

《人工智能导论》

主讲老师：李蓝天

课程内容

1. 神奇的人工智能
2. 认识你的脸
3. 倾听你的声音
4. 模仿你的行为

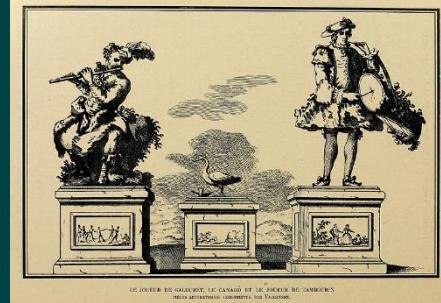


创造机器的梦想



1495年，列奥纳多·达·芬奇设计了仿人型的机械骑士，可以完成挥动手臂、摇头等简单动作。

—— 达芬奇机器人



1737年，雅克·迪·沃康松发明了一系列自动演奏机器，可以吹笛子和敲鼓。



18世纪末，田中久重发明了一系列复杂的玩偶，可以奉茶、写字。

梦想成真

机器人经历了模仿、感知和智能的进阶，如今已成为人类的伙伴。

人机对话



食品加工

消防灭火



自动驾驶

围棋高手 AlphaGo

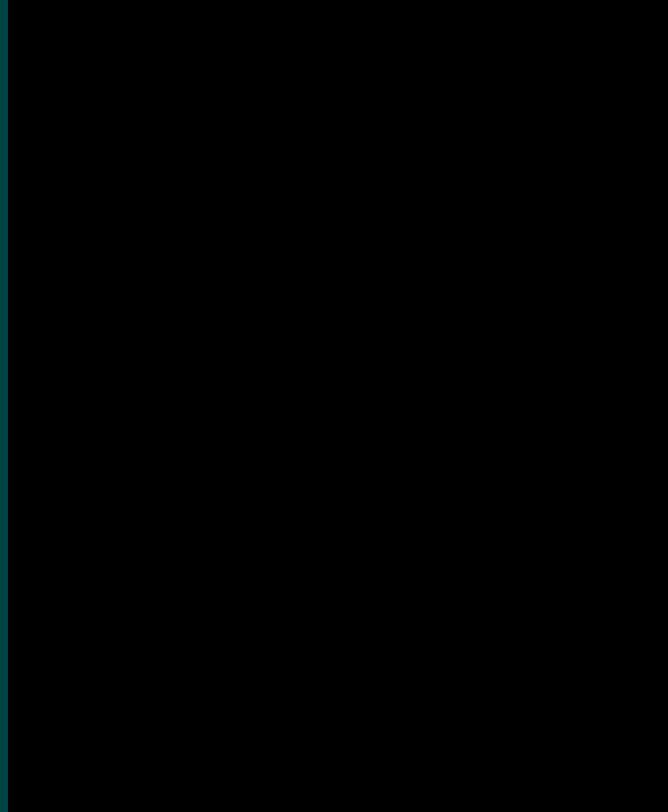
2015年至2017年，AlphaGo 机器人先后战胜了樊麾、李世石、柯洁等世界一流围棋高手。



游戏玩家

2013年11月，初创人工智能公司 DeepMind 开发了计算机玩壁球游戏。

两个月之后，谷歌花了6亿美元收购了 DeepMind 公司。



基于设计的机器人

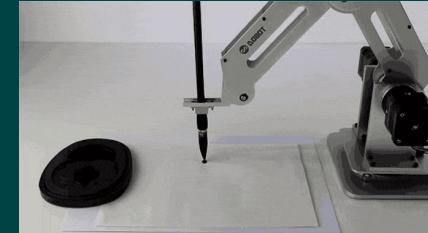
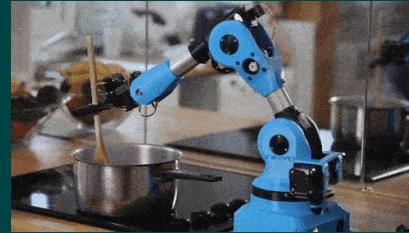
基于学习的机器人

