



人工智能原理与算法

第2章-智能体

中国科学院自动化研究所
国科大人工智能学院

雷 震

智能体 (Agent)

- 本章讨论智能体的本质，智能体是否完美，环境的多样性，及由此带来的各种智能体分类。

- 智能体和环境
- 良好行为：理性的概念
- 环境的本质
- 智能体的结构

- 智能体和环境
- 良好行为：理性的概念
- 环境的本质
- 智能体的结构

2.1 智能体和环境

- 任何通过传感器感知环境并通过执行器作用于该环境的事物都可以被视为智能体。

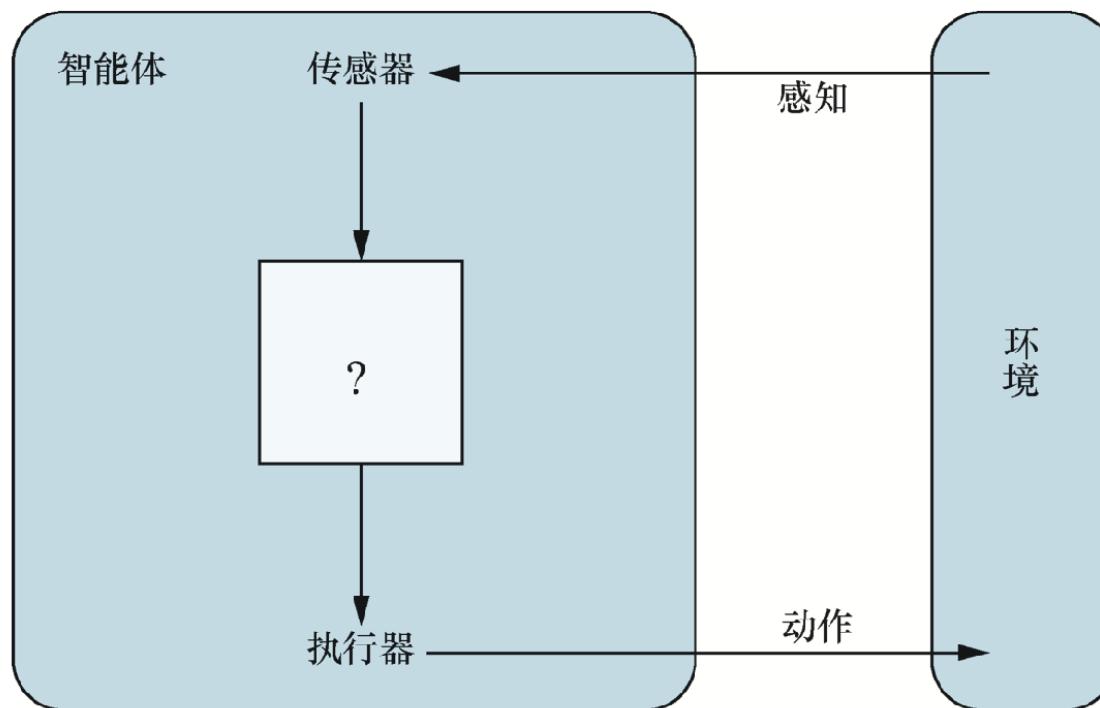


图 2-1 智能体通过传感器和执行器与环境交互

2.1 智能体和环境

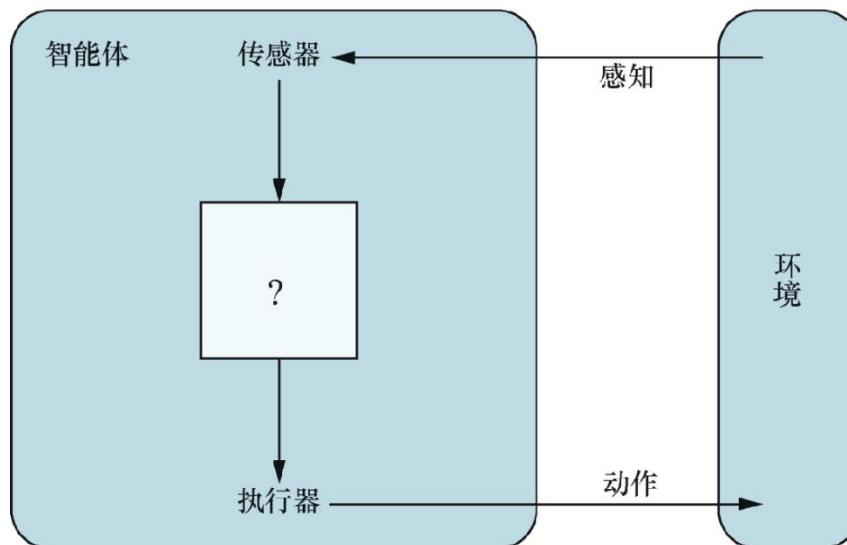


图 2-1 智能体通过传感器和执行器与环境交互

□ 人类智能体

- ▶ 传感器：眼睛、耳朵、其他器官
- ▶ 执行器：手、腿、声道

□ 机器人智能体

- ▶ 传感器：摄像头、红外测距仪
- ▶ 执行器：各种电动机

2.1 智能体和环境

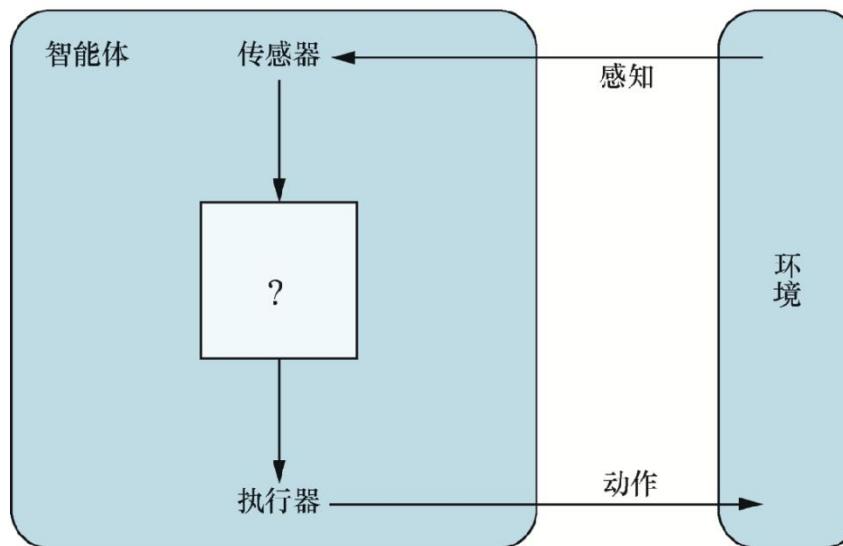


图 2-1 智能体通过传感器和执行器与环境交互

□ 软件智能体

- ▶ 传感器：键盘敲击、文件内容、网络数据包
- ▶ 执行器：屏幕显示、写文件、发送网络数据包

2.1 智能体和环境

- 感知：表示智能体的传感器正在感知的内容
- 感知序列：智能体所感知的一切的完整历史
- 一般地，一个智能体在任何给定时刻的动作选择可能取决于其内置知识和迄今为止观察到的**整个感知序列**，而不是它未感知到的任何事物
- 通过为每个可能的感知序列指定智能体的动作选择，则可以说我们了解了关于该智能体的所有内容

2.1 智能体和环境

- 从数学上讲，我们说智能体的行为由智能体函数描述，该函数将任意给定的感知序列映射到一个动作。
- 通过制成表格来描述任何给定的智能体及其函数
 - 对大多数智能体来说，这将是一个非常大的表，事实上是无限的（除非限制考虑的感知序列长度）。
 - 原则上，我们可以通过尝试所有可能的感知序列并记录智能体响应的动作来构建此表。该表是该智能体的外部特性。
 - 在内部，人工智能体的智能体函数将由智能体程序实现。

