Lecture-1
An Introduction

Instructions

- Please join Google Classroom to get updates.
- Attendance will be marked on Google Forms. You should have internet to mark the attendance.
- In case of any query, you can contact me on my email.

ishabansatti@gmail.com

- Please be punctual in class.

Marks Breakup

- Quizzes (25 = 5 X 5)
- Assignments (5 = 1 X 5)
- Mid Exam (30)
- Final Exam (40)

Instructions

- MCQs based quizzes on google forms.
- You should have a GitHub account to upload your assignment codes.
- All assignments will be followed by a relevant viva.

Agenda

01. Learning by Interaction

02. Reinforcement Learning (RL)

03. Elements of RL

04. Examples

05. Difference from Supervised and Unsupervised Learning

01.

Learning by Interaction

Learning by Interaction

- No explicit teacher.
- Direct interaction with the environment.
- Learn cause and effect, and about what to do to achieve goals.
- Examples: Infant movement, driving a car or holding a conversation.

02.

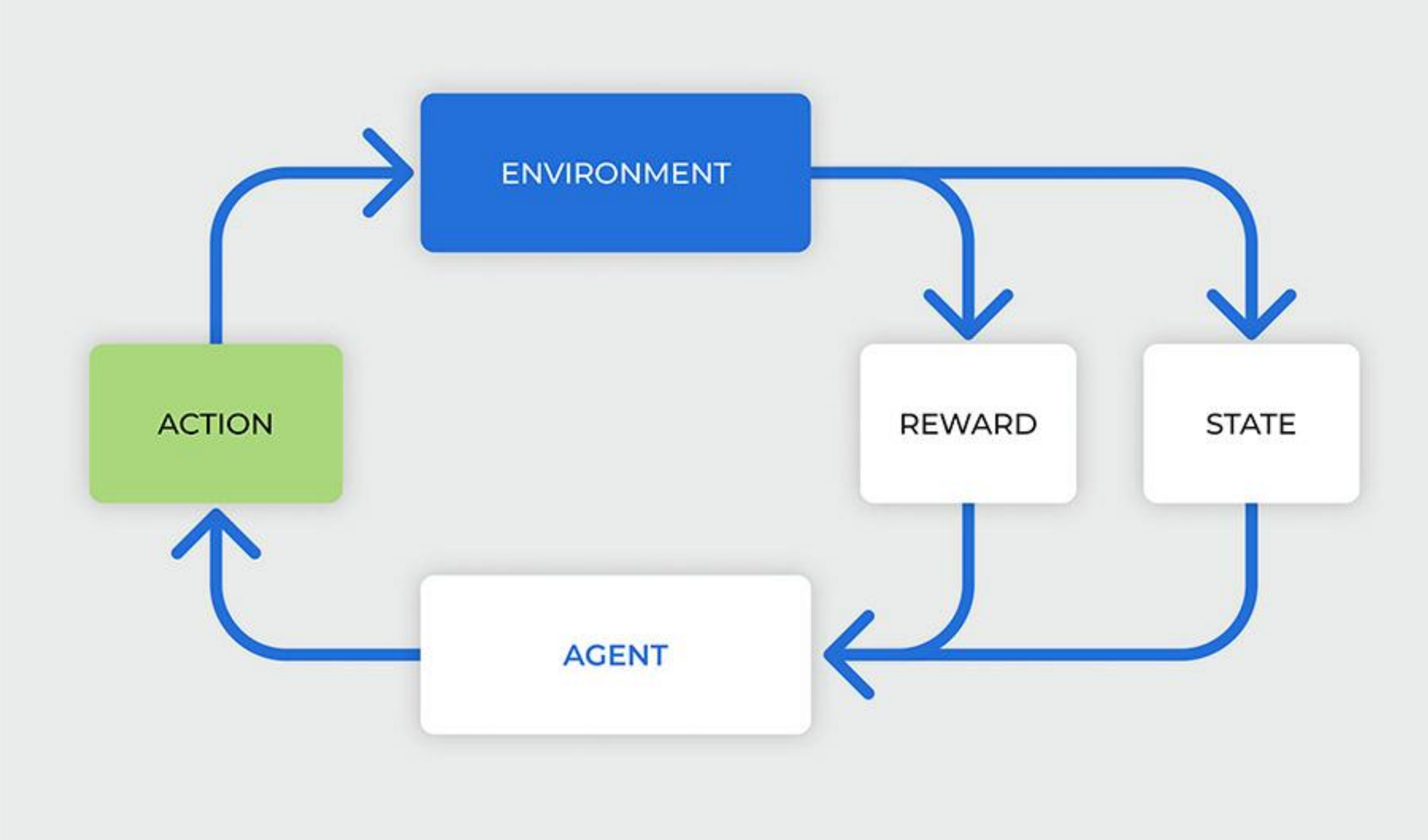
Reinforcement Learning (RL)

