

Machine Learning

Dr Changjiang He, Dr Kuo-Ming Chao
Computer Science | School of Art
University of Roehampton

Lesson 9.1

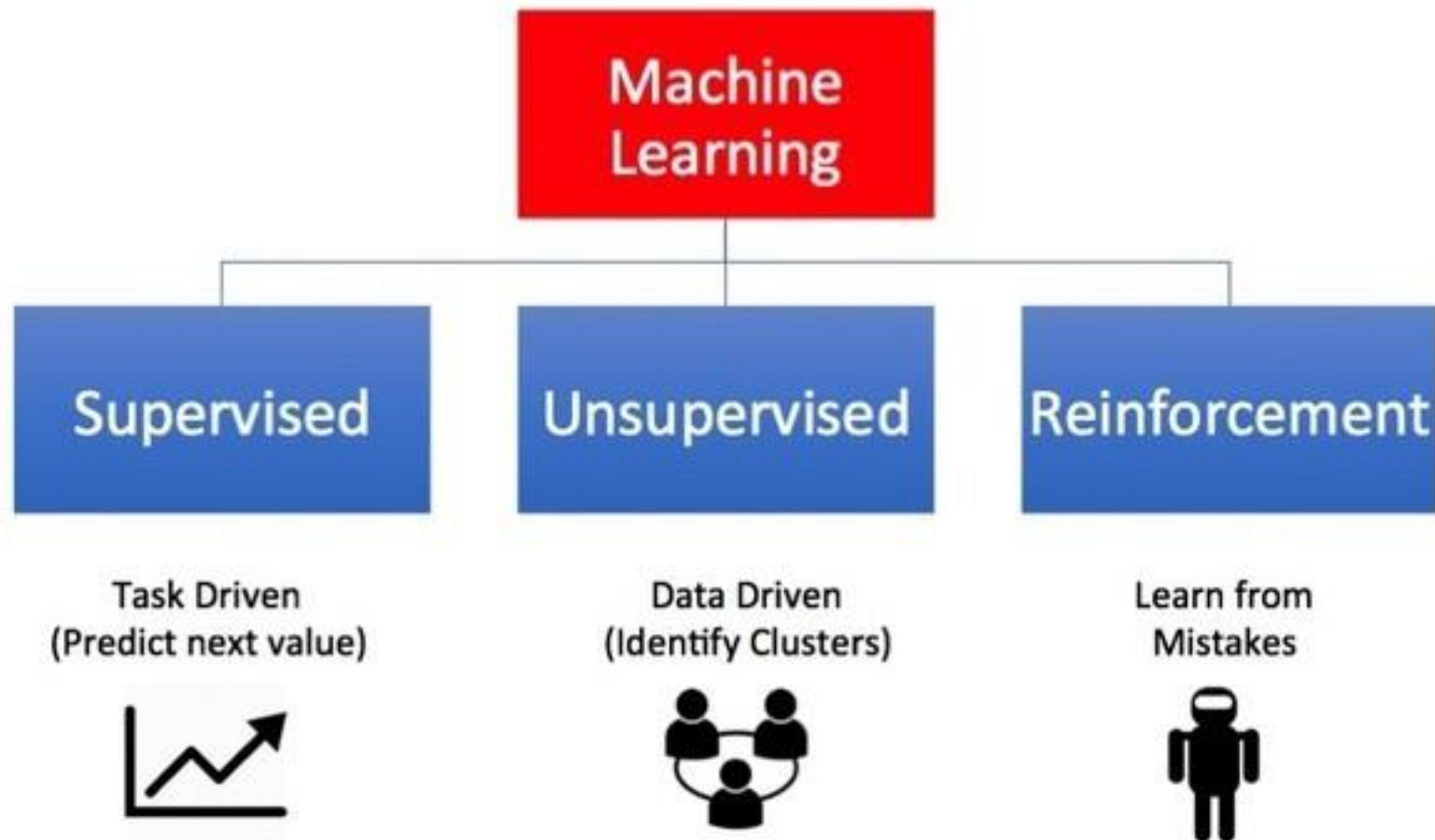
Reinforcement Learning

- What is reinforcement learning?
- Example of reinforcement learning
- RL Problem Basics

- Reinforcement Learning (RL) is a type of machine learning technique that enables an agent to learn in an interactive environment by trial and error using feedback from its own actions and experiences.
- reinforcement learning uses **rewards and punishments** as signals for positive and negative behaviour.

