Actions

Environment

Agent
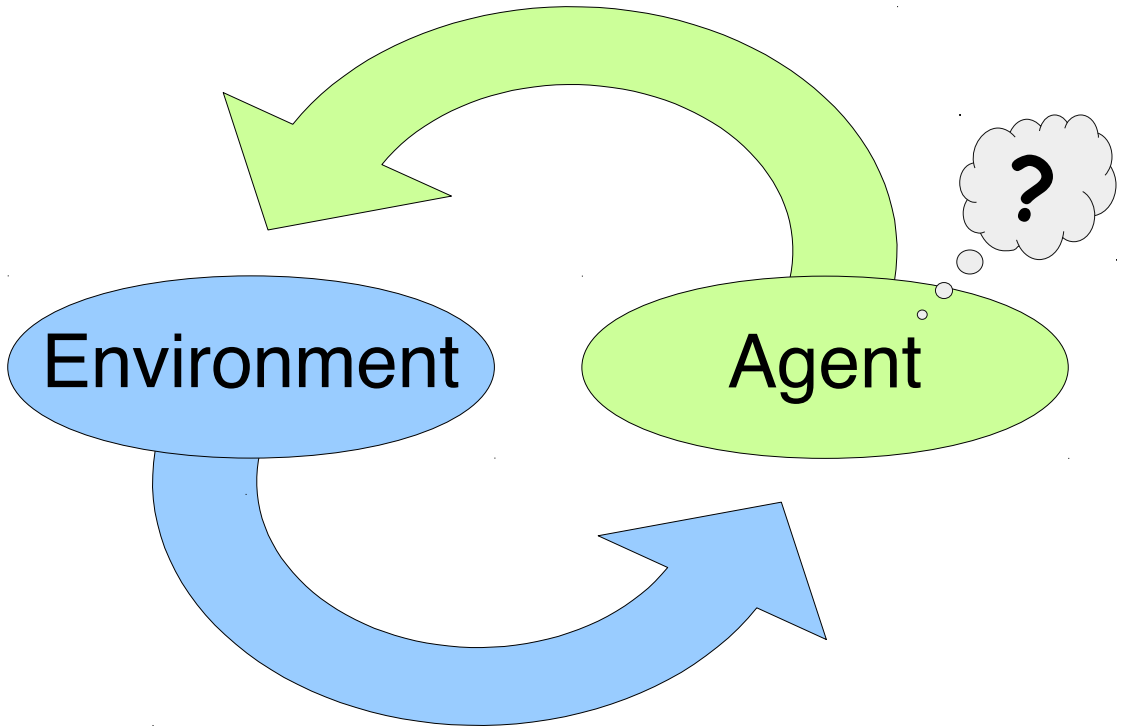
?

Observed states, rewards