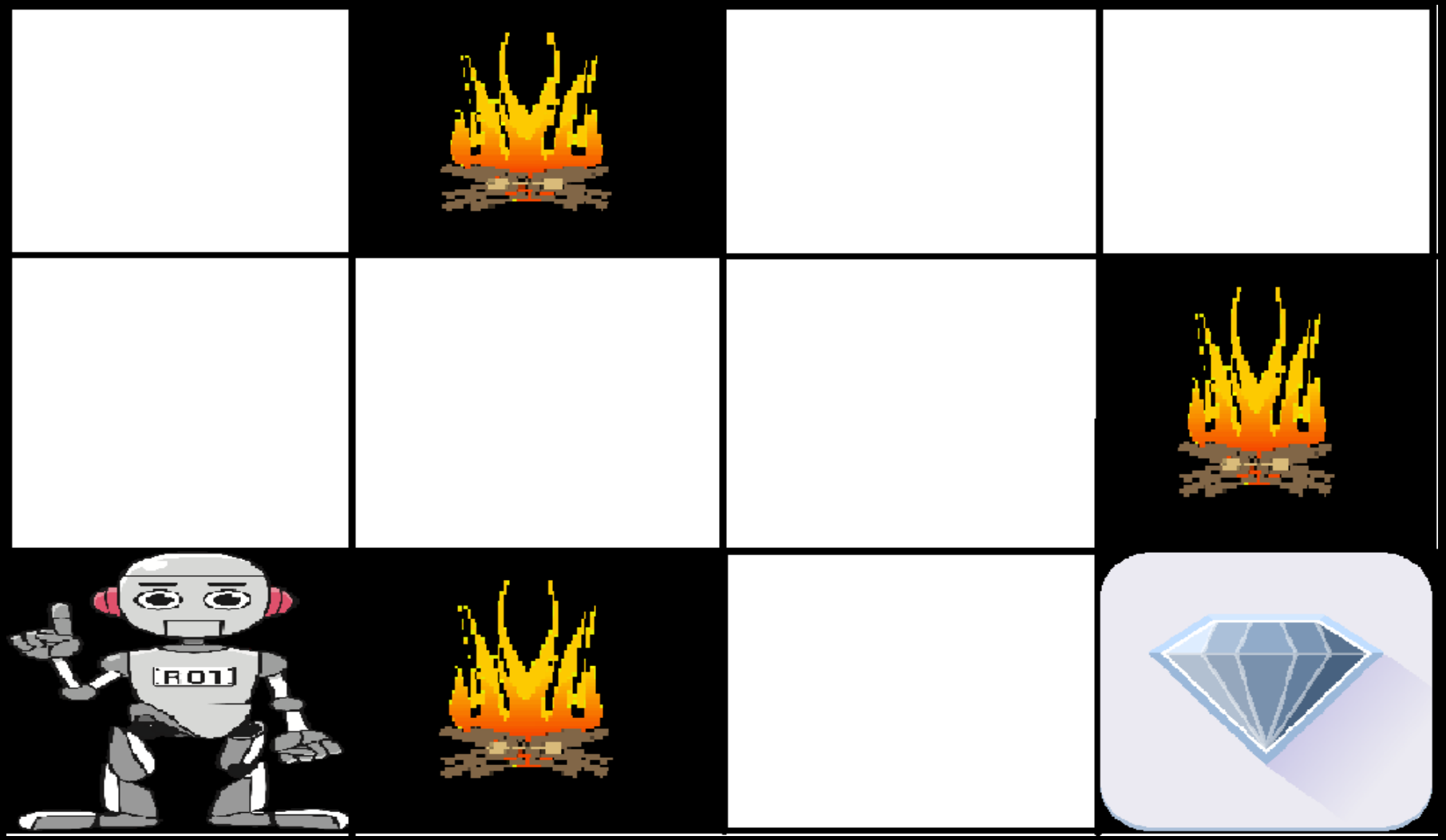
Typical RL scenario



Agent takes actions in an environment, which is interpreted into a reward and a representation of the state, which are fed back into the agent.



