

Reinforcement Learning

Muchang Bahng

Spring 2024

Contents

1

2

The following sources were used to make these notes.

1. David Silver's Google Deepmind Lectures on Reinforcement Learning

1

In reinforcement learning, there is no supervisor, only a reward signal. Feedback is also delayed, not instantaneous. Finally, time matters, and data is sequential and *not* iid.

Definition 1.1 (History)

The **history** is a sequence of observations, actions, and rewards.

$$H_t = A_1 O_1 R_1 \dots A_{t-1} O_{t-1} R_{t-1} A_t O_t R_t \quad (1)$$

Definition 1.2 (State)

A **state** is a function of the history.

$$S_t = f(H_t) \quad (2)$$