Data-driven and Learning-based Control

Reward in Reinforcement Learning

Erica Salvato

UNIVERSITÀ
DEGLI STUDI
DI TRIESTE

Table of Contents

The **reward** $r^{(k+1)}$ in RL is a scalar feedback signal $\in \mathbb{R}$ that indicates how well the controller is doing at step k.

The RL aim is to learn a controller π able to maximize the expected cumulative discounted reward

$$\sum_k \gamma^k r^{(k+1)}.$$

Therefore, if we are interested in training a controller for a specific task achievement:

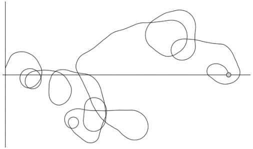
A task should be fully describable through the reward.

Power of reward

Therefore, reward plays a crucial role in RL.

Incorrect reward settings can lead to undesirable task performance behaviors.

- Example 1: Bicycle Riding
  - Positive rewards for progress led to suboptimal behavior (riding in circles).

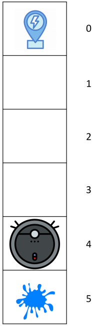


- Example 2: Soccer-Playing Robot
  - Positive rewards for ball touches led to suboptimal behavior (vibrating near the ball).

These examples suggest that shaping rewards must obey certain conditions if they are not to mislead the agent into learning suboptimal policies.

Table of Contents

Sparse rewards are those given to only a small handful of states.



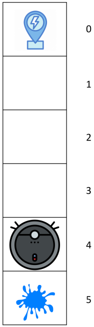
- **Reward function:**

$$\begin{cases} 5 & \text{if } x^{(k)} \neq 5 \text{ and } x^{(k+1)} = 5 \\ 1 & \text{if } x^{(k)} \neq 0 \text{ and } x^{(k+1)} = 0 \\ 0 & \text{otherwise} \end{cases}$$

Dense reward

Dense rewards aim to evaluate the agent in many different states.

$$r^{(k+1)} = -||x^{(k+1)} - x_{\text{des}}||$$



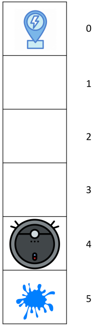
- **Reward function:**
  $r^{(k+1)} = -||x^{(k+1)} - 5||$

Potential-based reward

2 Types of reward

Potential-based reward shaping aims to improve the learning speed of RL agent by extracting and utilizing extra knowledge while performing a task.

$$r'^{(k+1)} = r^{(k+1)} + \gamma\phi\left(x^{(k+1)}\right) - \phi\left(x^{(k)}\right)$$



- **Reward function:**

$$\begin{cases} 5 & \text{if } x^{(k)} \neq 5 \text{ and } x^{(k+1)} = 5 \\ 1 & \text{if } x^{(k)} \neq 0 \text{ and } x^{(k+1)} = 0 \\ -\left(||x^{(k+1)} - 5|| - ||x^{(k)} - 5||\right) & \text{otherwise} \end{cases}$$

Questions’ time!

UNIVERSITÀ
DEGLI STUDI
DI TRIESTE