

Artificial Intelligence

2. 智能Agent

前文回顾

- 在第一章中，我们通过对于智能行为的若干定义来尝试刻画人工智能的目标。

像人一样思考	合理地思考
<p>“使计算机思考的令人激动的新成就，……按完整的字面意思就是：有头脑的机器”（Haugeland，1985）</p> <p>“与人类思维相关的活动，诸如决策、问题求解、学习等活动[的自动化]”（Bellman，1978）</p>	<p>“通过使用计算模型来研究智力”（Charniak 和 McDermott，1985）</p> <p>“使感知、推理和行动成为可能的计算的研究”（Winston，1992）</p>
像人一样行动	合理地行动
<p>“创造能执行一些功能的机器的技艺，当由人来执行这些功能时需要智能”（Kurzweil，1990）</p> <p>“研究如何使计算机能做那些目前人比计算机更擅长的事情”（Rich 和 Knight，1991）</p>	<p>“计算智能研究智能 Agent 的设计。”（Poole 等人，1998）</p> <p>“AI……关心人工制品中的智能行为。”（Nilsson，1998）</p>

引入

- 人工智能科学的目标是实现智能Agent
- 那么什么是智能Agent?
- 智能Agent的目标是什么?
- 智能Agent的困难是什么?
- 如何判断一个 ★★★ 是智能Agent?