2.1 智能体和环境

■ 智能体函数和智能体程序

- 智能体函数是一种抽象的数学描述
- 智能体程序是一个具体的实现，可以在某些物理系统中运行

2.1 智能体和环境

■ 示例：真空吸尘器世界

- 两个地点：方格A和B
- 真空吸尘器智能体可以感知它在哪个方格中，以及方格中是否干净
- 可选的操作：向右移动，向左移动，吸尘或什么都不做

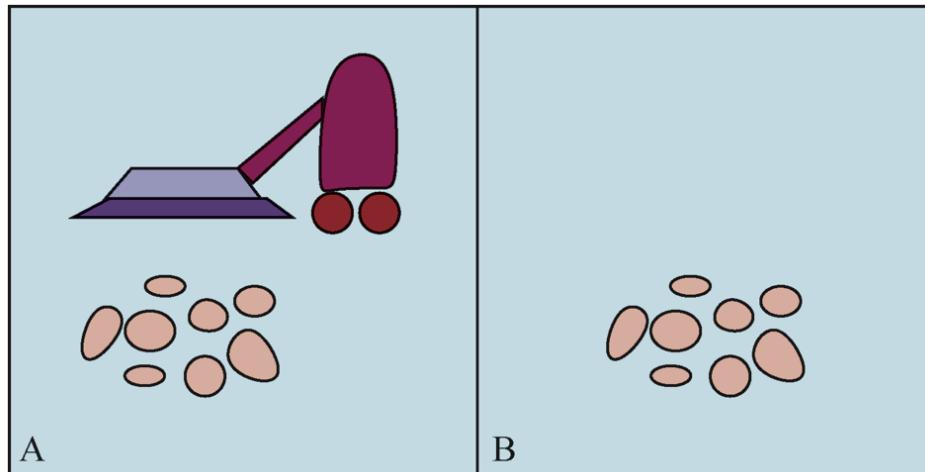


图 2-2 一个只有两个方格的真空吸尘器世界。每个位置可以是干净的，也可以是脏的，智能体可以向左移动或向右移动，可以清理它所占据的方格。不同版本的真空吸尘器世界允许不同的规则，例如智能体可以感知什么，它的动作是否总是成功等

2.1 智能体和环境

■ 示例：真空吸尘器世界

感知序列	动作
$[A, Clean]$	<i>Right</i>
$[A, Dirty]$	<i>Suck</i>
$[B, Clean]$	<i>Left</i>
$[B, Dirty]$	<i>Suck</i>
$[A, Clean], [A, Clean]$	<i>Right</i>
$[A, Clean], [A, Dirty]$	<i>Suck</i>
:	:
$[A, Clean], [A, Clean], [A, Clean]$	<i>Right</i>
$[A, Clean], [A, Clean], [A, Dirty]$	<i>Suck</i>
:	:

图 2-3 图 2-2 所示的真空吸尘器世界的简单智能体函数的部分表项。如果当前方格是脏的，智能体就会进行清理，否则它将移到另一个方格。注意，除非限制可能感知序列的长度，否则该表的大小是无限的

2.1 智能体和环境

■ 示例：真空吸尘器世界

- ▶ 右边一列的方法不同，就可以定义不同的真空吸尘器世界智能体
- ▶ 什么样的智能体是好的，或者是坏的，是智能的还是愚笨的？

感知序列	动作
[A, Clean]	Right
[A, Dirty]	Suck
[B, Clean]	Left
[B, Dirty]	Suck
[A, Clean], [A, Clean]	Right
[A, Clean], [A, Dirty]	Suck
:	:
[A, Clean], [A, Clean], [A, Clean]	Right
[A, Clean], [A, Clean], [A, Dirty]	Suck
:	:

图 2-3 图 2-2 所示的真空吸尘器世界的简单智能体函数的部分表项。如果当前方格是脏的，智能体就会进行清理，否则它将移到另一个方格。注意，除非限制可能感知序列的长度，否则该表的大小是无限的

- 智能体和环境
- 良好行为：理性的概念
- 环境的本质
- 智能体的结构

2.2 良好行为：理性的概念 - 性能度量

■ 理性智能体

□ 做正确事情的智能体。智能体函数表格的每一项都填写正确。

■ 什么是“做正确的事”？

- 结果主义：通过结果来评估智能体的行为
- 当智能体进入环境时，它会根据接受的感知产生一个动作序列。
这一动作序列会导致环境经历一系列的状态。如果序列是可取的，则智能体表现良好。
- 可取性：由性能度量描述，该度量评估任何给定环境状态的序列

2.2 良好行为：理性的概念 - 性能度量

- 对所有任务和智能体，没有一成不变的固定的性能度量，
具体问题具体分析
- **示例：真空吸尘器智能体**
 - 统计8小时工作时间内清理的灰尘总量来度量它的性能
 - 对理性智能体而言，你所要求即你所得
 - ▶ 一边吸尘，一边又把灰尘倒回地面，再吸尘，持续下去，从而使得用作性能度量的灰尘量最大化。
 - 性能度量2：奖励保持干净地面的智能体
 - ▶ 每个时间步，每个清洁的方格奖励一分（同时考虑电力消耗和产生噪音的惩罚）
- 一般原则：根据实际在环境中希望得到的结果来设计性能度量，而不是根据智能体应该如何表现

2.2 良好行为：理性的概念 - 性能度量

- 仍有棘手问题
- 例如：“干净地面”——基于一段时间内的平均整洁度
 - 两个不同的智能体可以达到相同的平均整洁度
 - ▶ 一个智能体工作始终保持一般水平
 - ▶ 另一个智能体短时间工作效率很高但需要长时间的休息
 - 哪个更好？

2.2 良好行为：理性的概念 - 理性

- 在任何时候，理性取决于以下4方面：
 - 定义成功标准的性能度量
 - 智能体对环境的先验知识
 - 智能体可以执行的动作
 - 智能体到目前为止的感知序列
- 理性智能体的定义：
 - 对于每个可能的感知序列，给定感知序列提供的证据和智能体所拥有的任何先验知识，理性智能体应该选择一个期望最大化其性能度量的动作。

2.2 良好行为：理性的概念 - 理性

■ 示例：真空吸尘器智能体

- 如果一个方格是脏的就吸尘，否则就移动到另一个方格，是否理性？
- 在下述条件下，的确是理性的
 - ▶ 在 1000 个时间步的“生命周期”内，性能度量在每个时间步为每个干净的方格奖励 1 分
 - ▶ 环境的“地理信息”作为先验知识是已知的，但灰尘的分布和智能体的初始位置未知，干净的方格保持干净，而吸尘动作会清理当前方格，右移或左移的动作使智能体移动一个方格，前提是这不会使智能体移出该环境，否则智能体会保持原位
 - ▶ 可用的动作仅有右移、左移和吸尘
 - ▶ 智能体能够正确感知其位置以及该位置是否有灰尘

2.2 良好行为：理性的概念 - 理性

■ 示例：真空吸尘器智能体

- 同样的智能体在不同的环境下会变得不理性
 - ▶ 例如：一旦所有的灰尘都被吸干净了，该吸尘器就会毫无必要地跑来跑去
 - ▶ 更好的智能体应该在它确定所有方格都已经干净了之后不做任何事情
 - ▶ 如果方格再次被弄脏了，该智能体应该不定期检查并在必要时重新清洁
 - ▶ 如果环境的地理信息是未知的，智能体则需要对其进行探索

2.2 良好行为：理性的概念 - 全知、学习和自主

■ 区别理性和全知的概念

□ 全知：全知的智能体能预知其行动的实际结果，并能据此采取行动。全知在现实中是不可能的。

□ 例子：

- ▶ 沿马路散步，看到街对面的老朋友
- ▶ 附近没有车辆，我也没有别的事情，所有根据理性，开始穿过马路
- ▶ 在33000英尺的高空一扇货舱门从一架飞过的飞机上掉了下来，在我到达马路对面之前拍扁了我

