

强化学习 Reinforcement Learning.

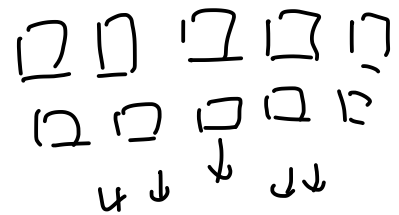
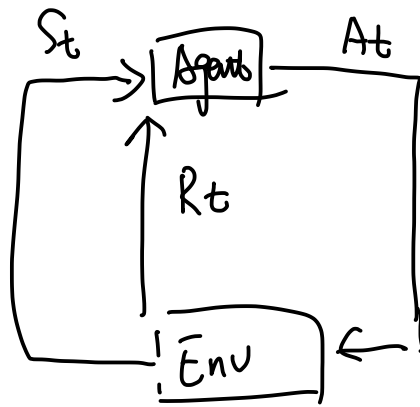
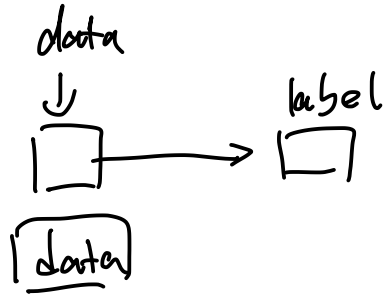
参考: David silver.
Sutton

1.1 ML

① supervised learning

② un

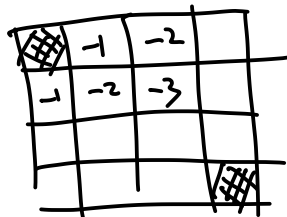
③ RL: learning what to do
map situation to action
to maximize expected Reward



ex <1> 太空侵略

<2> 下棋

<3> 网格世界



- 特点.
- ① no supervisor, only Reward.
 - ②. feedback delay
 - ③ sequential
 - ④ Env is affected by agent

1.2 Rewards

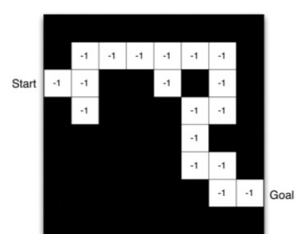
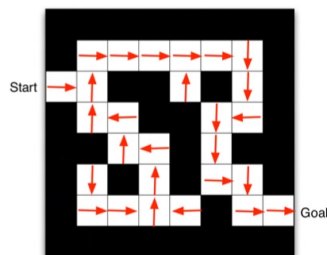
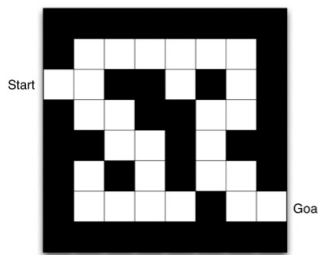
- ① scalar feedback
- ②. Rewards indicate how good at step t
- ③ maximize cumulative Reward.

Def: Reward Hypothesis

All goals can be described by maximizing expected cumulative reward

1.3 要素

- ① 策略
- ② 收益信号
- ③ 价值函数
- ④ 环境模型



价值