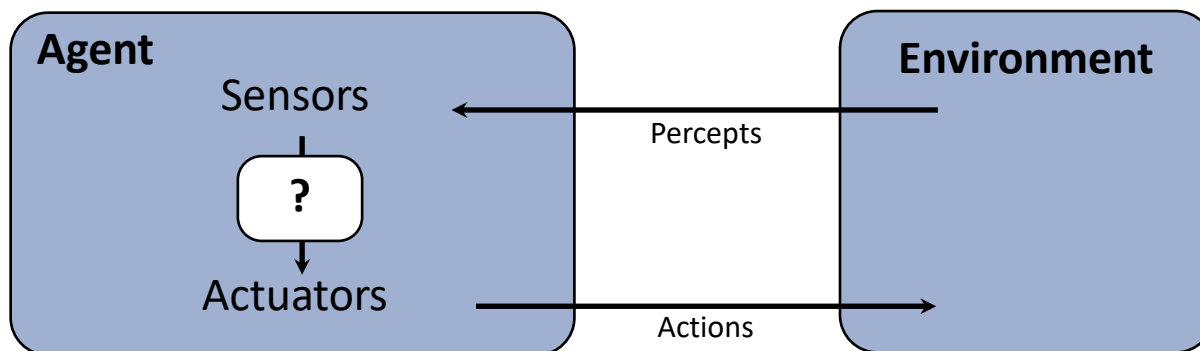
内容

- 2.1 Agent和环境
- 2.2 好的行为：理性的概念
- 2.3 环境的性质
- 2.4 Agent的结构

2.1 Agent和环境

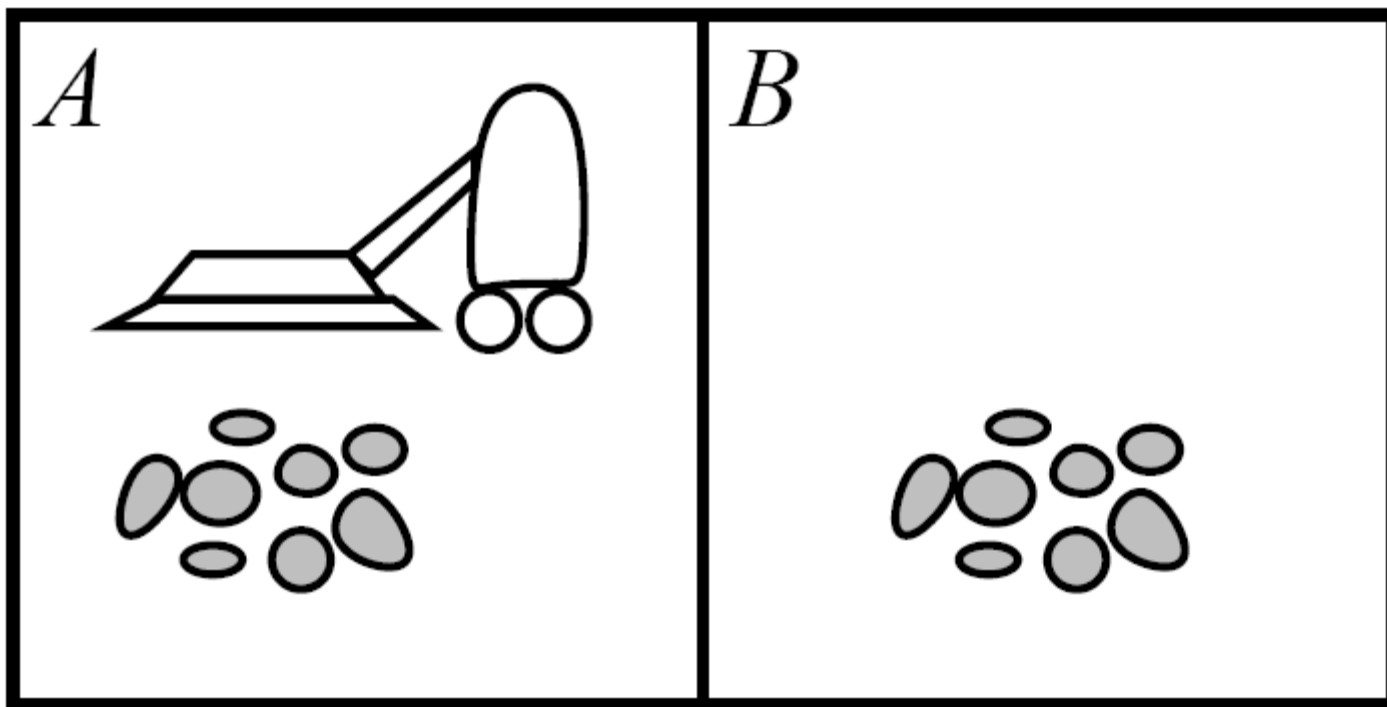
- 理性Agent的概念是人工智能方法的核心

分类	传感器	执行器
人类Agent	眼睛、耳朵、皮肤	手脚、声道
机器人Agent	摄像头、红外测距仪	马达
软件Agent	键盘鼠标、文件、web	屏幕、打印机、文件

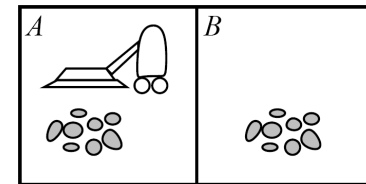


真空吸尘器的世界

- 传感器 (Percept) : 位置、内容
- 执行器 (Action) : 左、右、吸尘、等待



吸尘器agent函数



Percept sequence	Action
$[A, Clean]$	<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>
\vdots	\vdots

function REFLEX-VACUUM-AGENT($[location, status]$) **returns** an action

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

- 什么是**正确的**行为?

2.2 好的行为：理性的概念

理性 Agent 是做事正确的 Agent——从概念上讲, Agent 函数表格的每一项都填写正确。显然做正确的事要比做错事好, 但是到底什么是“做正确的事”呢?

我们用老方法来回答这个老问题: 考虑 Agent 行动的后果。当把 Agent 置于一个环境中后, 它针对收到的感知信息生成一个行动序列。这个行动序列导致环境经历一系列的状态变化。如果该系列正是渴望的, 那么这个 Agent 性能良好。这里的渴望, 通过**性能度量**表述, 它对环境状态的任何给定序列进行评估。

注意我们这里说的是环境状态, 而不是 Agent 状态。如果我们从 Agent 的角度定义其性能是否成功达到, Agent 可以达到完美理性, 它只需欺骗自己说环境是完美的。尤其人类 Agent 是典型的“酸葡萄”——如果是得不到的, 他们会相信他们并不在意 (如诺贝尔奖)。

然而, 数学模型最擅长的即是限定条件下求解

2.2 好的行为：理性的概念

显然，对所有的任务和 Agent，没有一成不变的固定的性能度量；典型地，设计人员会具体问题具体分析。这件事并不像说得这么容易。考虑上节提到的真空吸尘器 Agent。我们可以通过统计 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 Agent 而言，你所要求的即你所得。那么一个理性 Agent 可能一边吸尘，一边又把灰尘倒回地面，再吸尘，持续下去，从而使用作性能度量的灰尘量最大化。更合适的性能度量则是奖励保持干净地面的 Agent。例如，在每个时间步，每个清洁的方格奖励一分（也许要加上对电力消耗和产生噪音的惩罚）。作为一般原则，最好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 Agent 表现出的行为。