理性但不完美

2.2 良好行为：理性的概念 - 全知、学习和自主

- 理性不等于完美
 - 理性：期望性能最大化
 - 完美：实际性能最大化（现实中不可能）
- 理性的定义并不要求全知，理性的决策只取决于迄今为止的感知序列
- 理性智能体应该进行充分的信息收集
 - 例如，如果智能体在穿行繁忙的马路前没有观察道路两边的情况，则它就感知不到附近有大卡车在高速接近。理性的智能体不能在此情况下做出可以穿过马路的决定。根据信息不全的感知序列穿行马路是不理性的。

2.2 良好行为：理性的概念 - 全知、学习和自主

- 理性智能体不仅要能够收集信息，还要求能从它所感知的信息中尽可能多的学习。
- 智能体最初的设定可能反映的是环境的先验知识，随着智能体经验的丰富，这些知识会被修改或者增强。

2.2 良好行为：理性的概念 - 全知、学习和自主

■ 示例：蜣螂的例子

- 做窝并产卵后，会从附近的粪堆取回一个粪球堵住窝的入口。
- 如果粪球在路途中脱离了它的掌握，还会继续赶路，并作动作用不存在的粪球塞住入口，而不会注意到粪球已经不见了。
- 蜣螂在进化时在它的行为里内建了假设，当该假设被破坏时，就会产生不成功的行为。

2.2 良好行为：理性的概念 - 全知、学习和自主

■ 示例：黑足泥蜂的例子

- 雌蜂先挖一个洞，出去叮一只毛虫并拖回洞，再次进洞查看，再把毛虫拖到洞里，然后产卵。毛虫在黑足泥蜂孵卵期间作为食物来源。
 - 假如把毛虫挪开几英寸，雌蜂就会回到计划中“拖毛虫到地洞”的步骤，继续进行不做任何修改的计划。
-
- 如果智能体依赖于其设计者的先验知识，而不是它自身的感知和学习过程，这种情况我们会说该智能体缺乏自主性。

2.2 良好行为：理性的概念 - 全知、学习和自主

■ 理性智能体应该是自主的

- 它应该学习，以弥补不完整的或者不正确的先验知识
- 例如：能学习预测何时何地会出现额外灰尘的真空吸尘器比不能学习预测的要好

■ 实际中，很少要求智能体从一开始就完全自主

- 当智能体没有或者只有很少的经验时，行为往往随机
- 给人工智能智能体提供一些初始知识以及学习能力是合理的
- 当得到环境的充足经验后，理性智能体的行为才能独立于先验知识有效行动

- 智能体和环境
- 良好行为：理性的概念
- 环境的本质
- 智能体的结构

2.3 环境的本质

- 有了理性的定义，可以构建理性智能体，还需要考虑任务环境，它本质上是“问题”，理性智能体是“解决方案”。
- 2.3.1 指定任务环境
- 2.3.2 任务环境的属性

2.3 环境的本质 - 指定任务环境

■ 任务环境 (PEAS描述)

- Performance (性能)
- Environment (环境)
- Actuators (执行器)
- Sensors (传感器)

■ 设计智能体第一步就是尽可能完整地指定任务环境。

2.3 环境的本质 - 指定任务环境

■ 示例：自动驾驶出租车

- 全自动驾驶任务是完全开放的。环境组合会不断地产生新的状况，是无限的。

智能体类型	性能度量	环境	执行器	传感器
自动驾驶出租车司机	安全、速度快、合法、舒适旅程、最大化利润、对其他道路用户的影响最小化	道路、其他交通工具、警察、行人、客户、天气	转向器、加速器、制动、信号、喇叭、显示、语音	摄像头、雷达、速度表、GPS、发动机传感器、加速度表、麦克风、触摸屏

图 2-4 自动驾驶出租车司机任务环境的 PEAS 描述

2.3 环境的本质 - 指定任务环境

■ 示例：自动驾驶出租车

智能体类型	性能度量	环境	执行器	传感器
自动驾驶出租车司机	安全、速度快、合法、舒适旅程、最大化利润、对其他道路用户的影响最小化	道路、其他交通工具、警察、行人、客户、天气	转向器、加速器、制动、信号、喇叭、显示、语音	摄像头、雷达、速度表、GPS、发动机传感器、加速度表、麦克风、触摸屏

图 2-4 自动驾驶出租车司机任务环境的 PEAS 描述

□ 性能度量：

- ▶ 到达正确的目的地，尽量减少油耗和磨损，尽量减少行程时间或成本，尽量减少违反交通法规和对其他驾驶员的干扰，最大限度地提高安全性和乘客舒适度，最大化利润

□ 环境：

- ▶ 各种各样的道路；路上有其他车辆、行人、动物、道路施工、警车、石头和坑洞等；天气，驾驶规则等。环境越受限，设计问题就越简单。

2.3 环境的本质 - 指定任务环境

■ 示例：自动驾驶出租车

智能体类型	性能度量	环境	执行器	传感器
自动驾驶出租车司机	安全、速度快、合法、舒适旅程、最大化利润、对其他道路用户的影响最小化	道路、其他交通工具、警察、行人、客户、天气	转向器、加速器、制动、信号、喇叭、显示、语音	摄像头、雷达、速度表、GPS、发动机传感器、加速度表、麦克风、触摸屏

图 2-4 自动驾驶出租车司机任务环境的 PEAS 描述

□ 执行器：

- ▶ 通过加速器控制发动机以及控制转向和制动。输出到显示屏或语音合成器来与乘客对话，也许还需要其他途径同其他车辆进行交流。

□ 传感器：

- ▶ 一个或者多个可控制的视频摄像头，红外或声纳用来检测与其他车辆或障碍的距离，速度表，加速计，引擎、燃油与电子系统的传感器阵列，全球卫星GPS，最后需要一个触摸屏或麦克风供乘客说明目的地。

2.3 环境的本质 - 指定任务环境

■ 更多的不同类型智能体的PEAS描述

智能体类型	性能度量	环境	执行器	传感器
医学诊断系统	治愈患者、降低费用	患者、医院、工作人员	用于问题、测试、诊断、治疗的显示器	用于症状和检验结果的触摸屏/语音输入
卫星图像分析系统	正确分类对象和地形	轨道卫星、下行链路、天气	场景分类显示器	高分辨率数字照相机
零件选取机器人	零件在正确箱中的比例	零件输送带、箱子	有关节的手臂和手	摄像头、触觉和关节角度传感器
提炼厂控制器	纯度、产量、安全	提炼厂、原料、操作员	阀门、泵、加热器、搅拌器、显示器	温度传感器、气压传感器、流量传感器、化学传感器
交互英语教师	学生的考试分数	一组学生、考试机构	用于练习、反馈、发言的显示器	键盘输入、语音

图 2-5 智能体类型及其 PEAS 描述的示例

2.3 环境的本质 - 指定任务环境

- 区分“真实”环境和“人工”环境并不重要，重要的是智能体行为、环境产生的感知序列和性能度量之间关系的复杂性。“真实”环境有可能简单，“人工”环境也有可能复杂。
- 设计用来检测传送带上零件的机器人，可以做简单假设：
 - 照明不变，零件种类固定，只有两种可能行动（接受或者拒绝）
- 软件机器人：扫描互联网上的新闻来源并把有趣的条目发送给用户，同时也生成广告空间来换取收入
 - 自然语言处理能力，动态改变规划

2.3 环境的本质 - 任务环境的属性

- 人工智能中可能出现的任务环境范围非常广泛，定义少量维度对任务环境进行分类，决定适当的智能体设计及实现智能体的主要技术系列。
- 分类维度：
 - 完全可观测的与部分可观测的
 - 单智能体的和多智能体的
 - 确定性的与非确定性的
 - 回合式的与序贯的
 - 静态的与动态的
 - 离散的与连续的
 - 已知的与未知的

2.3 环境的本质 - 任务环境的属性

■ 完全可观测的与部分可观测的

- 完全可观测的：
 - ▶ 智能体的传感器能让它在每个时间点都能访问环境的完整状态。
- 有效的完全可观测的：
 - ▶ 传感器检测到与动作选择相关的所有方面，相关取决于性能度量标准。
- 噪声、不精确的传感器、或者传感器数据中缺少部分状态，都可能导致环境成为部分可观测的。
- 若智能体根本没有传感器，环境则是不可观测的。