基于设计的机器人

机器人按行为分类

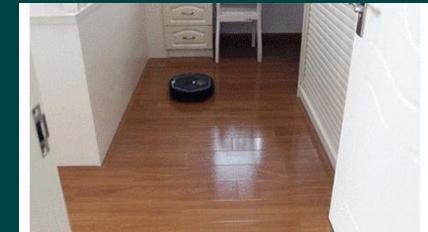
- 操作机器人

- 关注如何通过关节运动完成动作
- 机械臂



- 移动机器人

- 关注如何规划路线以到达目的地
- 无人驾驶汽车

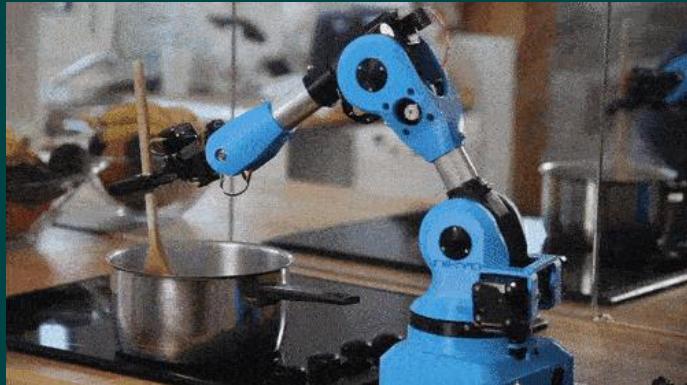


操作机器人

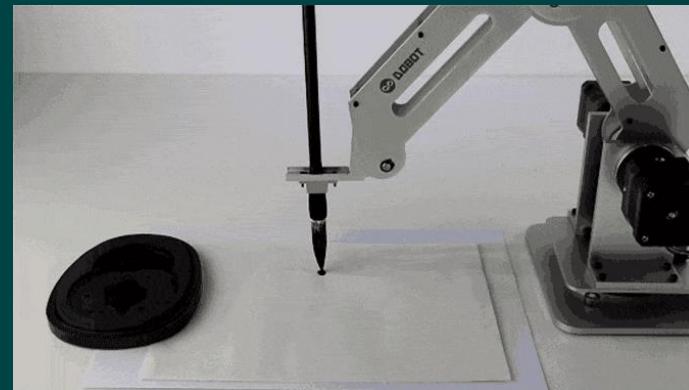
目标：完成某种指定的动作，如抓取、焊接等。

关键问题：如何将机械臂的末端（工作点）放置到预设位置。

背景知识：运动学、动力学、控制论



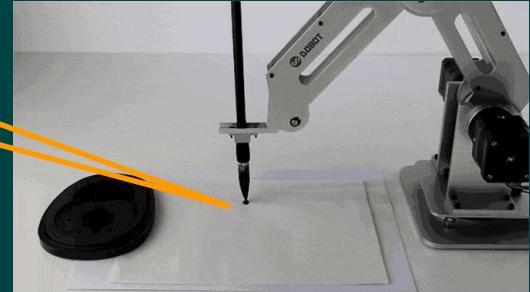
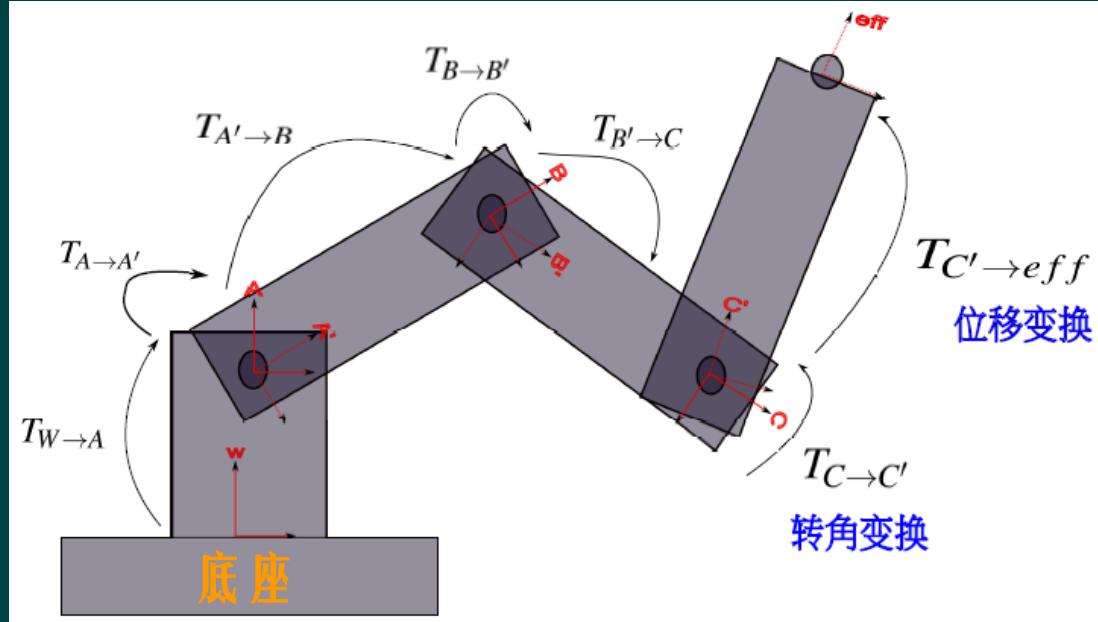
抓住汤勺，搅拌汤汁



抓住毛笔，撰写文字

连接系统

末端（工作点）



1. 如何设置连接杆的夹角，使末端放到指定位置？
2. 为了实现1，如何对各个连接杆施力？
3. 如何监控夹角的改变过程，实时调整施力？

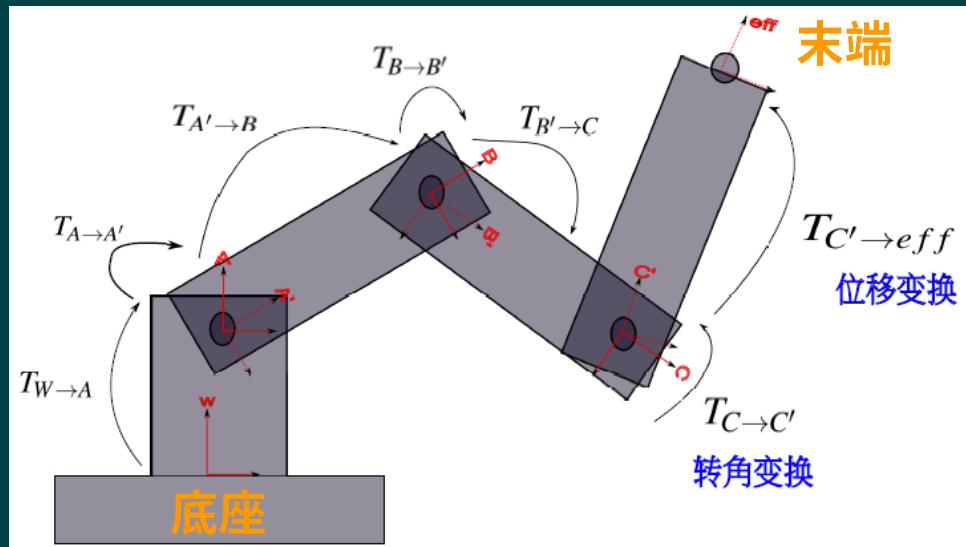
问题1：基于运动学计算目标夹角

位形：所有夹角的一组取值称
为一个位形

变换：一个位形决定了底座到
末端的变换

$$T_{W \rightarrow eff} = T_{w \rightarrow A} T_{A \rightarrow A'} \dots T_{C \rightarrow eff}$$

