

# 《亲手用 AI 教会小龟机器狗陪我玩》

--- 强化学习入门实践课

## 第一部分 来自生活的强化学习

**目标：**通过学习小狗表演的案例，对人工智能中的强化学习有个形象理解。

**段 1：从生活案例“教小狗表演欢迎动作”，开始认识强化学习。**

养狗的小伙伴通常都很喜欢教小狗表演欢迎动作，比如：

主人伸出左手，它就滚一圈；

主人伸出右手，它则起立不停拱手拜年；

课堂问题 1：上述动作小狗是不是需要一个动作一个动作来教？

课堂问题 2：如何让小狗记住做出这个动作是对的？

课堂问题 3：我们已经教会小狗对伸出左手的反应，但是发现希望它改成更有趣的反应，怎么办？

**段 2：总结小狗训练的过程**

小狗根据人类的指示，做出对应的动作，主人根据动作完成情况给予奖励。

强化学习本身就来自对动物的训练，那么让我们一比一的认识强化学习里每个概念对应的小狗的哪部分。

强化学习的基本流程：环境告诉智能体当前状态，智能体根据当前状态以随机概率做出反应动作，然后环境中的人类给予智能体评价得分。

## 第二部分 认识强化学习

**目标：**结合第一部分小狗训练的案例，认识人工智能中强化学习的各种基本概念，并对强化学习有个整体理解。

**段 1：和小狗训练案例一比一对应的讲解强化学习的基本流程环和基本概念。**

罗列强化学习基本概念图：智能体(Agent)、环境(Environment)、行动(Action)、奖励(Reward)

环境：在这里人对小狗的触发动作就是环境状态

智能体：在这里小狗就是智能体

动作：在这里小狗对人动作的反应动作就是动作

评价：在这里我们给小狗的奖励就是评价。

## 段 2：形象讲解环境

“环境”更好的表达应该是“环境状态集”，就是能触发智能体或者被用来判断的状态集合。

在小狗训练的案例的是 主人做出的动作。

课堂讨论一：自动驾驶汽车的环境状态集有哪些？

## 段 3：形象讲解动作

“动作”就是智能体在获取环境状态集后，随机做出的反应动作。

小狗案例里就是：翻滚、起立、蹲下、伸手等等

课堂讨论二：自动驾驶汽车的动作有哪些？

## 段 4：形象讲解评价

“评价”就是智能体做出反应动作后，获得的评价奖励得分。

小狗案例里就是：主人对它的食物表扬。

## 段 5：简单讲解算法和 Q-Table

智能体需要有个地方，把每个环境状态做出反应动作的奖励得分记下来，

如果不记下来，那么不然就变成，妈妈口中的“这个人教不起来的”。

这个记下来的地方就叫 Q-Table 它是一个得分表格。

小狗其实也会把这个表格记在脑子里，形成条件反应。

## 段 6：复盘小狗的课堂问题

问题 1 表达 强化学习是为各种环境状态下，寻找最优解。

问题 2 表达 评价

问题 3 表达 强化学习训练其实是在不停寻找最高奖励的过程

# 第三部分 入门实践

**目标：基于小龟机器狗完成一次强化学习训练入门实践。**

## 段 1：介绍小龟机器狗训练平台使用操作

如何启动训练任务

如何对动作评分

## 段 2：开始实践

参考小狗案例，我们训练当用右手去戳机器狗左侧的时候，机器狗表演我们最期望的动作（因

人而异)。细节如下：

- 1、打开平台工具切换小狗到训练状态
- 2、用一根手指头去戳机器狗左侧身体到微微倾斜状态，然后放开机器狗
- 3、等待机器狗做出动作
- 4、在平台工具上对机器狗做出的反应动作打分
- 5、重复训练 10 次，为各个反应动作表演打分
- 6、切换到执行模式，验证效果

## **第四部分 教会我的机器狗陪我玩**

**目标：**参考小狗表演案例，假设机器狗一个性格特征，并通过各种触碰手势动作对其训练出性格表达。

**实践：训练机器狗**

- 1、设定机器狗的性格：凶凶的狼狗、粘人的贵宾、傻傻的哈士奇等等。
- 2、至少完成从五个角度触碰机器狗的反应。
- 3、成果展示。

## **作业 让机器狗完成一段连续舞蹈**

**目标：**通过教会小龟机器狗跳一段你觉得酷炫的舞蹈，认识强化学习也能完成类似走迷宫一样的全路径最优解答。

认识强化学习的连续性模式，为每个当前表演动作选择一个最优后续表演动作。