2.3 环境的本质 - 任务环境的属性

■ 单智能体的与多智能体的

□ 单智能体的：

- ▶ 独自解决纵横字谜的智能体处于单智能体环境中

□ 多智能体的：

- ▶ 下国际象棋的智能体处于二智能体环境中

多智能体环境中不同智能体是否寻求各自性能度量最大化

- ▶ 竞争性

- 国际象棋中的对手实体B正试图最大化其性能度量，根据国际象棋规则，这将最小化智能体A的性能度量

- ▶ 合作的

- 在出租车驾驶环境中，避免碰撞使所有智能体的性能度量最大化

2.3 环境的本质 - 任务环境的属性

■ 确定性的与非确定性的

□ 确定性的：

- ▶ 如果环境的下一个状态完全由当前状态和智能体执行的动作决定

□ 非确定性的：

- ▶ 环境是部分可观测的

- 出租车驾驶环境，无人能精确预测交通状况，车辆爆胎或者引擎失灵

□ 区分随机的与非确定性的：

- ▶ 随机的：环境模型显式地处理概率

- 明天的降雨可能性为25%

- ▶ 非确定性的：可能性没有被量化

- 明天有可能下雨

2.3 环境的本质 - 任务环境的属性

■ 回合式的与序贯的

□ 回合式的：

- ▶ 智能体的经验被划分为原子式的回合
- ▶ 在每一回合中，智能体接收一个感知，然后执行单个动作。下一回合并不依赖于前几回合采取的动作。
 - 在装配流水线上检测缺陷零件的智能体

□ 序贯的：

- ▶ 当前决策可能会影响未来所有决策，短期行为可能会产生长期影响
 - 国际象棋、出租车驾驶

2.3 环境的本质 - 任务环境的属性

■ 静态的与动态的

□ 动态的：

- ▶ 环境在智能体思考时发生了变化，动态环境会不断地询问智能体想要采取什么行动
 - 驾驶出租车

□ 静态的：

- ▶ 相对容易处理，因为智能体在决定某个操作时不需要一直关注世界，也不需要担心时间的流逝
 - 填字游戏

□ 半动态的：

- ▶ 环境本身不会随着时间的推移而改变，但智能体的性能分数会改变
 - 在用时钟计时的情况下国际象棋

2.3 环境的本质 - 任务环境的属性

■ 离散的与连续的

- 环境的状态、处理时间多方式以及智能体的感知和行动，都有离散/连续之分
 - ▶ 国际象棋环境具有有限数量的不同状态（不包括时钟），也有一组离散的感知和动作
 - ▶ 驾驶出租车是一个连续状态和连续时间的问题，出租车和其他车辆的速度和位置是一系列连续的值，并随着时间平稳地变化，出租车的驾驶动作也是连续的（转向角等）

2.3 环境的本质 - 任务环境的属性

■ 已知的与未知的

- 不是指环境本身，而是指智能体（或设计者）对环境“物理定律”的认知状态
 - ▶ 已知环境中，所有行动的结果（如果环境是非确定性的，则对应结果的概率）都是既定的
 - ▶ 未知环境中，智能体将不得不了解它是如何工作的，才能做出正确的决策
 - ▶ 已知环境可以是部分可观测的
 - 在纸牌游戏中，知道规则但仍然无法看到尚未翻转的牌
 - ▶ 未知环境可以是完全可观测的
 - 一个全新的电子游戏，屏幕可能会显示整个游戏状态，但在尝试之前并不知道各个按钮的作用

- 智能体和环境
- 良好行为：理性的概念
- 环境的本质
- 智能体的结构

2.4 智能体的结构

- 人工智能的工作是设计一个智能体程序实现智能体函数，即从感知到动作的映射
- 假设该程序将运行在某种具有物理传感器和执行器的计算设备上，我们称之为智能体架构

智能体=架构+程序

2.4 智能体的结构

■ 选择的程序必须适合相应架构

- 如果程序打算推荐步行这样的动作， 那么对应的架构最好有腿
- 架构可能只是一台普通PC， 也可能是一辆带有多台车载计算机、摄像头和其他传感器的机器人汽车
- 通常， 架构使程序可以使用来自传感器的感知， 然后运行程序，并将程序生成的动作选择反馈给执行器。

2.4 智能体的结构 - 智能体程序

- 本书中讨论的智能体程序具有同样的框架：
 - 将**当前**感知作为传感器的输入，并将动作返回给执行器
- 智能体程序以当前感知作为输入，智能体函数以整个感知历史作为输入

```
function TABLE-DRIVEN-AGENT(percept) returns 一个动作
  persistent: percepts, 初始为空的序列
            table, 以感知序列为索引的动作表, 初始为完全确定
```

将*percept*添加到*percepts*的末尾

action \leftarrow LOOKUP(*percepts*, *table*)

return *action*

图 2-7 每个新感知都会调用 TABLE-DRIVEN-AGENT 程序，并且每次返回一个动作。它在内存中保留了完整的感知序列

2.4 智能体的结构 - 智能体程序

- 表驱动方法：以跟踪感知序列为索引，到行动表里查询以做出决策
 - 自动驾驶出租车：1小时驾驶对应的查找表条目超过 $10^{600000000000}$ 个
 - 国际象棋： 10^{150}
- 不可行：
 - 宇宙中没有任何物理智能体有空间存储表
 - 设计者没有时间创建表
 - 任何智能体都无法从其经验中学习所有正确的记录
 - 即使环境足够简单，表的容量可以接受，设计者仍然没有向导来填写表中条目

2.4 智能体的结构 - 智能体程序

- 如果不考虑上述因素，Table-Driven-Agent确实做到了我们所需要的：
 - 实现了所需的智能体函数
- 人工智能面临的关键挑战是找出编写程序的方法，尽可能从一个小程序而不是从一个大表中产生理性行为。