总结：计算一个位形，依此位
形可将末端放置于目标位置



问题2：基于动力学设计力矩

理论基础：牛顿力学

处理步骤

- 设计加速-减速（角加速度）方案，完成从起点到终点的变化；
- 基于角加速度，计算每一时刻施于连接杆（手臂）上的力矩；
- 基于连接杆（大臂和小臂）之间的连接关系，计算对每个连接杆所施加的力矩。



问题3：基于控制论进行过程调节

误差

- 质量分布、重力、摩擦力等
- 影响动作精度

反馈机制

- 加传感器监视动作完成情况
 - 压力传感器应用于擦玻璃机器人
 - 视觉传感器应用于焊接机器人
- 基于前一个时刻的位置和速度，调整当前时刻的力矩大小



移动机器人

任务

- 实现有效的移动
- 对环境建立地图，并估计自身位置



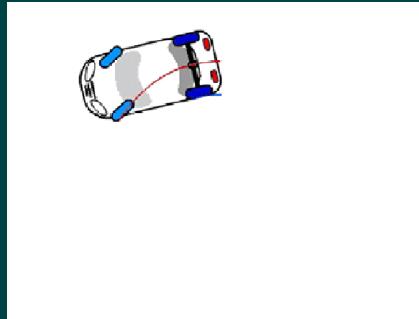
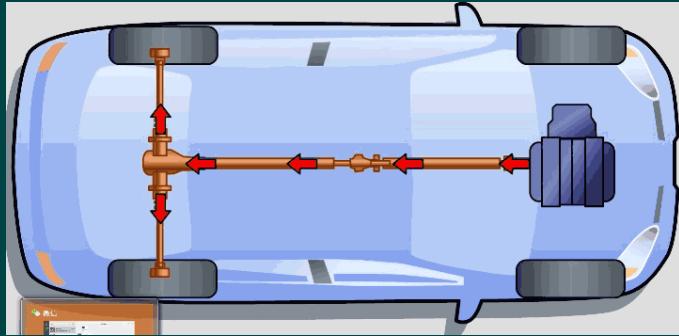
无人驾驶



