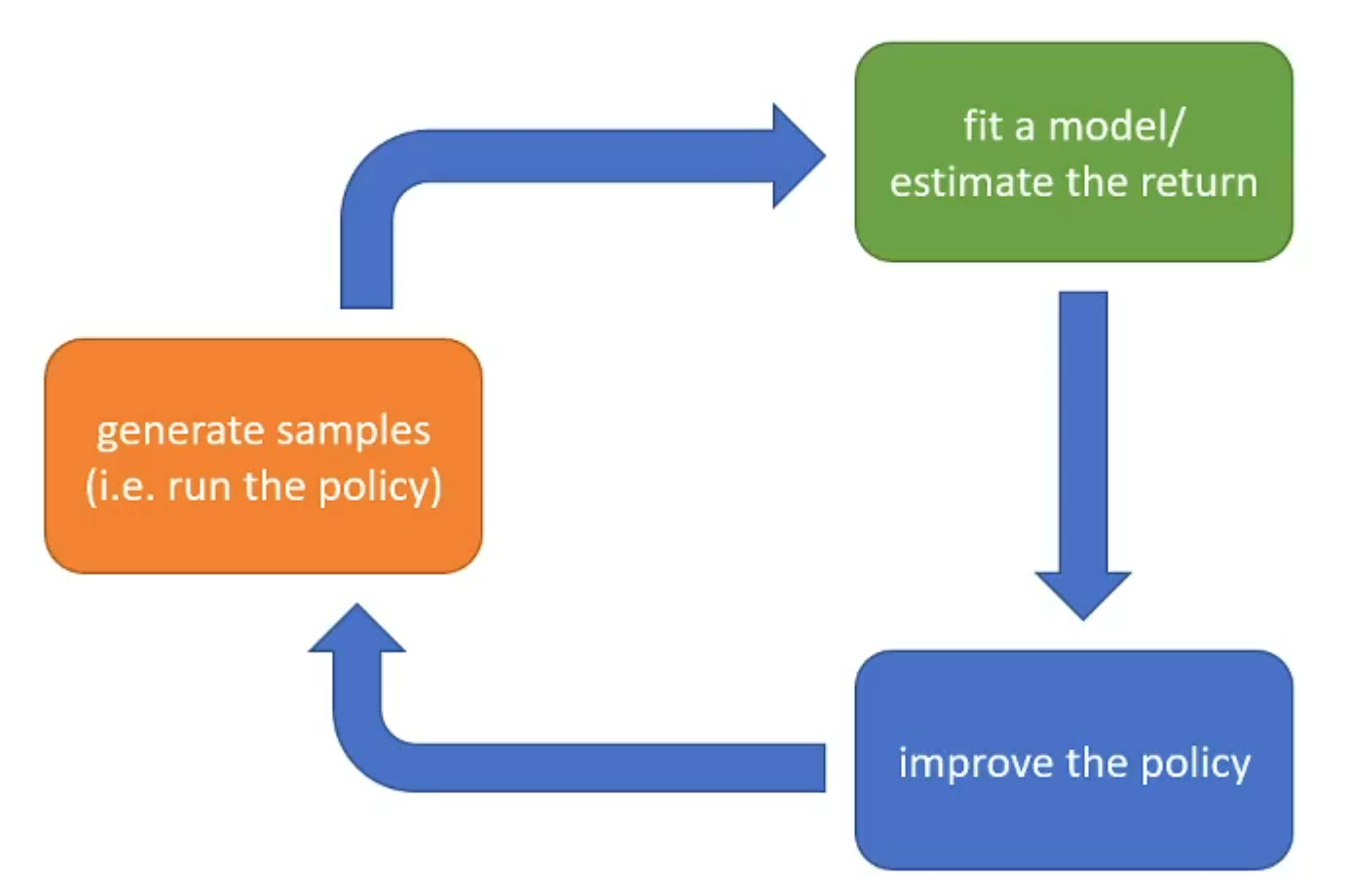
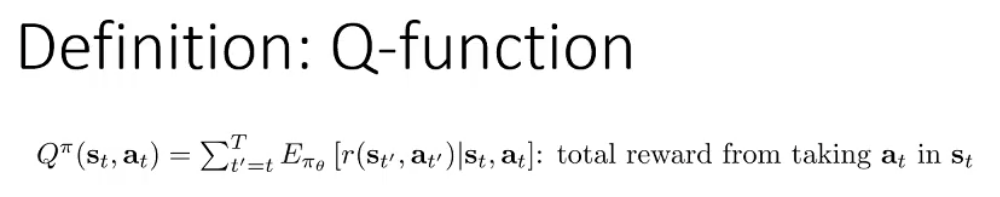
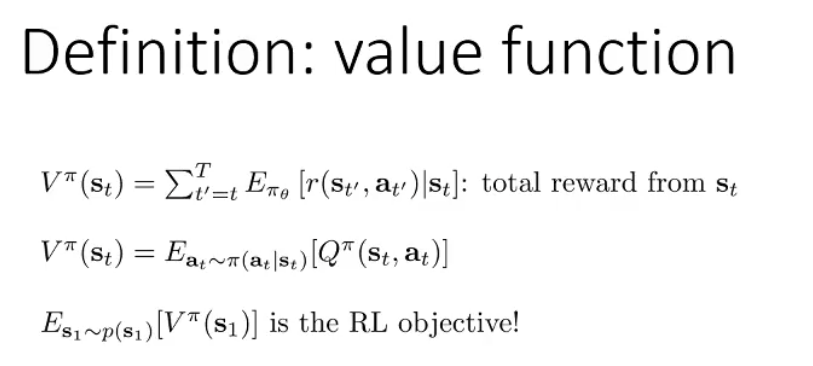
* **Lecture 4 Foundation**