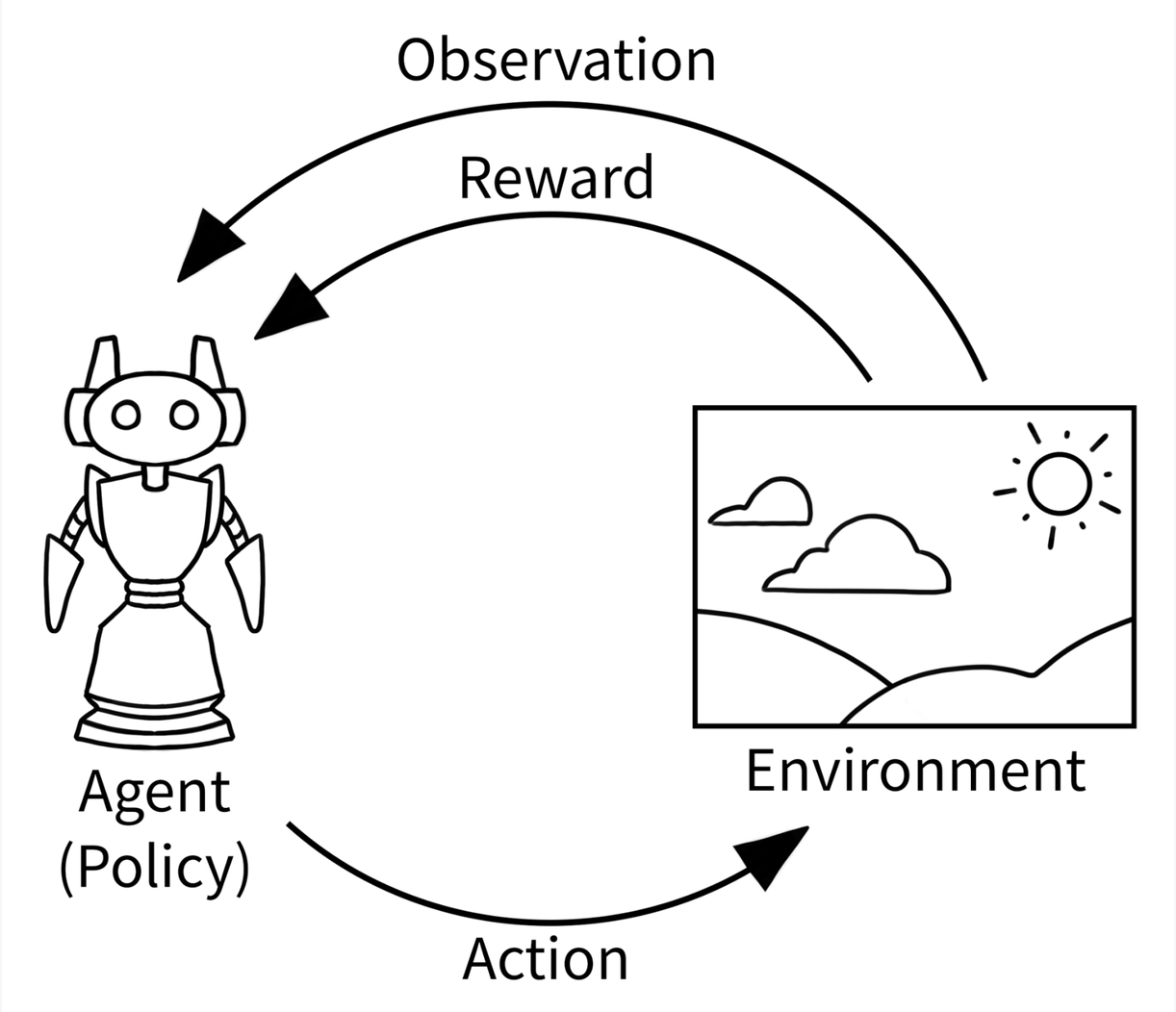
Reinforcement Learning

- A computational approach to learn from interaction.
- RL is learning
  - what to do i.e. how to map situations to actions
  - To maximize a numerical reward signal.
- Two distinguishing features of RL
  - trial-and-error search
  - delayed reward

03.

Elements of RL


ENVIRONMENT
ACTION
REWARD
STATE
AGENT


Observation
Reward
Agent
(Policy)
Environment
Action

Agent and Environment

- The environment is the world that the agent lives in and interacts with.
- At every step of interaction, the agent sees a (possibly partial) observation of the state of the world and then decides on an action to take. The environment changes when the agent acts on it but may also change on its own.

States and Observations

- A state is a complete description of the state of the world. There is no information about the world which is hidden from the state. An observation is a partial description of a state, which may omit information.
- The state of a robot might be represented by its joint angles and velocities.
- When the agent is able to observe the complete state of the environment, the environment is fully observed. When the agent can only see a partial observation, we say that the environment is partially observed.

State Space

- The state space is the set of all possible states that the environment can be in.
- A state is a complete description of the situation in the environment at a particular point in time.
- For a robot on a grid, if the grid is 3x3, each cell on the grid can be a possible position of the robot. The state space would include all possible positions the robot can occupy, i.e., 9.

Reward

- A reward R_t is a scalar feedback signal
- Indicates how well agent is doing at step t
- The agent's job is to maximise cumulative reward