扫地机器人

有效移动

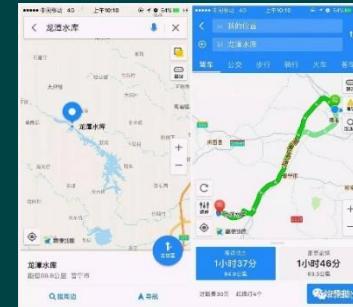
移动是有条件的



环境建立地图，估计自身位置

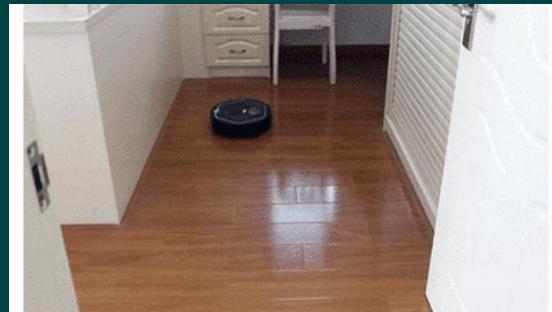
方式一：人为预先指定机器人运动路线，并记录下所见的场景

- 自动驾驶汽车

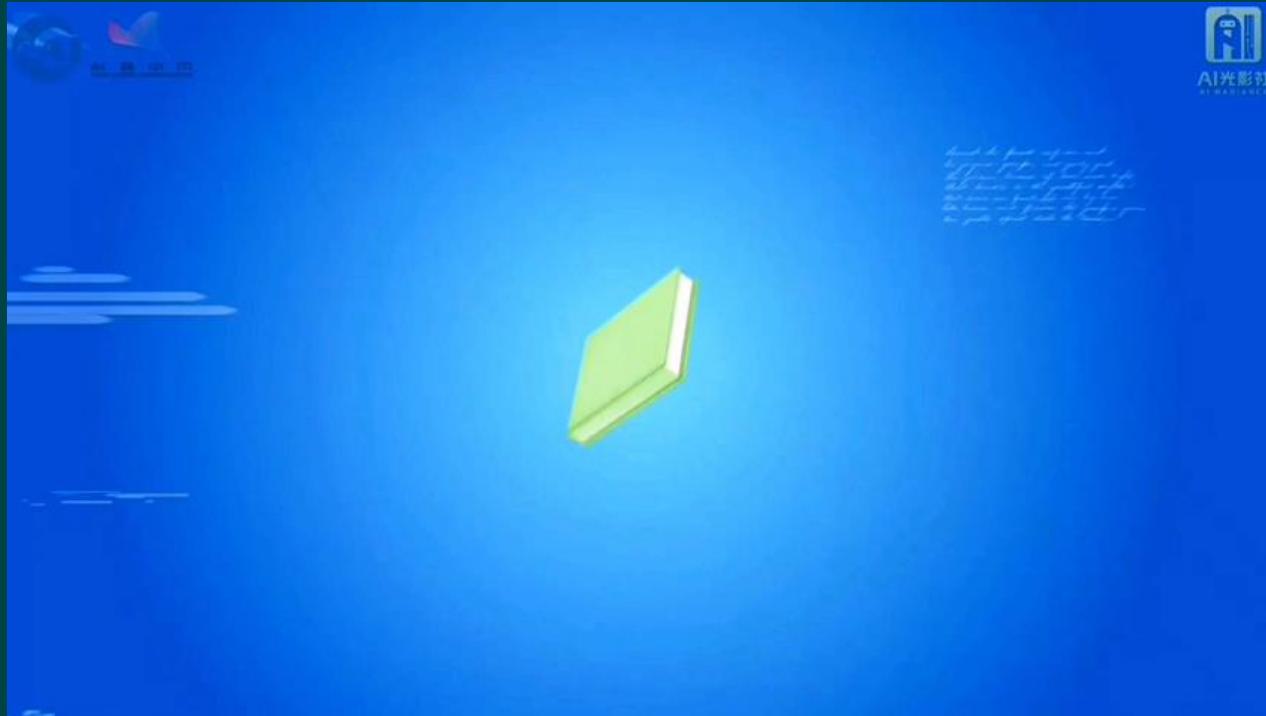


方式二：同步定位与地图构建

- 家用扫地机器人



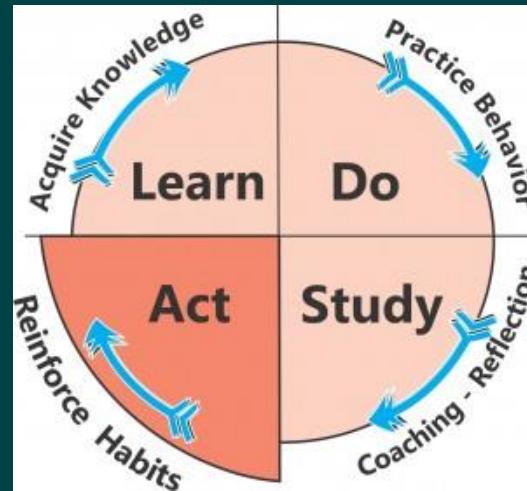
AI光影社：扫地机器人是如何工作的



人是怎么学习的？



制定目标



不断尝试

婴儿学习走路

学会走路（目标）

不断地尝试迈步、抬腿（动作）

当偶尔站立或扶着挪动时，得到糖果（奖励）

当姿势错误摔倒了，感受到疼痛（惩罚）

父母会适当给出一些手势和引导（观察）

最后学会站立和迈步的技巧，不再摔倒



儿童学习骑脚踏车

从 A 顺利地骑到 B (目标)

不断地尝试脚蹬车前行 (动作)

平稳正确地前行 (奖励)

摔倒、停滞不前 (惩罚)

看其它小伙伴们们的骑行 (观察)