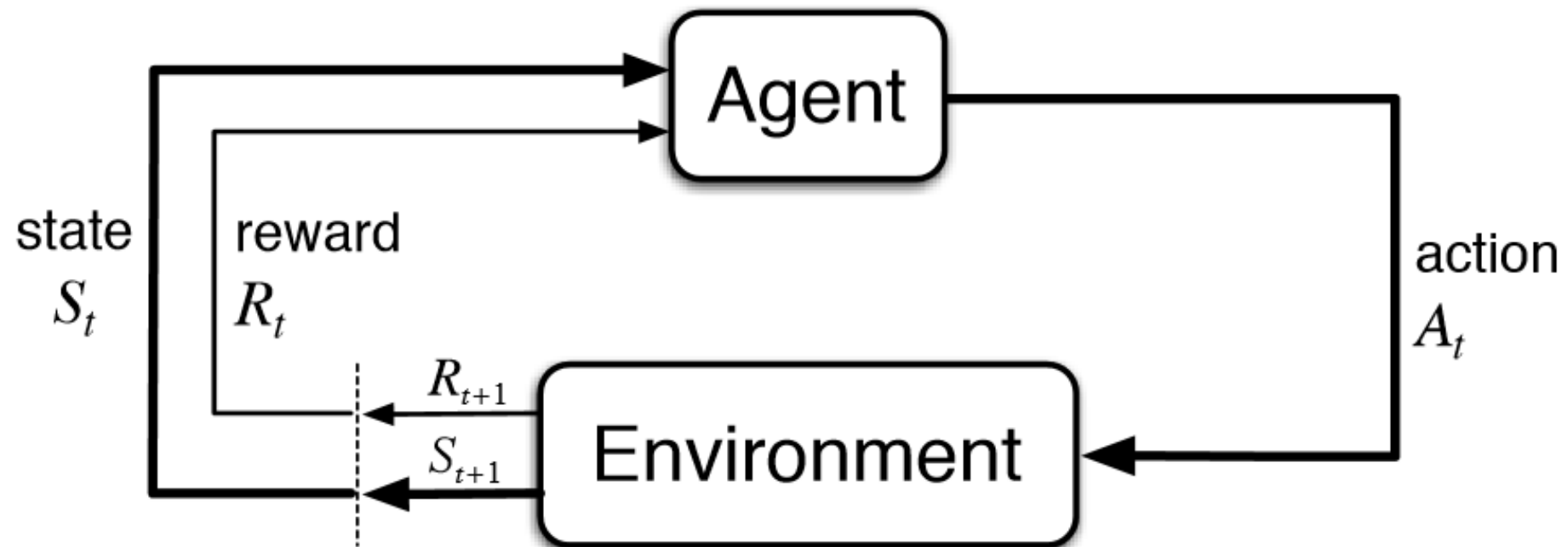
What is Reinforcement Learning

Types of Machine Learning



What is Reinforcement Learning

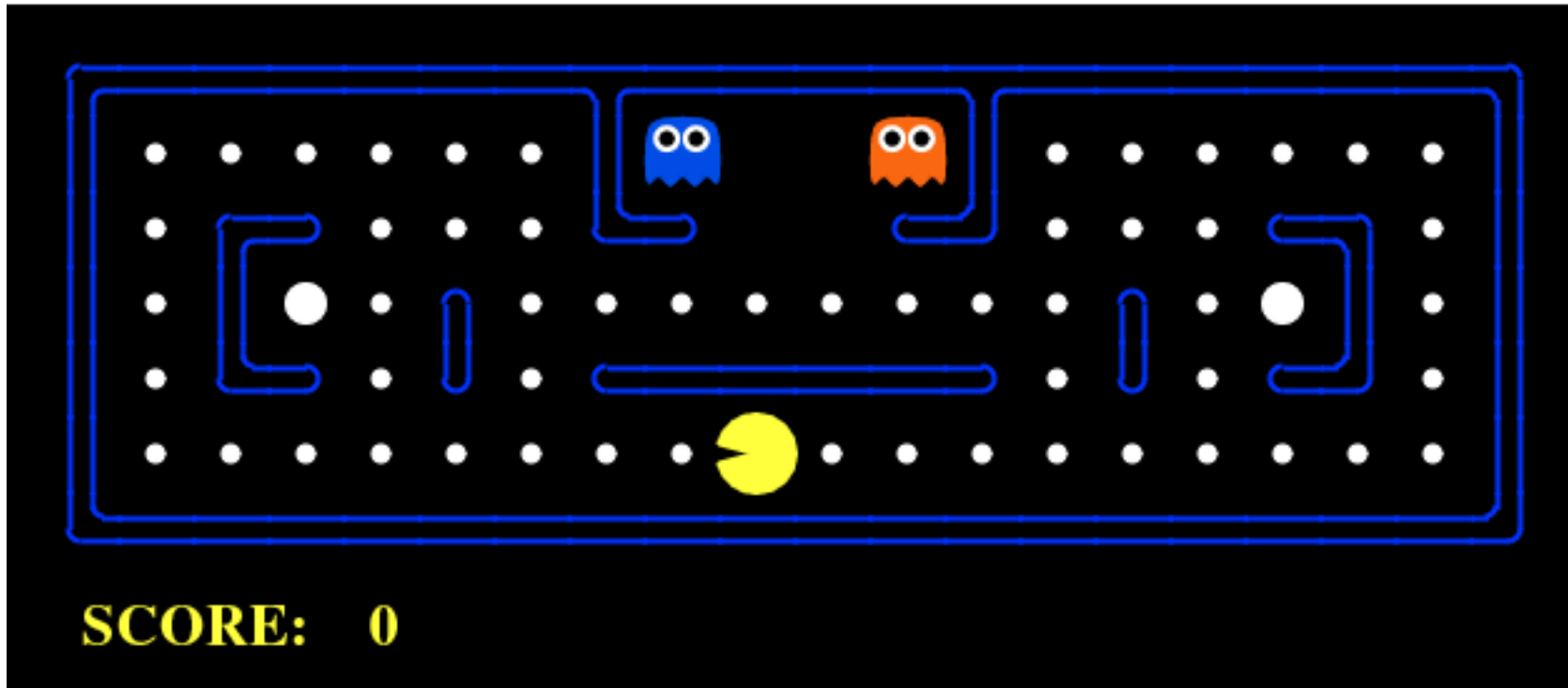
the goal is to find a suitable action model that would maximize the **total cumulative reward** of the agent.



Some key terms that describe the basic elements of an RL problem are:

- 1.Environment** — Physical world in which the agent operates
- 2.State** — Current situation of the agent
- 3.Reward** — Feedback from the environment
- 4.Policy** — Method to map agent's state to actions
- 5.Value** — Future reward that an agent would receive by taking an action in a particular state

Let's take the game of [PacMan](#) where the goal of the agent (PacMan) is to eat the food in the grid while avoiding the ghosts on its way.



- In this case, the grid world is the interactive environment for the agent where it acts.
- Agent receives a reward for eating food and punishment if it gets killed by the ghost (loses the game).
- The states are the location of the agent in the grid world and the total cumulative reward is the agent winning the game.

- In order to build an optimal policy, the agent faces the dilemma of exploring new states while maximizing its overall reward at the same time.
- This is called **Exploration vs Exploitation** trade-off.
- To balance both, the best overall strategy may involve short term sacrifices.
- Therefore, the agent should collect enough information to make the best overall decision in the future.

- Since, RL requires a lot of data, therefore it is most applicable in domains where simulated data is readily available like gameplay, robotics.
- For example, RL is quite widely used in building AI for playing computer games. **AlphaGo zero** is the first computer program to defeat a world champion in the ancient Chinese game of Go.
- Others include ATARI games, Backgammon ,etc