Policy

- A policy is the agent's behavior.
- It is a map from state to action, e.g.

  Deterministic policy: $a = \pi(s)$

  Stochastic policy: $\pi(a|s) = \mathbb{P}[A_t = a|S_t = s]$

Action Space

- The action space defines all possible actions an agent can take within an environment.
- The action space can be either discrete or continuous.
- For example, in a game like chess, the action space is discrete—each move corresponds to a specific, countable option. In contrast, a self-driving car's steering angle or acceleration might operate in a continuous action space, where actions are real-valued and infinite in possibility.

Value Function

- Value function is a prediction of future reward.
- Used to evaluate the goodness/badness of states.
- And therefore, to select between actions, e.g.

$$v_\pi(s) = \mathbb{E}_\pi \left[R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \ldots \mid S_t = s \right]$$

Model

- A model predicts what the environment will do next
- P predicts the next state
- R predicts the next (immediate) reward, e.g.

$$\mathcal{P}^a_{ss'} = \mathbb{P}[S_{t+1} = s' \mid S_t = s, A_t = a]$$
$$\mathcal{R}^a_s = \mathbb{E}\left[R_{t+1} \mid S_t = s, A_t = a\right]$$

04.

Examples

Gymnasium Env.

Frozen Lake

This environment is part of the Toy Text environments which contains general information about the environment.

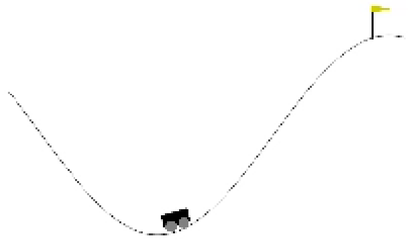
Action Space	`Discrete(4)`
Observation Space	`Discrete(16)`
import	`gymnasium.make("FrozenLake-v1")`

Frozen lake involves crossing a frozen lake from start to goal without falling into any holes by walking over the frozen lake. The player may not always move in the intended direction due to the slippery nature of the frozen lake.

https://gymnasium.farama.org/

Gymnasium Env.