最后学会骑行的技巧，不再摔倒



学生学习知识

学会知识（目标）

不断地学习（动作）

学的好，分数高（奖励）

学不会，分数低（惩罚）

与老师同学交流（观察）

最后掌握高效学习知识的方法



人类学习方法：强化学习

人类学习是通过**目标**、**动作**、**奖励**、**惩罚**、**观察**等来实现的。

这类学习方法通过主动和环境进行**交互**，利用环境给出的**反馈**（奖励或惩罚）学习行为策略，使得最终**收益最大化**，称之为**强化学习**。

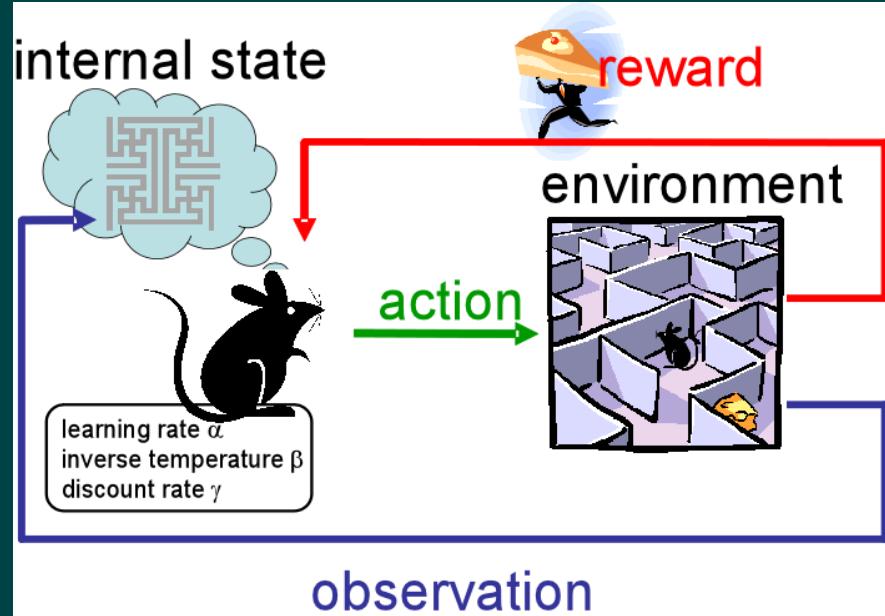
强化学习是人类的学习方式，将其应用至机器人学习中。



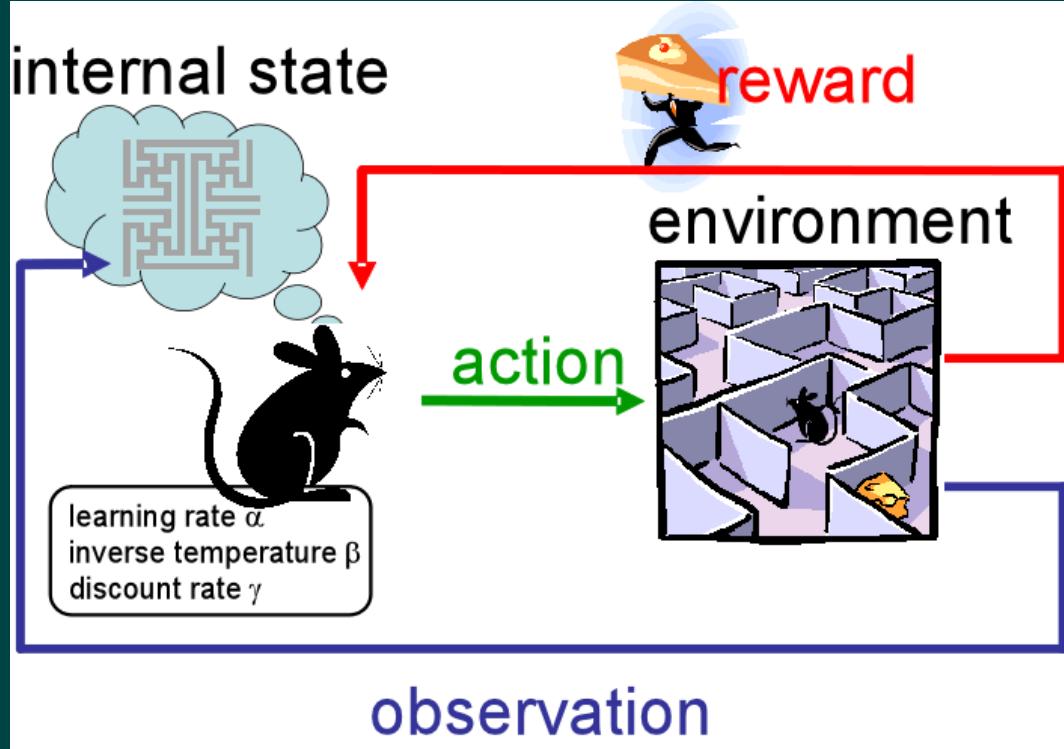
什么是强化学习？

强化学习是指在一系列环境中，通过多步恰当的决策来达到一个目标的学习过程，是一种序列多步决策的问题。

强化学习的目标就是要寻找一个能使得我们获得最大累积收益的策略。



什么是强化学习?



环境：老鼠所处的位置

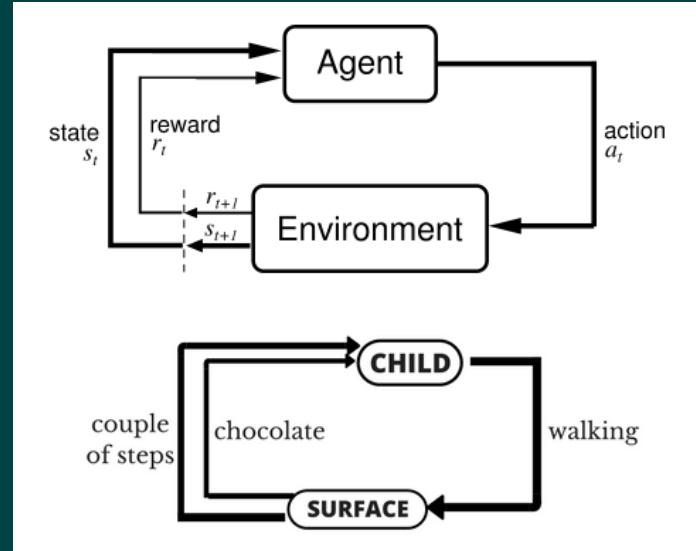
动作：老鼠行进的方向

反馈：吃到奶酪加分；没
吃到奶酪减分

策略：学习如何选择动作，
更容易吃到奶酪

强化学习的特点

以最终目标为学习动力，而非每一步目标，目光长远
等价地，学习带有很强的时序性，行为后果具有时序滞后性
决策会影响环境，进而影响后序决策



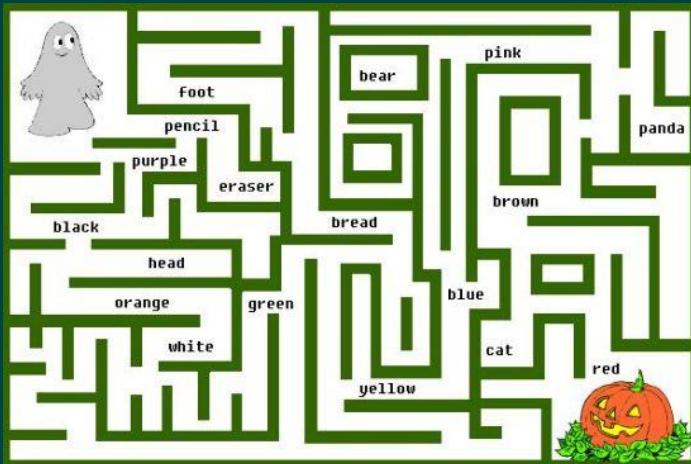
强化学习的分类

多轮任务：成本低、时间短，收益最后显现

- 象棋、围棋等游戏

连续任务：成本高、时间长，奖惩随时可见

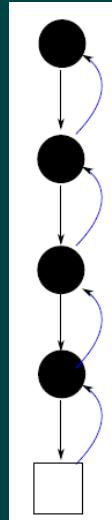
- 股票操盘、行走机器人



强化学习的方法

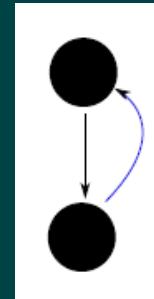
蒙特卡洛算法：完整模拟一次任务执行过程，通过任务结束后的收益调整行为策略，用于多轮任务。