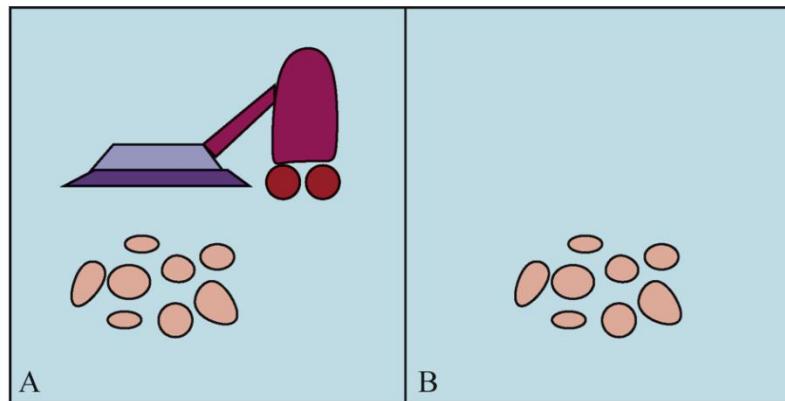
2.4 智能体的结构 - 智能体程序

■ 4种基本的智能体程序：

- 简单反射型智能体
- 基于模型的反射型智能体
- 基于目标的智能体
- 基于效用的智能体

2.4 智能体的结构 - 简单反射型智能体

- 智能体根据当前感知选择动作，忽略感知历史的其余部分



function REFLEX-VACUUM-AGENET([*location,status*]) **returns** 一个动作

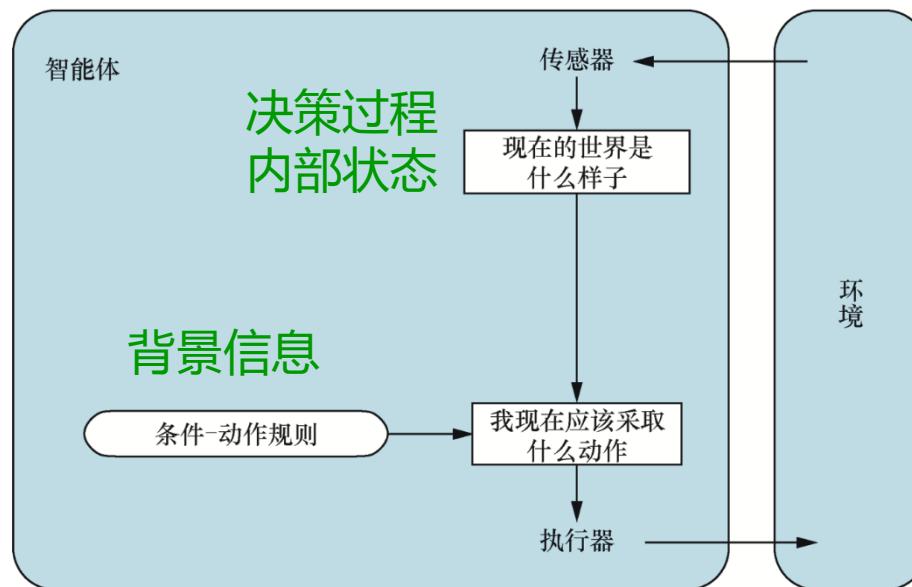
```

if status = Dirty then return Suck
else if location = A then return Right
else if location = B then return Left
    
```

图 2-8 在只有两个位置的真空吸尘器环境中，简单反射型智能体的智能体程序，该程序实现图 2-3 中列出的智能体函数

2.4 智能体的结构 - 简单反射型智能体

- 简单反射行为也会发生在更加复杂的环境中
 - 如果前面的车正在刹车，则启动刹车（这种联结称为**条件-动作规则**）



简单反射型智能体具有极好的简单特性，但是智能有限，要求环境完全可观测

```

function SIMPLE-REFLEX-AGENT(percept) returns 一个动作
  persistent: rules, 一组条件-动作规则

  state  $\leftarrow$  INTERPRET-INPUT(percept)
  rule  $\leftarrow$  RULE-MATCH(state, rules)
  action  $\leftarrow$  rule.ACTION
  return action
  
```

图 2-10 简单反射型智能体。它根据一条规则进行操作，该规则的条件与感知定义的当前状态相匹配

2.4 智能体的结构 - 基于模型的反射型智能体

- 处理部分可观测性的最有效方法——让智能体根据某种依赖于感知历史的内部状态，追踪它现在观测不到的部分世界。
- 更新这些内部状态信息需要在智能体程序中以某种形式编码两种知识
 - 世界如何随时间变化——世界的转移模型
 - 智能体行为的影响和世界如何独立于智能体而发展
 - 当智能体顺时针转动方向盘时，汽车就会向右转
 - 下雨时，汽车的摄像头就会被淋湿
 - 世界状态如何反映在智能体感知中——传感器模型
 - 当前面的汽车开始刹车时，前向摄像头的图像中会出现亮起的红色区域
 - 当摄像头被淋湿时，图像中会出现水滴状物体并部分遮挡道路
- 转移模型和传感器模型结合在一起让智能体能够在传感器受限时尽可能地跟踪世界的状态，使用此类模型的智能体称为**基于模型的智能体**

2.4 智能体的结构 - 基于模型的反射型智能体

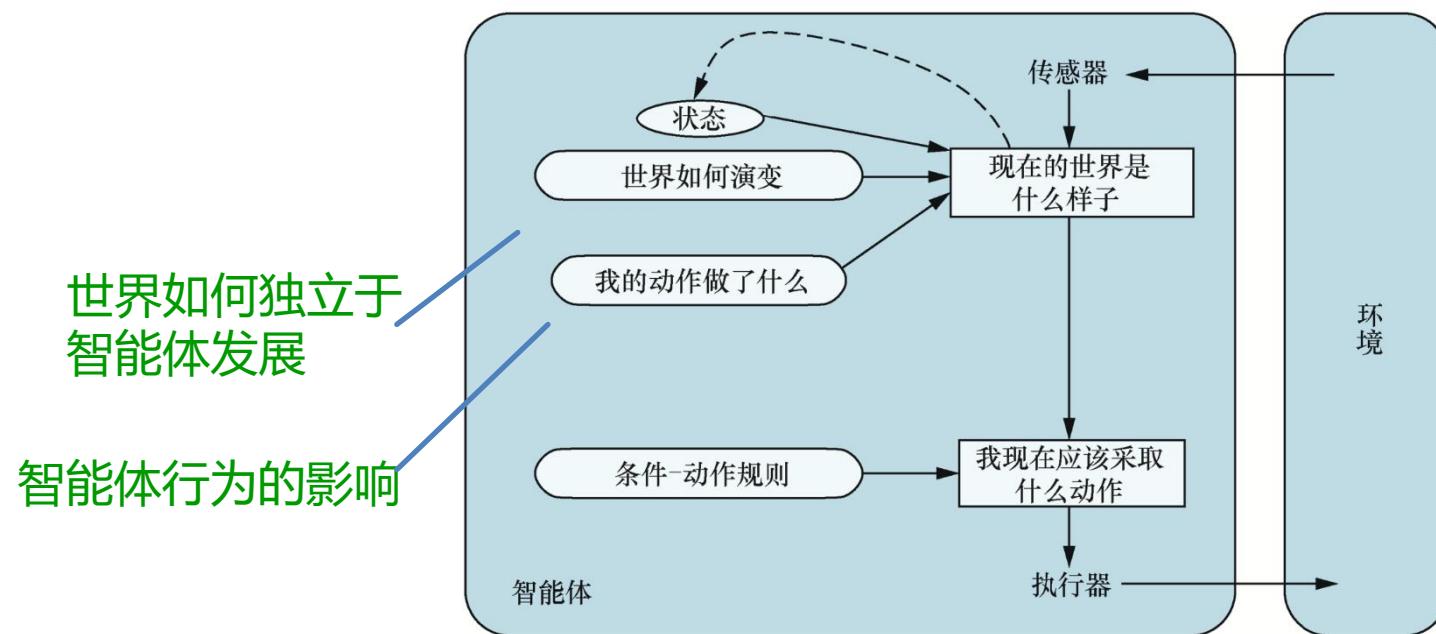


图 2-11 基于模型的智能体

function MODEL-BASED-REFLEX-AGENT(*percept*) **returns** 一个动作

persistent: *state*, 智能体对世界状态的当前理解

transition_model, 关于下一个状态如何决定于当前状态和动作的描述

sensor_model, 关于当前世界状态如何反映到智能体感知的描述

rules, 一组条件-动作规则

action, 最近的动作, 初始为空

—— 负责创建描述新的内部状态

state \leftarrow UPDATE-STATE(*state*, *action*, *percept*, *transition_model*, *sensor_model*)

rule \leftarrow RULE-MATCH(*state*, *rules*)

action \leftarrow *rule.ACTION*

return *action*

图 2-12 基于模型的反射型智能体。它使用内部模型追踪世界的当前状态, 然后以与反射型智能体相同的方式选择动作

2.4 智能体的结构 - 基于目标的智能体

- 了解环境的现状并不总是足以决定做什么
 - 例如，在一个路口，出租车可以左转、右转或直行。正确的决定取决于出租车要去哪里。
- 除了当前状态的描述之外，智能体还需要某种描述理想情况的**目标**信息
- 基于目标的智能体结构：
 - 将目标信息与模型相结合（与基于模型的反射型智能体中使用的信息相同），并选择实现目标的动作
- 搜索，规划