Mountain Car Continuous

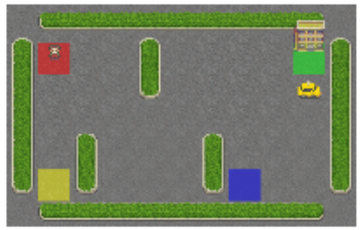


This environment is part of the Classic Control environments which contains general information about the environment.

Action Space	`Box(-1.0, 1.0, (1,), float32)`
Observation Space	`Box([-1.2 -0.07], [0.6 0.07], (2,), float32)`
import	`gymnasium.make("MountainCarContinuous-v0")`

Gymnasium Env.

Taxi



This environment is part of the Toy Text environments which contains general information about the environment.

Action Space	`Discrete(6)`
Observation Space	`Discrete(500)`
import	`gymnasium.make("Taxi-v3")`

The Taxi Problem involves navigating to passengers in a grid world, picking them up and dropping them off at one of four locations.

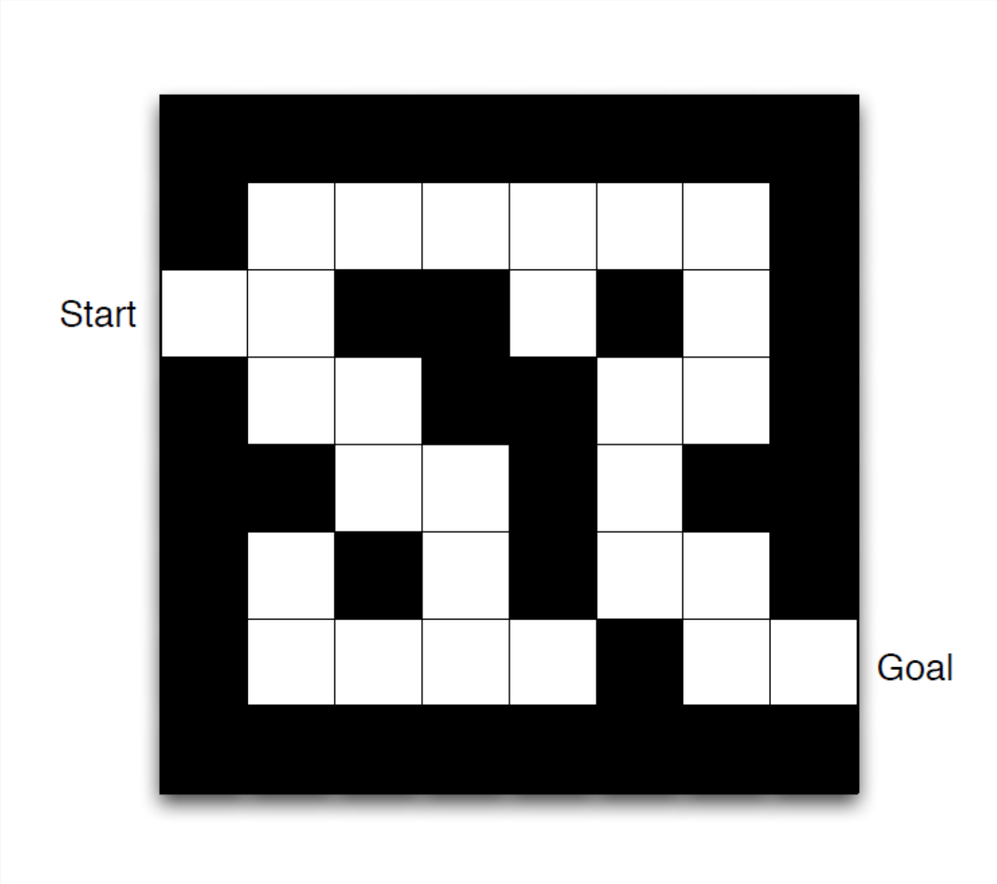
Bioreactor

- **Actions:** Target temperatures and target stirring rates that are passed to lower-level control systems.
- **States:** thermocouple and other sensory readings, perhaps filtered and delayed, plus symbolic inputs representing the ingredients.
- **Rewards:** moment-by-moment measures of the rate at which the useful chemical is produced.