一系列行为
(全路径)



时序差分算法：基于当前行为策略生成一个动作，这个动作通常会立即得到一个反馈，基于这一反馈对行为策略进行即时调整，用于连续任务

一步或多步
行为



强化学习的应用（1）

机器人：自主学习如何行动，如何完成制定目标



Boston Dynamics



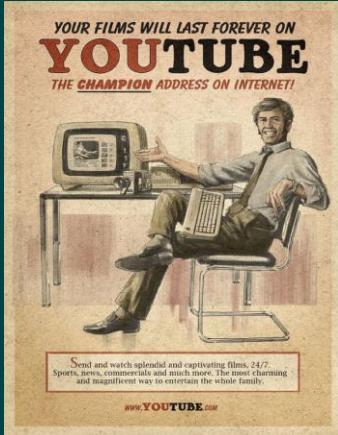
Meet Atlas, the latest humanoid robot from
Alphabet Inc.'s Boston Dynamics.

WSJ

强化学习的应用 (2)

生产与生活

- 如何摆放食品销量最大?
- 如何投放广告吸引更多客户?
- 如何配置信号灯疏导交通?



强化学习的应用（3）

REINFORCEMENT LEARNING

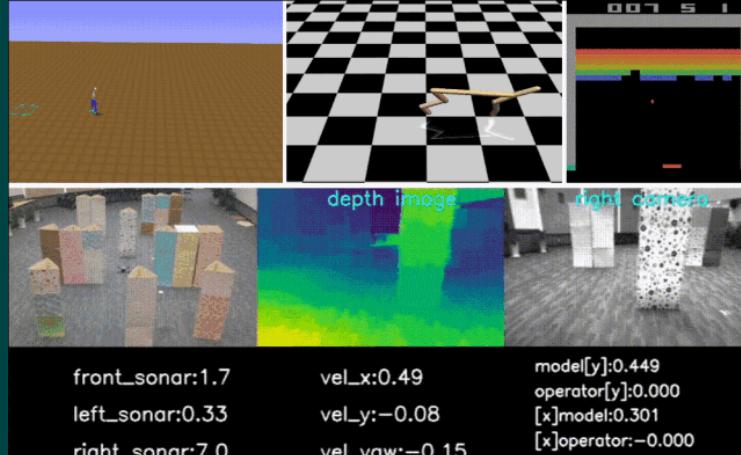
Before training
peaceful swimming

复杂环境下的强化学习

强化学习需要对环境状态进行判断

在大量实际任务中，环境输入的原始信号非常复杂

- 摄像头图片、雷达反馈信号；围棋棋盘

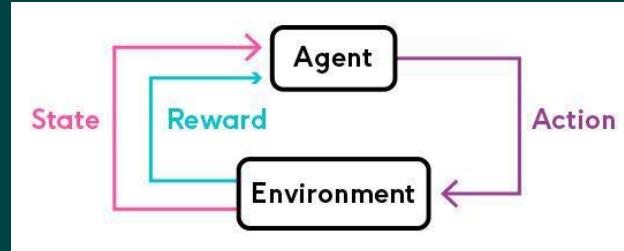
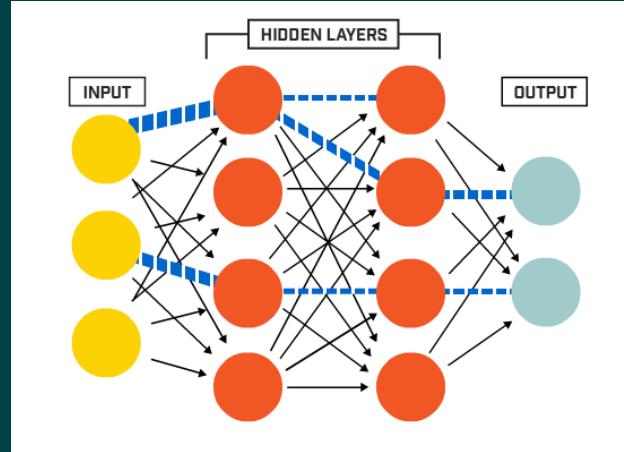