2.4 智能体的结构 - 基于目标的智能体

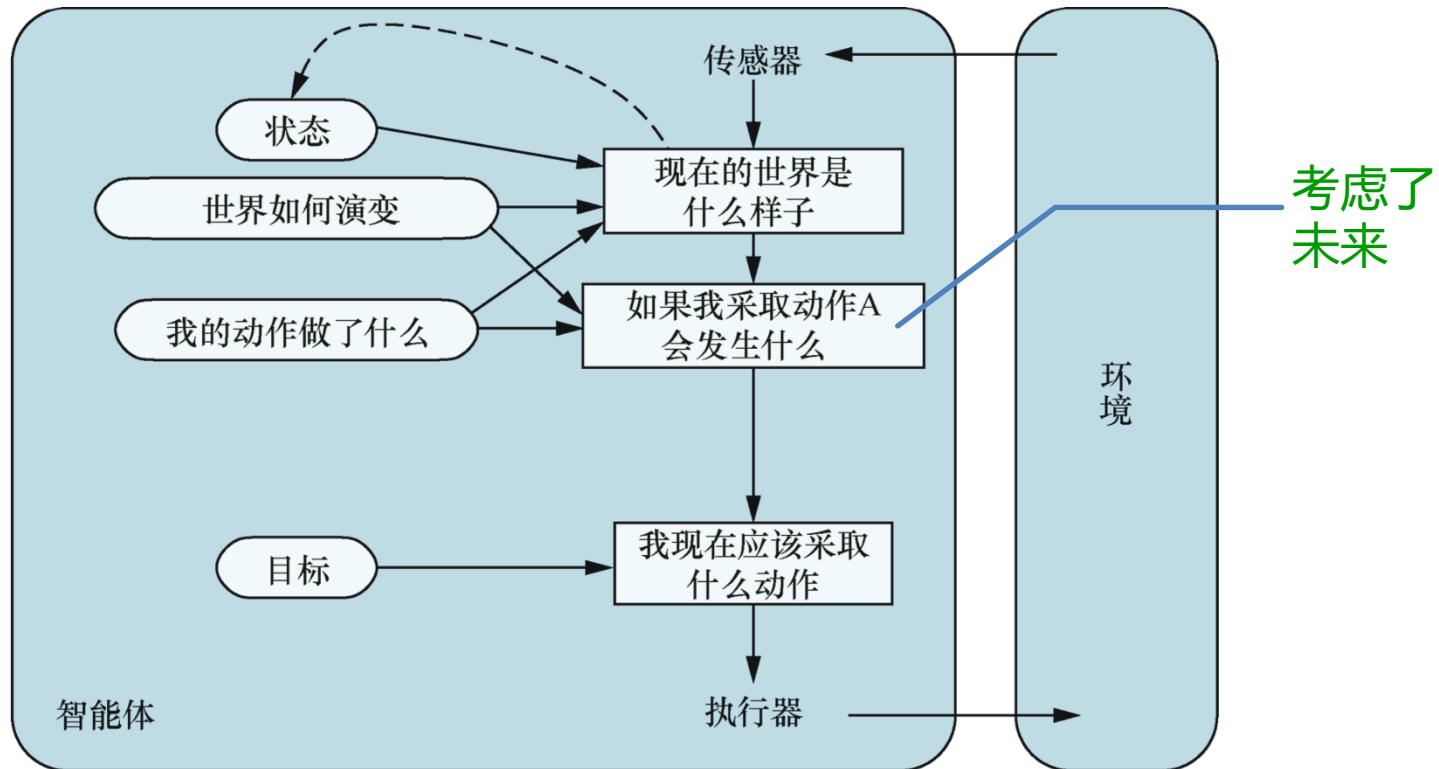


图 2-13 基于模型、基于目标的智能体。它追踪世界状态以及它试图实现的一系列目标，并选择一项最终能够实现目标的动作

2.4 智能体的结构 - 基于目标的智能体

- 这类决策从根本上不同于前面描述的条件-动作规则，因为它涉及对未来的考虑
 - 如果我这样做会发生什么？
 - 这会让我快乐吗？
- 反射型智能体
 - 在看到刹车灯时刹车，但它不知道为什么
- 基于目标的智能体
 - 在看到刹车灯时会刹车，因为这是它预测的唯一动作，这个动作可以实现不撞到其他汽车的目标
 - 相比于反射型智能体，效率较低，但更灵活，因为支持其决策的知识是显式表示的，并且可以修改

2.4 智能体的结构 - 基于效用的智能体

- 仅靠目标并不足以产生高质量的行为，只提供了一个粗略的二元区别，更快、更安全、更可靠或更便宜
- 性能度量会给任何给定的环境状态序列打分，因此它可以很容易地区分到达出租车目的地所采取的更可取和更不可取的方式
- 智能体的**效用函数**本质上是性能度量的内部化。如果内部效用函数和外部性能度量一致，那么根据外部性能度量选择动作，以使其效用最大化的智能体是理性的。
 - 当存在相互冲突的目标时，效用函数会进行适当的权衡
 - 当智能体有多个目标实现，但没有一个目标可以确定地实现时，效用提供了一种方法，可以权衡目标的重要性和成功的可能性

2.4 智能体的结构 - 基于效用的智能体

- 基于效用的理性智能体会选择能够最大化其动作结果**期望效用**的动作

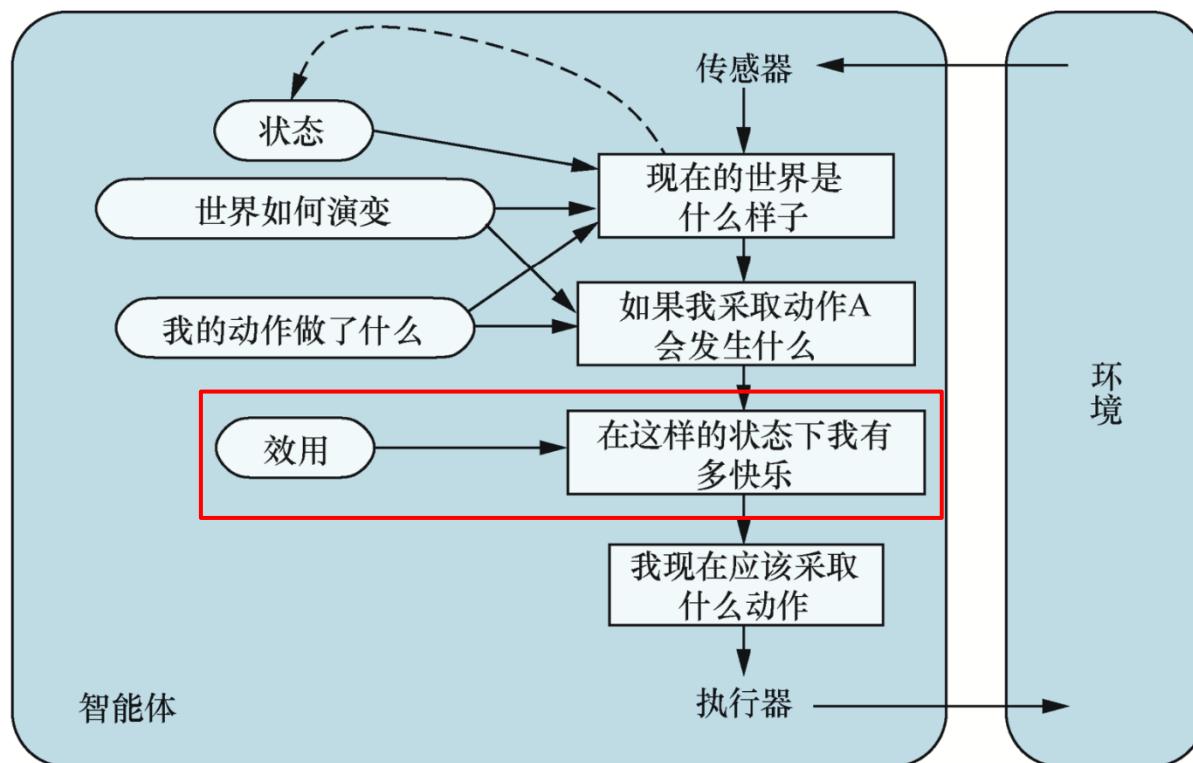


图 2-14 基于模型、基于效用的智能体。它使用了一个世界模型以及一个效用函数来衡量它在各状态之间的偏好，然后选择产生最佳期望效用的动作，其中期望效用是通过对所有可能的结果状态和对应概率加权所得