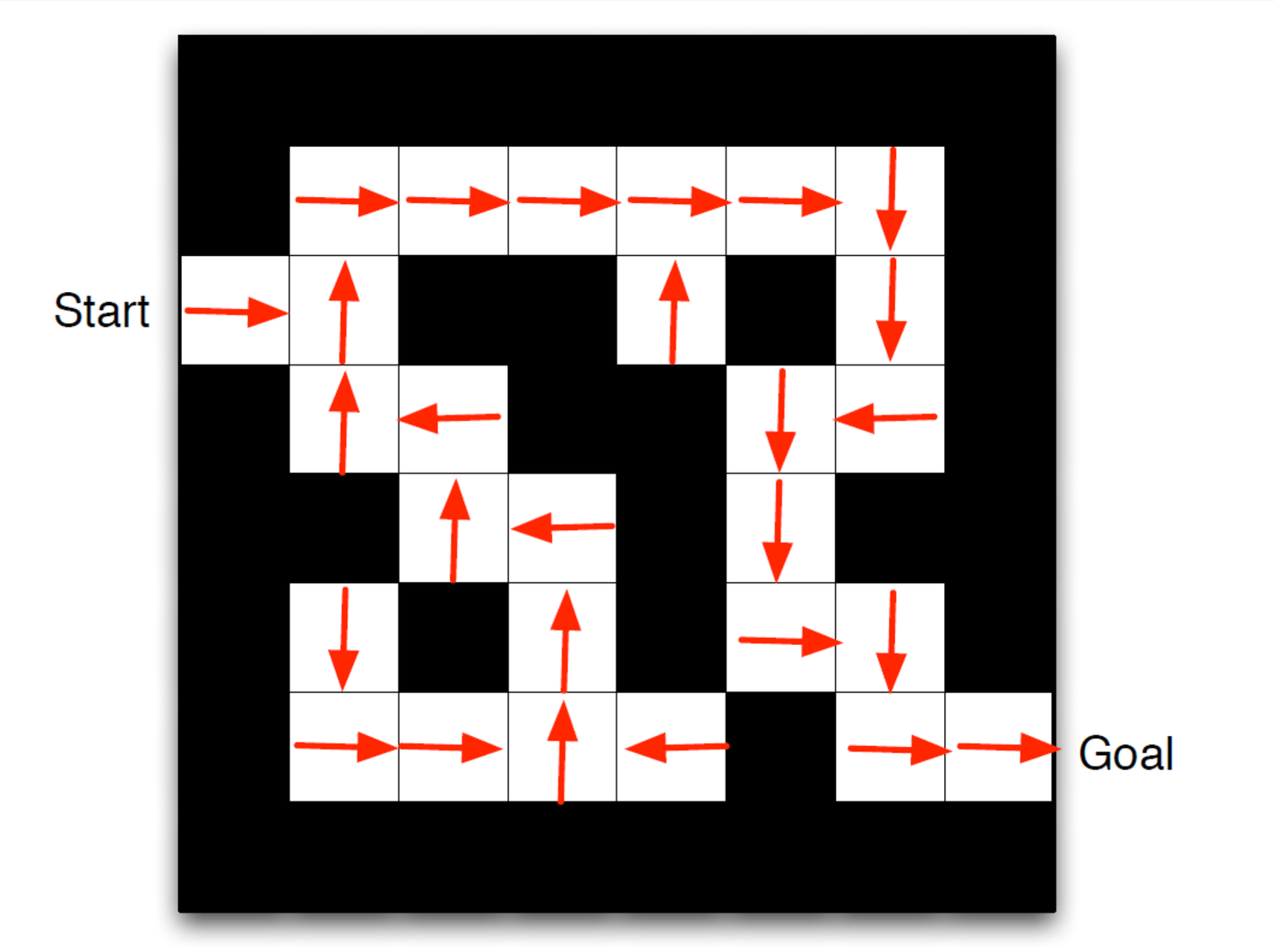
Pick-and-Place Robot

- **Actions:** the voltages applied to each motor at each joint.
- **States:** latest readings of joint angles and velocities.
- **Rewards:** The reward might be +1 for each object successfully picked up and placed. To encourage smooth movements, on each time step a small, negative reward can be given as a function of the moment-to-moment "jerkiness" of the motion