深度强化学习

第一部分：深度学习，将原始感知信号逐层处理，提取环境状态。

第二部分：强化学习，利用强化学习目标更新深度神经网络参数。

深度强化学习是当前机器学习乃至整个人工智能领域研究的热点之一，推动了近年来人工智能的热潮。

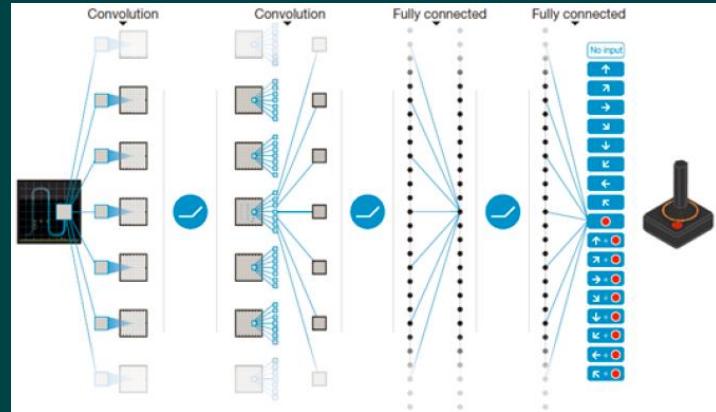


Atari 游戏

DeepMind 利用深度 Q-learning 网络 (DQN) 教会机器玩游戏

输入为游戏画面，经过多层卷积神经网络 (CNN) 来提取状态信息，得到状态空间

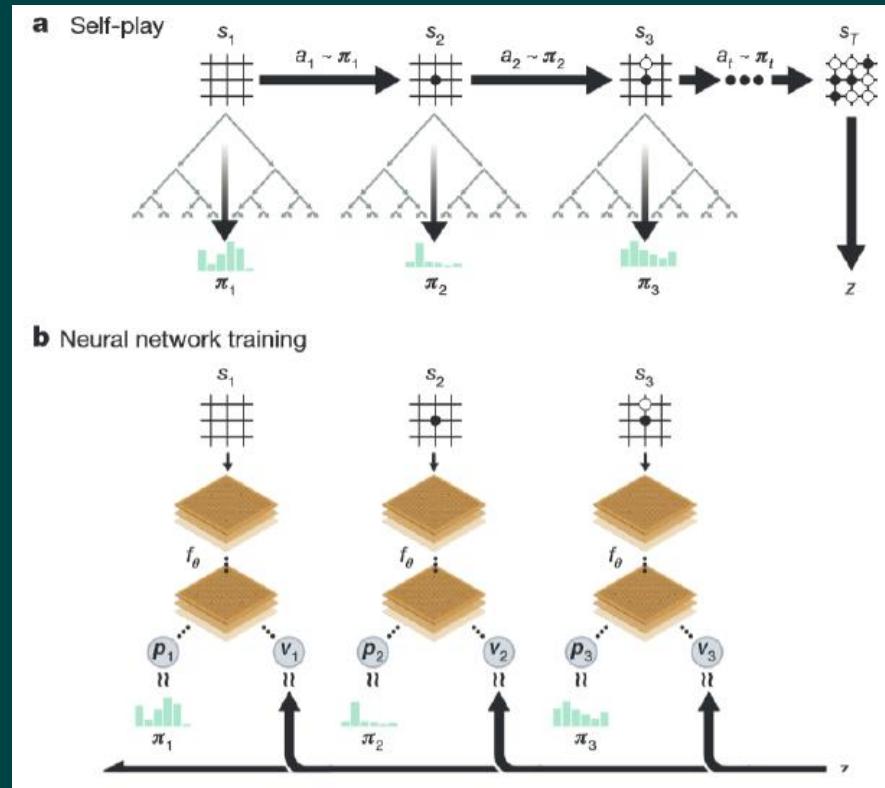
基于该状态空间，学习不同状态下采取某一操作的价值，摇动手柄



AlphaGo Zero

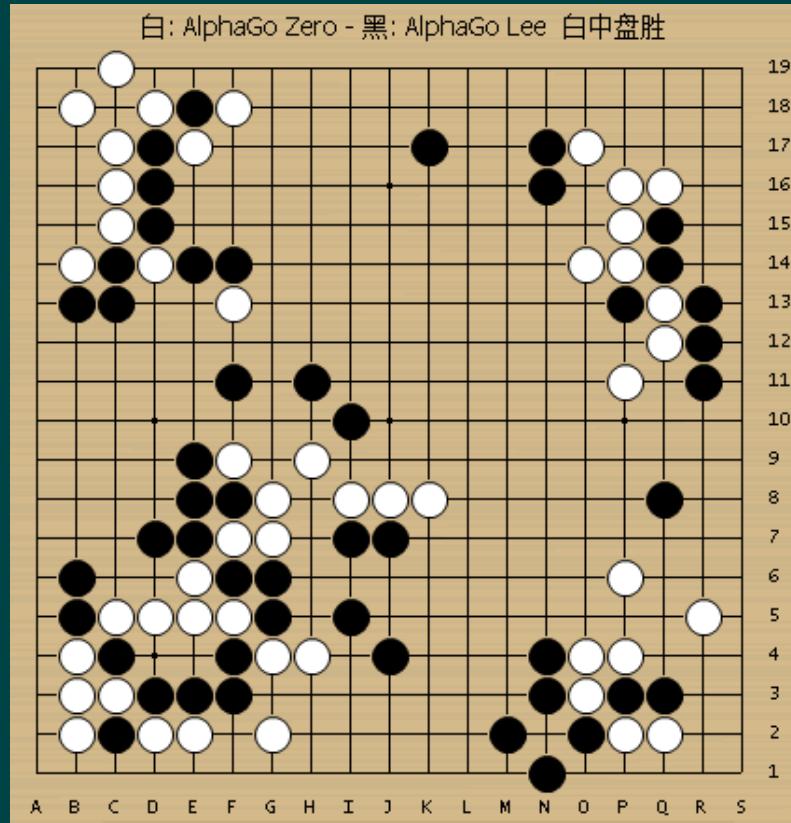
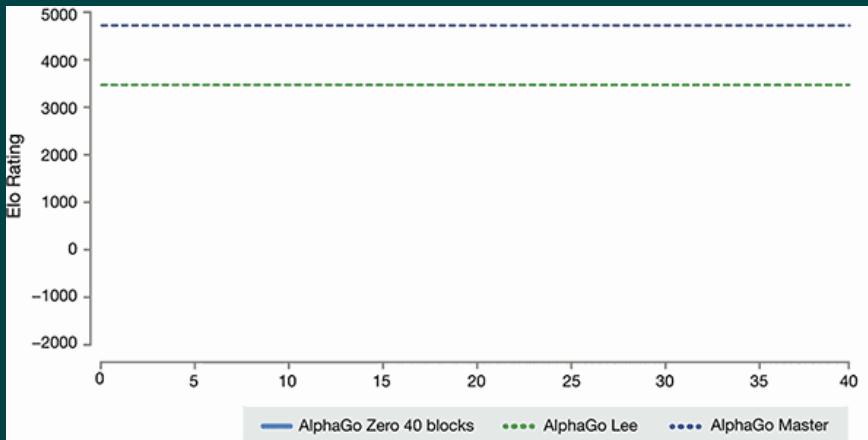
不再学习人类的棋谱，而是完全依赖深度强化学习，通过自我对弈学习机器自己的走棋方式。