2.4 智能体的结构 - 学习型智能体

- 学习让智能体能够在最初未知的环境中运作，并变得比其最初的知识可能允许的能力更强

评估者：

反馈评价，并确定如何修改性能元素以在未来做得更好

学习元素：
负责提升

问题生成器：
可以得到全新和有信息丰富的经验的动作建议

性能元素：
负责选择外部行动

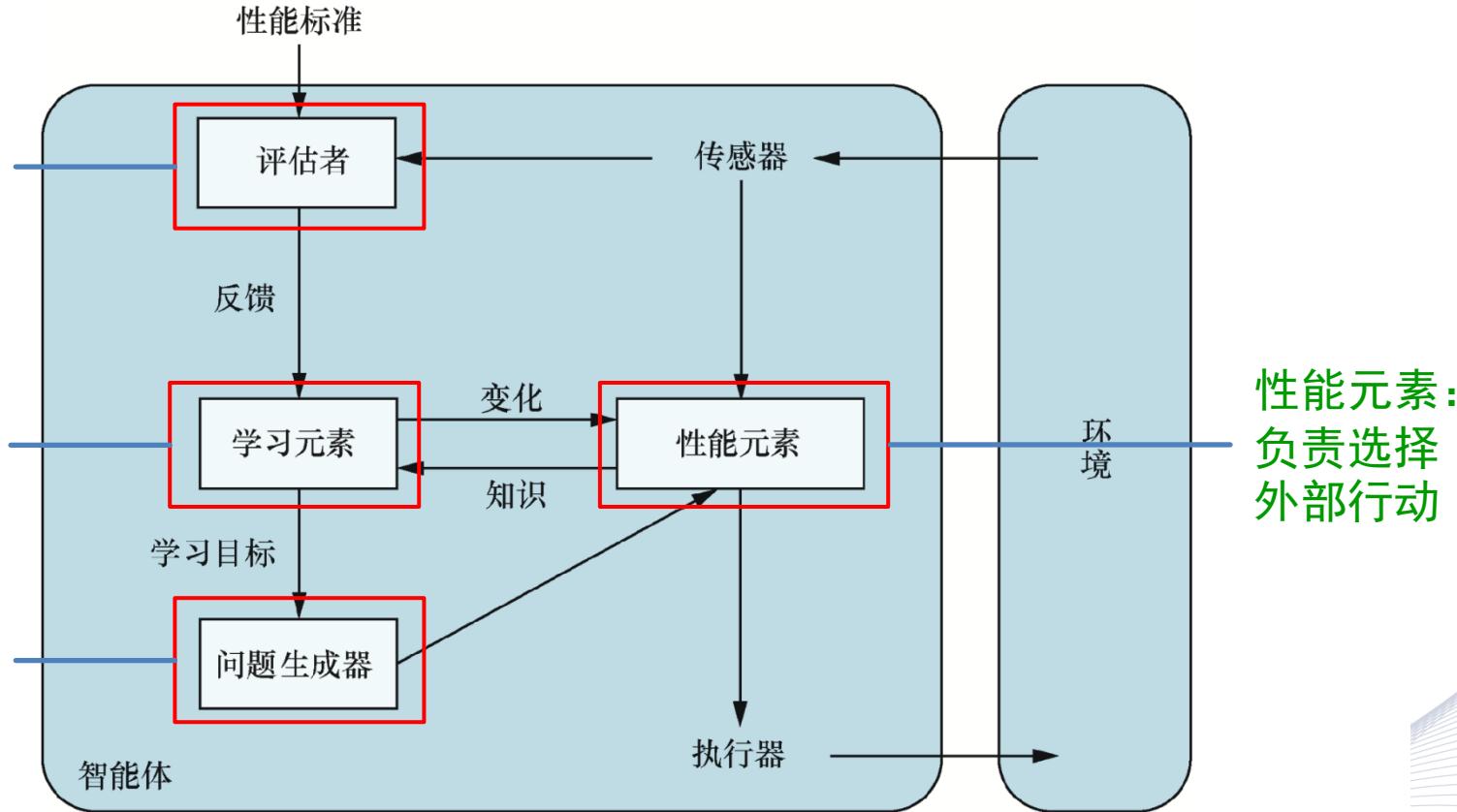


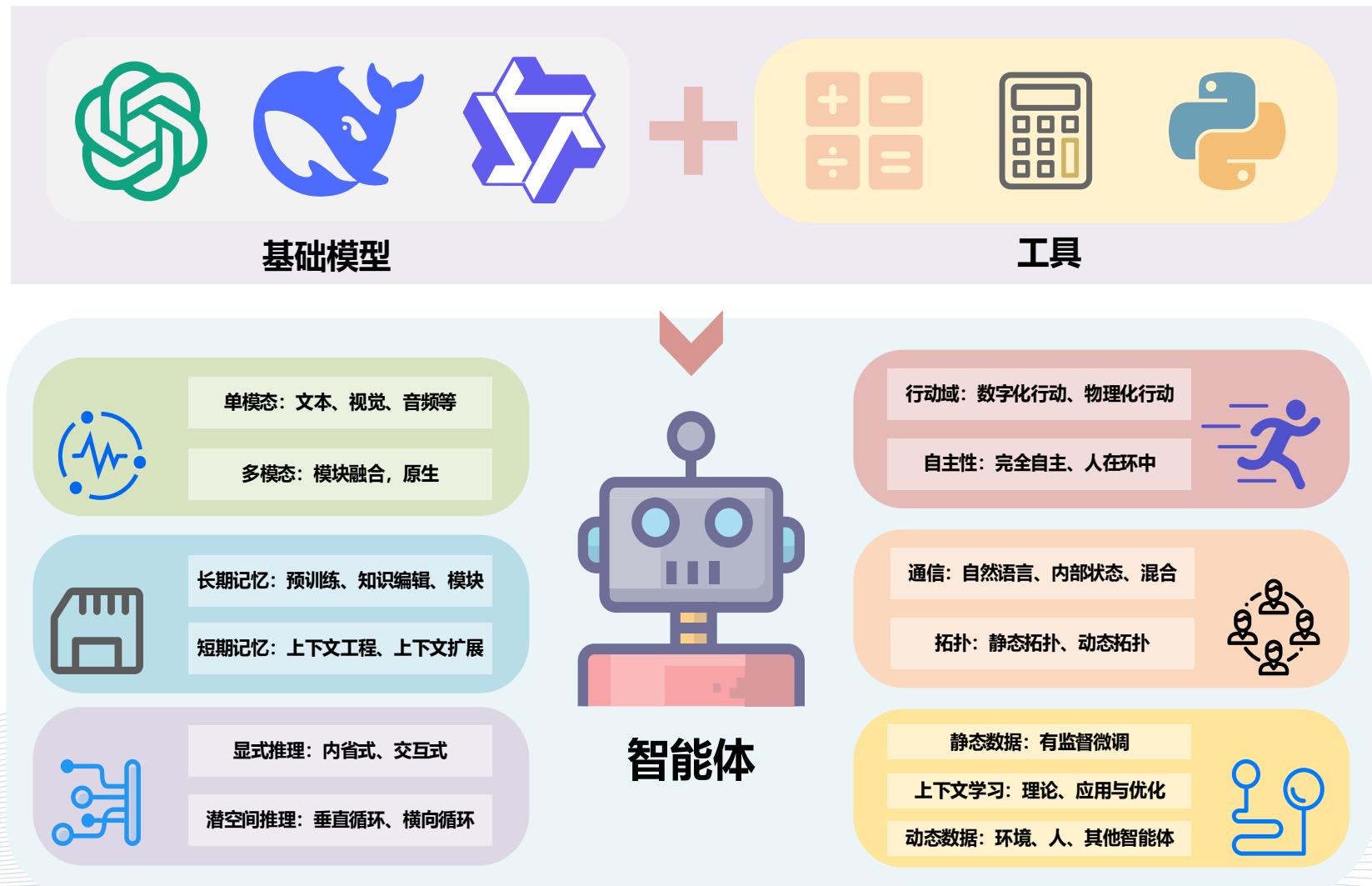
图 2-15 通用学习型智能体。“性能元素”框表示我们之前认为的整个智能体程序，现在“学习元素”框可以修改该程序以提升其性能