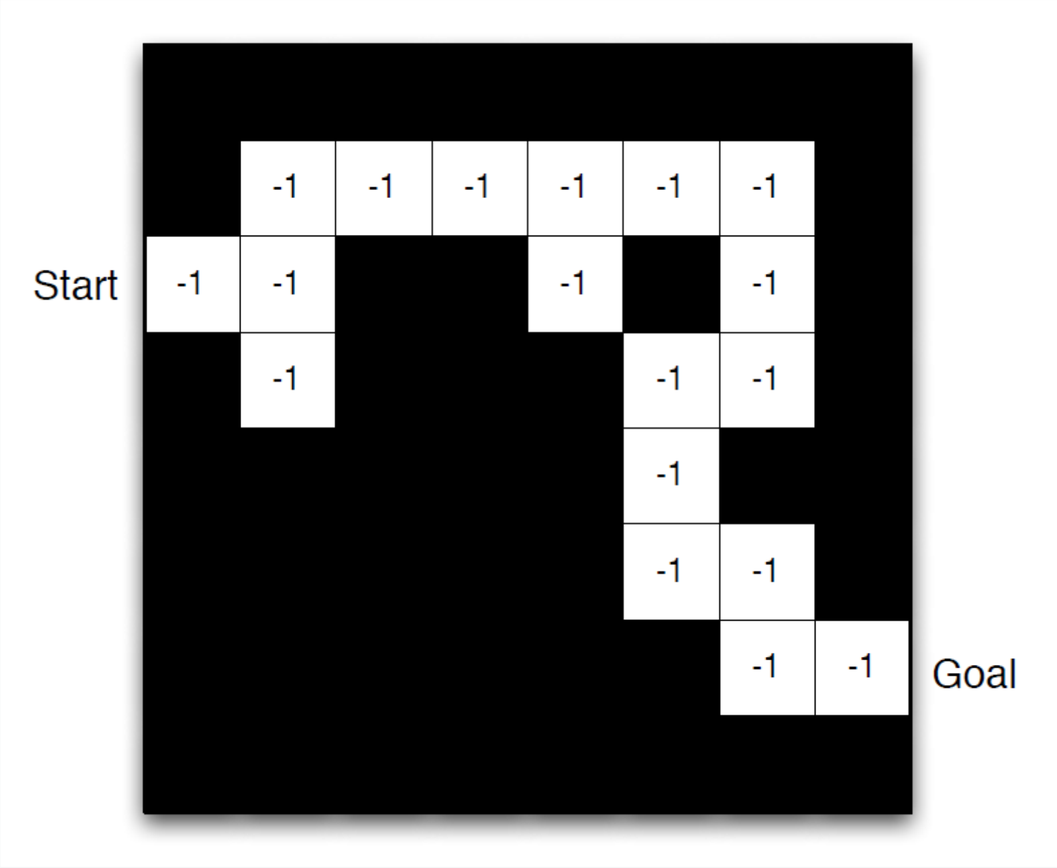
A Maze



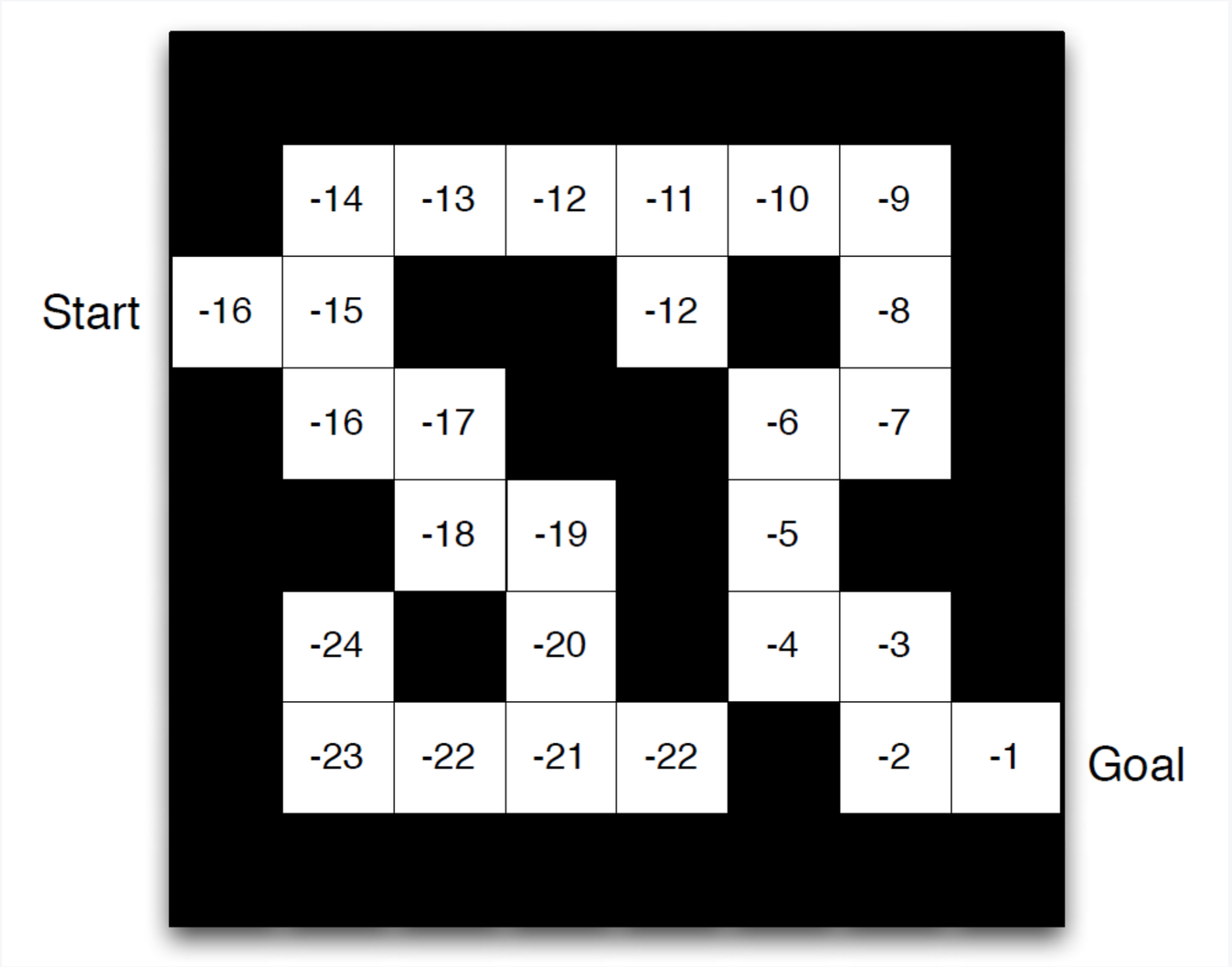
A Maze



A Maze



A Maze


Start
-14
-13
-12
-11
-10
-9
-16
-15
-12
-8
-16
-17
-6
-7
-18
-19
-5
-24
-20
-4
-3
-23
-22
-21
-22
-2
-1
Goal

05.

Difference from Supervised and Unsupervised Learning

Supervised Learning

- Supervised learning is learning from a training set of labeled.
- Objective of SL is for the system to extrapolate, or generalize, its responses to act correctly in situations not present in the training set.
- In interactive problems it is often impractical to obtain examples of desired behavior that are both correct and representative of all the situations in which the agent has to act.

Unsupervised Learning

- Unsupervised learning is typically about finding structure hidden in collections of unlabeled data.
- Reinforcement learning as a kind of unsupervised learning because it does not rely on examples of correct behavior.
- Reinforcement learning is trying to maximize a reward signal instead of trying to find hidden structure.

Thank You!