

Term	Definition
environment	A mathematical model, computer simulation, or real world in which the agent interacts. The environment is described by a current <i>state</i> which is observed by the agent and influenced by the agent's <i>actions</i> . The agent also experiences a reward (cost or benefit) from interacting with the environment, depending on the states and the actions taken.
agent	A software algorithm which generates proposed <i>actions</i> in response to observations of the current <i>state</i> of the <i>environment</i> .
neural network training	Process by which an agent improves its performance through repeated episodes of interaction with the environment. Training effectively allows estimation of parameters describing weights of the agent's neural networks.
tuning	Process by which a researcher selects the hyper-parameters governing the behavior of agent training.
policy function	A function which maps observations of the current <i>state</i> to the <i>actions</i> selected by an <i>agent</i> .
reward	The cost or benefit an agent receives for any interaction with the environment. Note that an agent seeks to maximize the discounted cumulative reward over an episode, and that a reward may be given after every action or potentially only at some end condition (e.g. winning or losing a chess match). Researchers can often improve agent learning by appropriate reward design during training.
value function	3 value functions that are commonly used across RL algorithms: the state-value function, $V^\pi(s_t)$ (expected value of being in a given state), the state-action value function, $Q^\pi(s_t, a_t)$ the expected value of a given state conditional on a certain action, often called the Q-value, and the advantage function, $A^\pi(s_t, a_t)$, the Q-value relative to some baseline.
state	The state of the world (environment) with which the agent interacts. The state can be discrete or continuous and in an arbitrary number of dimensions. An agent may observe the current state completely, or only partially.
action	The space of possible actions the agent can take. Actions can be discrete or continuous and in an arbitrary number of dimensions. Both the state and action space are part of the environment definition – the agent may not necessarily know what the valid range of possible actions is.
hyper-parameters	parameters controlling behavior of the RL algorithm that are fixed over the course of training, e.g. the learning rate, discount rate, or size and type of neural network(s) used by the agent.
model	We use the term “model” to describe a mathematical model of the ecological process (environment). An RL algorithm is considered “model-based” if the agent seeks to ‘learn’ this underlying model, and “model-free” if it does not. Confusingly, some frameworks (stable-baselines) also refer to the agent as the “model” (even when the algorithm being used is model-free).