2.4 智能体的结构 - 学习型智能体

■ 示例：智能出租车

- 性能元素包括用来选择驾驶行为的全部知识和过程集合，出租车使用性能元素在公路上行驶
- 评估者观察世界并把信息传递给学习元素（比如，横穿三条车道，其他司机惊呼）
- 学习元素制定出一条规则表示行动优劣，安装该规则以修改性能元素
- 问题生成器为了改进的需要而确定需要修改一定范围内的行为，并提议进行试验（不同路面不同条件试验刹车）

由LLM驱动的智能体



2.4 智能体的结构 - 智能体程序的组件如何工作

- 通过一个复杂性和表达能力不断增加的横轴来描述表示，即原子表示、因子化表示和结构化表示

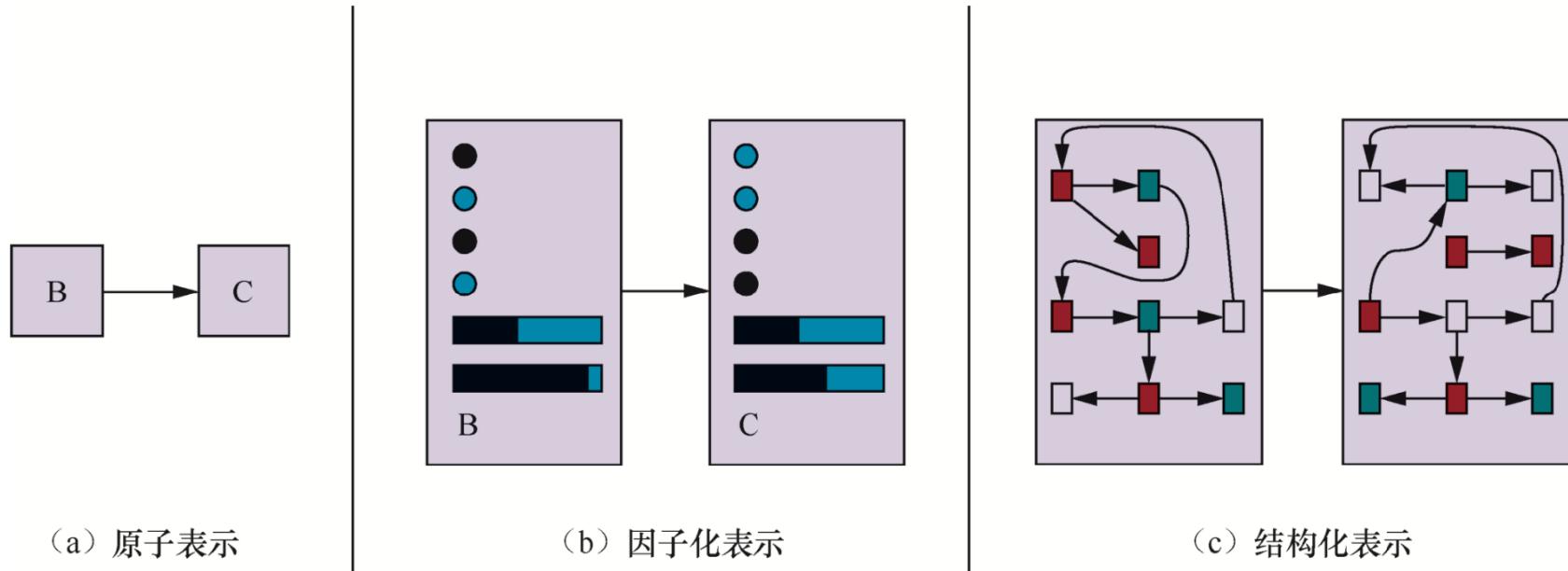


图 2-16 表示状态及其之间转移的 3 种方法：(a) 原子表示一个状态（如 B 或 C）是没有内部结构的黑盒；(b) 因子化表示状态由属性值向量组成，值可以是布尔值、实值或一组固定符号中的一个；(c) 结构化表示状态包括对象，每个对象可能有自己的属性以及与其他对象的关系

2.4 智能体的结构 - 智能体程序的组件如何工作

- 原子表示：世界的每一个状态都是不可分割的，它没有内部结构。
 - 搜索和博弈中的标准算法，隐马尔科夫模型，马尔可夫决策过程
- 因子化表示：将每个状态拆分为一组固定的变量或属性，每个变量或属性都可以有一个值。
 - 约束满足算法，命题逻辑，规划，贝叶斯网络，机器学习算法
- 结构化表示：明确地描述对象及其各种不同的关系。
 - 关系数据库，一阶逻辑，一阶概率模型，基于知识的学习，自然语言理解

2.4 智能体的结构 - 智能体程序的组件如何工作

- 原子表示、因子化表示和结构化表示所在的轴是表达性增强的轴
- 可以通过简洁的描述捕捉到更具表达性的表示，表达性差的表示也可以捕捉到一切，但需要更多描述。通常，表达性更强的语言更简洁
 - 例如：描述国际象棋规则
 - 一两页结构化表示语言（如一阶逻辑）
 - 数千页因子化表示语言（如命题逻辑）
 - 10^{38} 页的原子语言（如有限状态自动机）
- 随着表示能力的增强，推理和学习变得更加复杂
- 为了避免缺点的同时获得表达性表示的好处，真实世界中的智能系统可能需要轴上的所有点同时运行

本章小结

- 智能体是在环境中感知和行动的事物。智能体的智能体函数指定智能体在响应任意感知序列时所采取的动作。
- 性能度量评估智能体在环境中的行为。给定到目前为止所看到的感知序列，理性智能体的动作是为了最大化性能度量的期望值。
- 任务环境规范包括性能度量、外部环境、执行器和传感器。在设计智能体时，第一步必须始终是尽可能完整地指定任务环境。

本章小结

■ 任务环境在几个重要维度上有所不同

- 完全可观测的或部分可观测的
- 单智能体的或多智能体的
- 确定性的或非确定性的
- 回合式的或序贯的
- 静态的或动态的
- 离散的或连续的
- 已知的或未知的

本章小结

- 在性能度量未知或难以正确指定的情况下，智能体优化错误目标的风险很大。在这种情况下，智能体设计应该反映真实目标的不确定性。
- 智能体程序实现智能体函数。存在各种基本的智能体编程，反映了决策过程中明确使用的信息类型。这些设计在效率、紧凑性和灵活性方面各不相同。智能体程序的适当设计取决于环境的性质。
- 简单反射型智能体直接响应感知，而基于模型的反射型智能体保持内部状态以跟踪当前感知中不明晰的世界状态。基于目标的智能体采取行动来实现目标，而基于效用的智能体试图最大化自己期望的“快乐”。
- 所有智能体都可以通过学习提升性能。

课后作业

■ 对于下列活动（任选3个），分别给出任务环境的PEAS描述，并按2.3.2节列出的性质进行分析：

- 足球运动
- 探索Titan的地下海洋
- 在互联网上购买AI旧书
- 打一场网球比赛
- 对着墙壁练网球
- 完成一次跳高
- 织一件毛衣
- 在一次拍卖中对一个物品投标

课后作业

■ 本道习题讨论的是Agent函数和Agent程序的区别：

- 是否有不止一个Agent程序可以实现给定的Agent函数？请举例说明，或者说明为什么不可能
- 有没有无法用任何Agent程序实现的Agent函数？
- 给定一个机器体系结构，能使每个Agent程序刚好实现一个Agent函数吗？
- 给定存储量为n比特的体系结构，可以有多少种可能的不同Agent程序？
- 假设我们让Agent程序固定不变，但机器速度提高，这会改变Agent函数吗？

THANKS