Making decisions sequentially



What to do ?
What not to do?

Main points in Reinforcement learning

- Input: The input should be an initial state from which the model will start
- Output: There are many possible output as there are variety of solution to a particular problem.
- Training: The training is based upon the input, The model will return a state and the user will decide to reward or punish the model based on its output.
- Sequential decision making
- There is no supervisor.
- The best solution is decided based on the maximum reward.

Application of RL

- Robotics for industrial automation.
- Business strategy planning.
- Machine learning.
- Data processing.
- Aircraft control .
- robot motion control.

Learning Models of Reinforcement

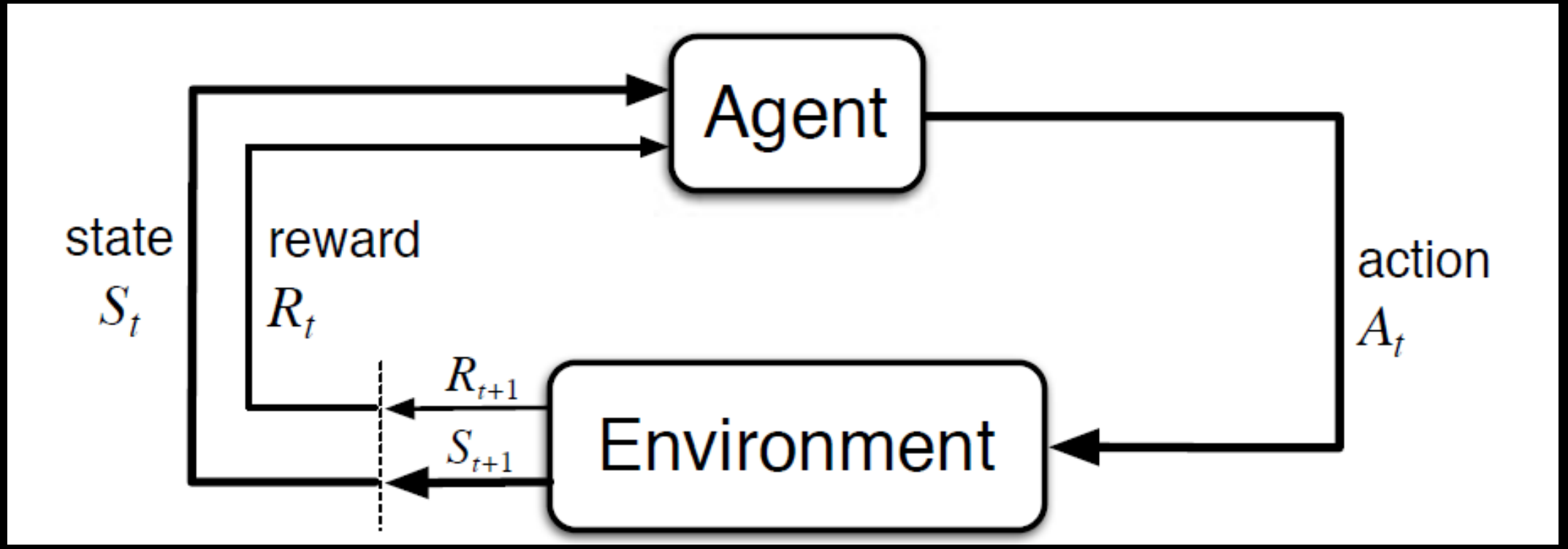
- Markov Decision Process
- Q learning

Markov Decision Process

The following parameters are used to get a solution:

- Set of actions- A
- Set of states -S
- Reward- R
- Policy- n
- Value- V

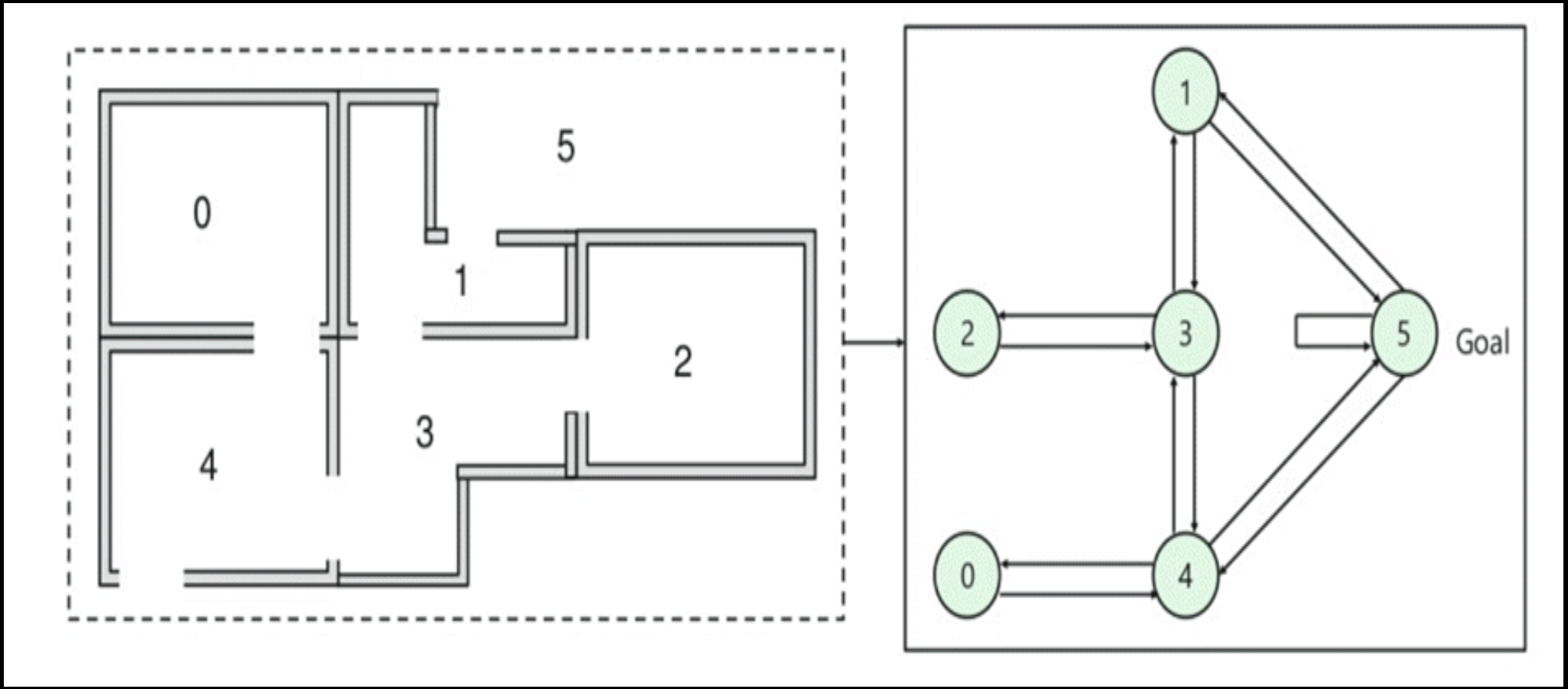
The mathematical approach for mapping a solution in reinforcement Learning is recon as a Markov Decision Process or (MDP).



