

Sequential

Learner
Chose Action
 $A_t \in \mathcal{A}$

Environment
Reveal Reward
 $X_t \in \mathbb{R}$