Essentials of Reinforcement Learning

RL involves an agent acting within an environment.

The environment returns two types of information to the agent:

Reward: This is a scalar value that provides quantitative feedback on the action that the agent took at a timestep t. The agent’s objective is to maximize the rewards it accumulates, and so the rewards are what reinforce productive behaviors that the agent discovers under environmental conditions.

State: This is how the environment changes in response to an agent’s actions. During the forthcoming timestep (t + 1) these will be the conditions for the agent to choose an action in.

Repeating the above two steps in a loop until reaching some terminal state. This terminal state could be reached by, for example, attaining the maximum reward possible, attaining some specific desired outcome, running out of allotted time, using up the maximum number of permitted moves in a game, or the agent dying in a game.

Reinforcement learning problems are sequential decision-making problems.