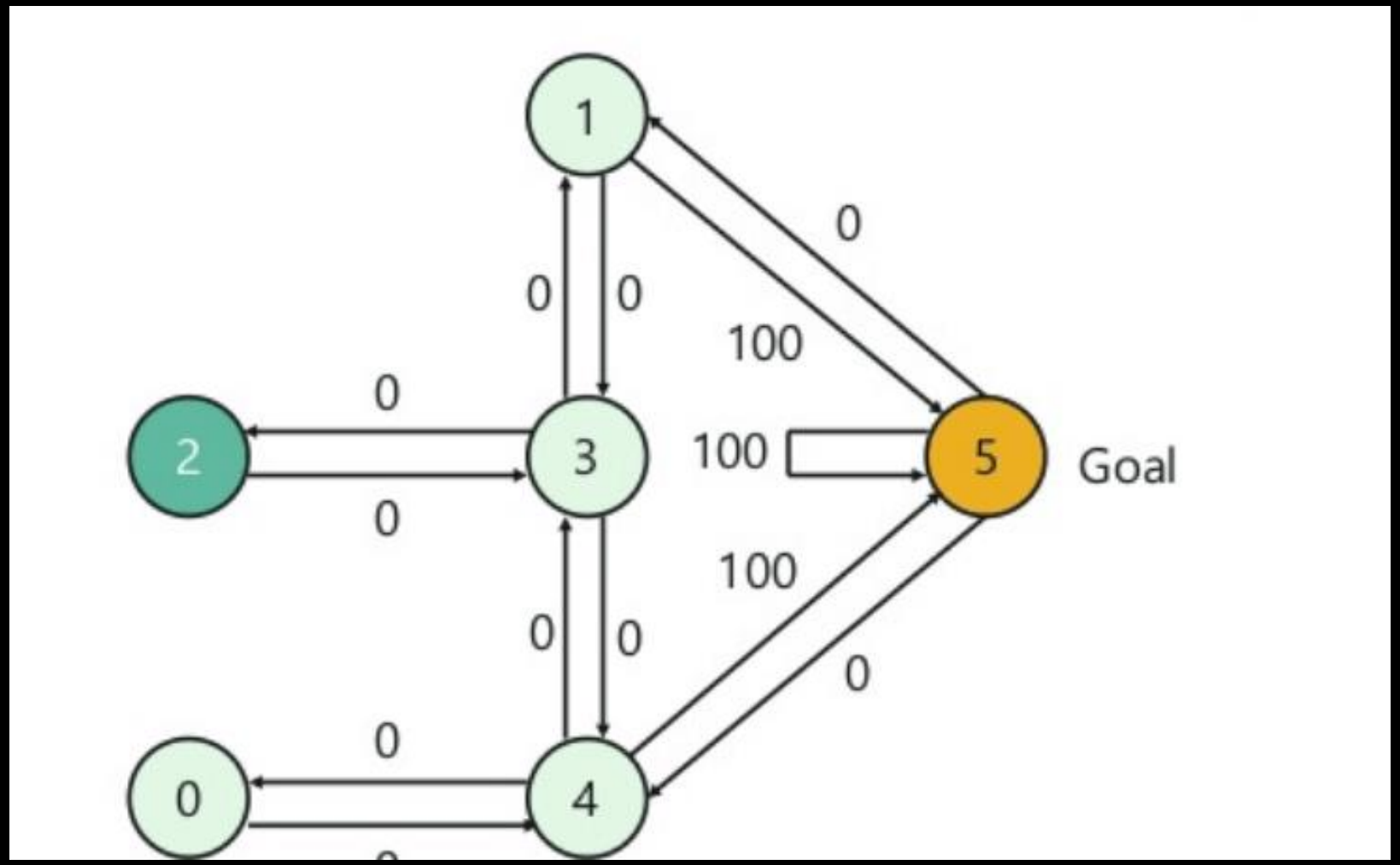
Q learning

Q learning is a value-based method of supplying information to inform which action an agent should take. Let's understand this method by the following example:

- There are five rooms in a building which are connected by doors.
- Each room is numbered 0 to 4
- The outside of the building can be one big outside area (5)
- Doors number 1 and 4 lead into the building from room 5



- Next, you need to associate a reward value to each door:
- Doors which lead directly to the goal have a reward of 100
- Doors which is not directly connected to the target room gives zero reward
- As doors are two-way, and two arrows are assigned for each room
- Every arrow in the above image contains an instant reward value



For example, an agent traverse from room number 2 to 5

Initial state = state 2
State 2-> state 3
State 3 -> state (2,1,4)
State 4-> state (0,5,3)
State 1-> state (5,3)
State 0-> state 4

Thanks 😀