卷积神经网络描述棋局状态，蒙特卡洛搜索模拟走棋过程，反馈不同走棋的输赢概率。



AlphaGo Zero

使用 64 个 GPU 和 19 个 CPU，用
三天时间完成了自我对弈 490 万局，
成为绝世高手！

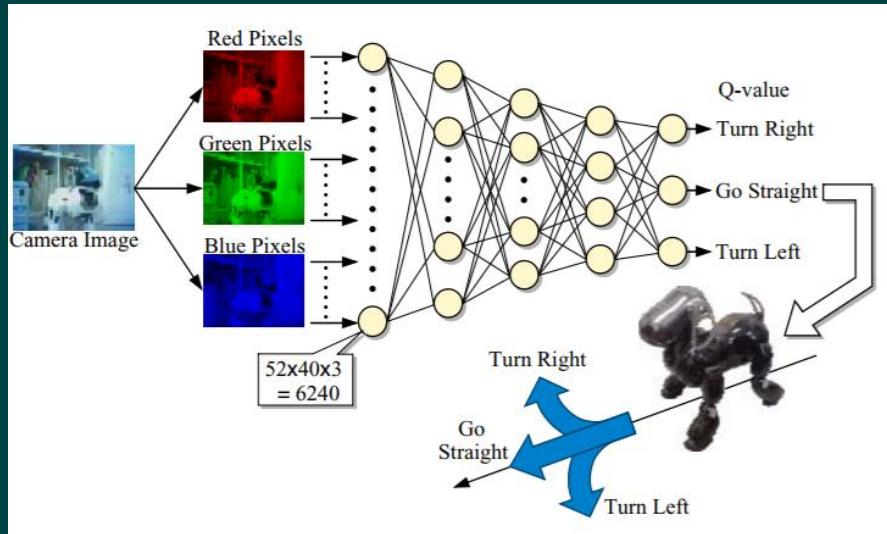


AIBO 接吻机器狗

初始状态



目标状态



AIBO 接吻机器狗

Kissing AIBO Task (2008)

The world's first end-to-end
Deep Reinforcement Learning (DQN)
applied to a Real Robot

自动驾驶汽车

输入：摄像头在一段时间内拍摄的图像以及采取的动作

输出：该动作可能产生的后果，包括发生碰撞的概率和系统的安全状态

Self-supervised Deep Reinforcement Learning with
Generalized Computation Graphs for Robot Navigation



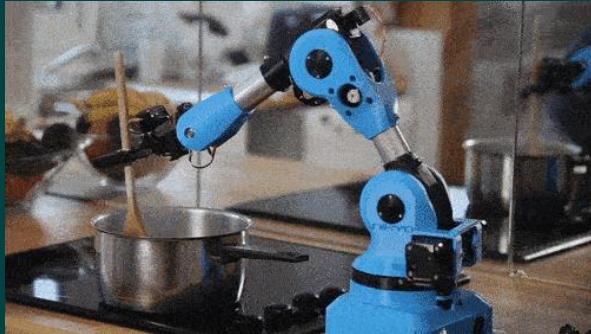
Gregory Kahn, Adam Villaflor, Bosen Ding, Pieter Abbeel, Sergey Levine



两种机器人的比较

基于设计的机器人

- 人为设计，安全稳定；易解释、可预判
- 让它干啥就干啥，复杂一点就白搭 😞



基于学习的机器人

- 行为方式来源于外界经验，需要大量尝试
- 制定目标，自主学习，可适应复杂环境



机器人与伦理

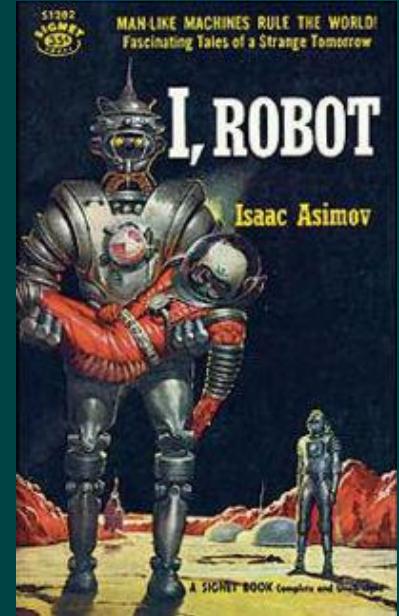
机器人是人工智能技术走向现实世界的载体。

如何规范机器人的行为？

科幻小说家艾萨克·阿西莫夫在小说《我，机器人》

(I, Robot) 中订立了著名的“机器人三定律”：

- 机器人不得伤害人类，且确保人类不受伤害；
- 在不违背第一法则的前提下，机器人必须服从人类的命令；
- 在不违背第一及第二法则的前提下，机器人必须保护自己。



AI光影社：无人车撞了人，谁来负责？



谢谢观看！