1. A core fact is that the single reward is not smooth but the expectation of rewards is smooth in parameter space.
2. The main modules in RL are generate samples, fit a model/estimating reward and improving the policy.

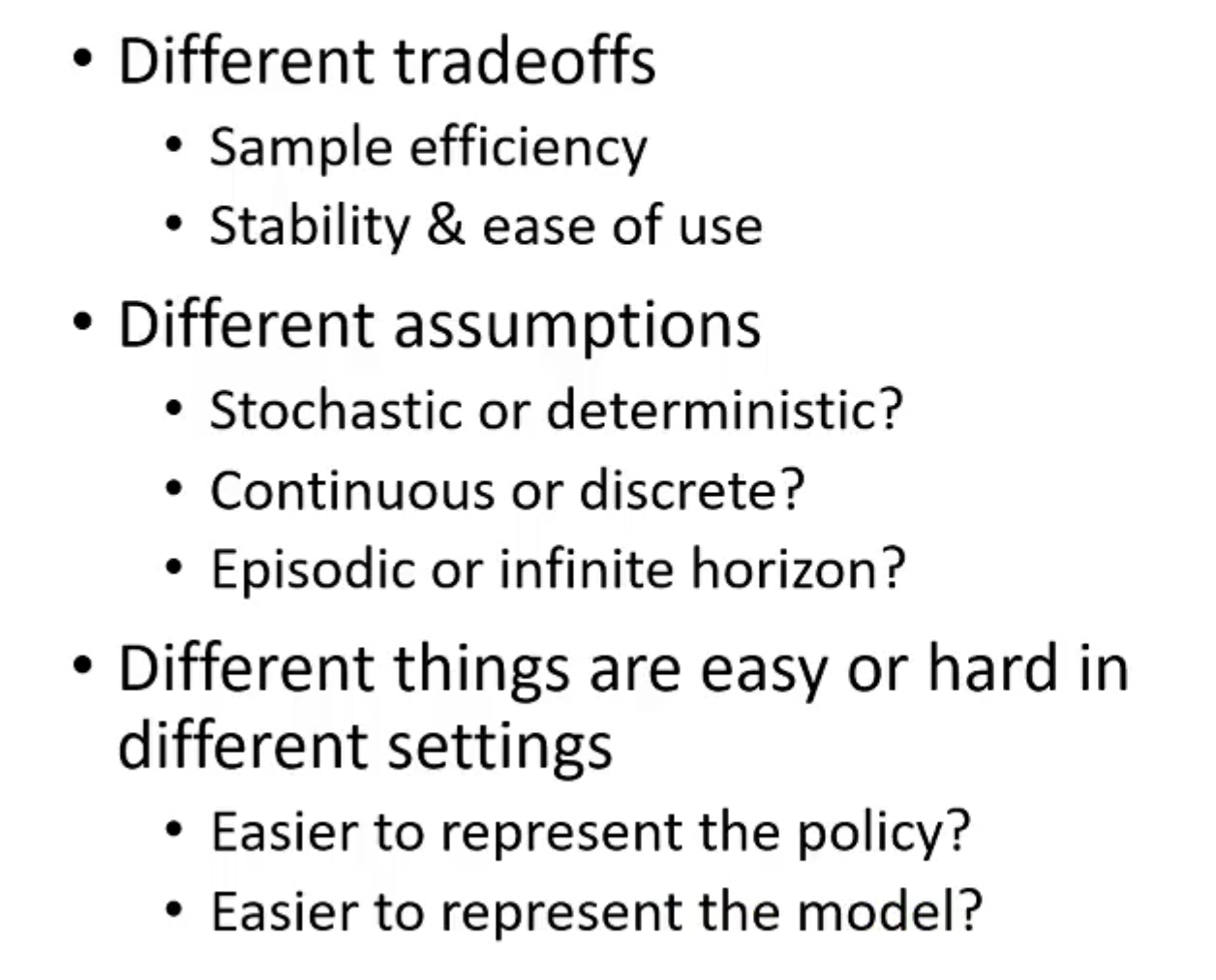


1. Three important functions: Q-function (based on state and action), V-function (only based on state), A-function (advantage, Q - V).





1. RL algorithms mainly conclude policy gradients, value-based, actor-critic and model-based RL algorithm.
2. Some important factors that determine the choice of algorithm:



1. Sample efficiency

