

聊一聊

汽车是怎样学会自动驾驶的？

郭雅欣

第一个问题：人是怎么开车的？

人会考虑的因素

安全：

- 1 红绿灯
- 2 左右交叉等个方向来车
- 3 行人

效率：

最短时间到达

第二个问题:

汽车怎样像人一样自行驾驶？

汽车自动驾驶的一种思路

汽车是否可以从以往的交通事故(经验)中学会如何避免交通事故

或者 通过学习能够更高效地优化它的行车路线？

强化学习

强化学习中有状态(state)、动作(action)、奖赏(reward)这三个要素。

智能体 (Agent) 需要根据当前状态来采取动作, 获得相应的奖赏之后, 再去改进这些动作, 使得下次再到相同状态时, 智能体能做出更优的动作。

Q-Learning

$$Q(\text{state}, \text{action}) = R(\text{state}, \text{action}) + \text{Gamma} * \text{MaxQ}(\text{next state}, \text{all actions})$$

$Q(\text{state}, \text{action})$:

动作效用函数, 用于评价特定状态下采取某个动作的优劣。每个 (state, action) 都有对应的效用值, 所有 (state, action) 的效用值形成了 (state, action) 效用表, 即 Q-table

$\text{MaxQ}(\text{next state}, \text{all actions})$:

记忆中下一个状态 (next state) 的动作中效用值的最大值

Q-Learning

$$Q(\text{state}, \text{action}) = R(\text{state}, \text{action}) + \text{Gamma} * \text{Max}Q(\text{next state}, \text{all actions})$$

$R(\text{state}, \text{action})$:

此动作的立即回报。与Q-table类似，也存在R-table

Gamma :

折扣因子，表示以往经验的重要程度

最终的结果

根据最终形成的Q-table, 选择当前状态下, 值最大的动作。

智能车上的应用

这只是简单的模型。

只考虑下面这些因素：

准备行驶的方向

红绿灯

对向、左侧、右侧方向是否有来车，以及它们即将行进的方向

程序简介

程序会模拟交通规则，根据智能车选择的动作的合法程度，给与奖励或惩罚。

智能车的动作选择 -- 随机 / Q-table中值最大的动作

程序简介

经过数轮模拟(每一轮是指 从发车开始 到到达目的地位置), 可以形成最终的Q-table。之后小车就可以根据该表, 选择当前状态下的最优动作了(当前状态下值效用值最大的动作)。

QA & 谢谢