Multi Armed Bandit

Shuvrajeet Das | Rajshekhar Rakshit

25-01-2022

1 Basics of Reinforcement Learning:

- RL problem mostly comprises of two components:
  - Environment
  - Agent

  Most of the problems we deal with have a single agent. On the other hand, if there is more than one agent ,that problem is called **Multi-Agent RL**, or **MARL** for short. In MARL, the relationship between the agents could be cooperative, competitive or cooperative-competitive.
- The objective of a RL problem is the agent learning what to do in the environment. This defines which **action** to take in the environment.
- The set of all the information that precisely and suffiently describes the situation in the environment is called **State**.
- In some problems, the knowledge of the state is fully known to the agent. In other problems, the knowledge of the state is partially known to the agent. Thus defining the two types of states namely:
  - **Fully Known State**
  - **Partially Known State**

So far, we have not really defined what makes an action good or bad. In RL, every time the agent takes an action, it receives a reward from the environment (albeit it is sometimes zero). Reward could mean many things in general, but in RL terminology, its meaning is very specific: it is a scalar number. The greater the number is, the higher the reward also is. In an iteration of an RL problem, the agent observes the state the environment is in (fully or partially) and takes an action based on its observation.

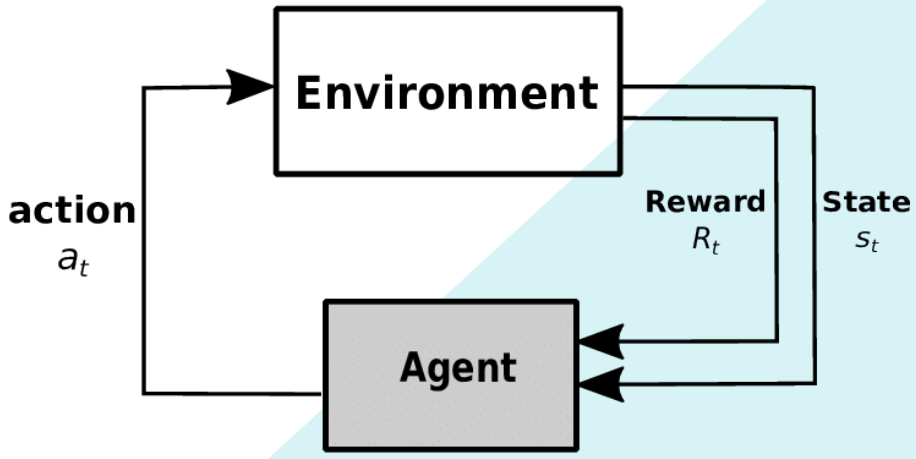

Figure 1: RL process