即使避免了明显的陷阱，依然有些棘手问题难以处理。例如，前一段提到过“干净地面”，是指随时间变化的平均清洁度。然而两个不同的 Agent 可能得到同样的平均清洁度，一个一直在做普通的清洁工作，而另一个短时间内积极清洁然后休息很长时间。哪种工作方式更可取似乎是清洁科学的好课题，但实际上它是一个有着更深含义的哲学问题。哪个更好——起伏不定的不计后果的生活，还是安全但单调的存在？哪个更好——人人都生活在适中的贫困经济，还是有人生活富足而其他人非常贫穷的经济？

AI 没有对于好的定义，人类才有。

2.2.1 理性

任何指定的时刻，什么是理性的判断依赖于以下 4 个方面：

- 定义成功标准的性能度量。
- Agent 对环境的先验知识。
- Agent 可以完成的行动。
- Agent 截止到此时的感知序列。

从这可以导出**理性 Agent** 的定义：

对每一个可能的感知序列，根据已知的感知序列提供的证据和 Agent 具有的**先验知识**，理性 Agent 应该选择能使其性能度量最大化的行动。

2.2.1 理性

4个方面	
性能度量	每个时间单位对每块清洁的方块奖励1分
环境先验知识	只有左右移动，且不能移出方格
可采取的行动	左、右、吸尘
感知序列	位置、所在方格是否有灰尘

显而易见的是，同样的 Agent 在不同的环境下会变成非理性。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；如果性能度量包含对左右移动罚 1 分，该 Agent 的性能评价就会相当糟糕。这种情况下，一个更好的 Agent 应该在它确信所有的地方已经干净了以后不做任何事情。如果方格再次被弄脏了，该 Agent 应该不定期地检查并在必要的时候重新清洁。如果环境的地形未知的话，该 Agent 还需要去探查其他区域而不是固守方格 A 和 B。

2.2.2 全知者、学习和自主性

- 理性 \neq 全知
- 理性：使期望的性能最大化
- 权知：使实际的性能最大化

一天我沿着香榭丽舍大道散步，这时我看到了街对面的一位老朋友。当时附近没有车辆，我也没有别的事情，所以根据理性，我开始穿过马路。同时，在 33000 英尺的高空一扇货舱门从一架路过的飞机上掉了下来¹，并且在我到达马路对面之前拍扁了我。我穿过马路的决定难道是不理性的么？我的讣告中不太可能写上“试图穿行马路的傻瓜”。

- 在已有的数据基础上，尽可能挖掘足够多的信息

2.3 环境的性质

- 任务环境PEAS描述:
- 性能 (Performance)
- 环境 (Environment)
- 执行器 (Actuator)
- 传感器 (Sensor)

智能出租车Agent

Performance	Environment	Actuators	Sensors

智能卫星图像分析Agent

Performance	Environment	Actuators	Sensors
correct image categorization. 正确的图像归类	downlink from orbiting satellite. 轨道卫星的下行信道	display of scene categorization. 场景归类的显示	color pixel arrays. 颜色像素阵列

智能购物Agent

Performance	Environment	Actuators	Sensors
price 价格 quality 质量 appropriateness 合理性 efficiency 效率	Websites 网站 vendors 厂商 shippers 货主	display to user 商品展示 follow URL 跟随URL fill in form 填单	Webpages 网页 (text, (文本、 image, 图像、 scripts) 脚本)

2.3.2 任务环境的性质：可观察 (Observable)

完全可观察的与部分可观察的：如果 Agent 的传感器在每个时间点上都能获取环境的完整状态，那么我们就说任务环境是完全可观察的。如果传感器能够检测所有与行动决策相关的信息，那么该任务环境是有效完全可观察的；而相关的程序则取决于性能度量。完全可观察的环境很方便，因为 Agent 不需要维护任何内部状态来记录外部世界。噪音、不精确的传感器，或者传感器丢失了部分状态数据，都可能导致环境成为部分可观察的——例如，只有一个本地灰尘传感器的真空吸尘器 Agent 无法知道另一个方格是否有灰尘，自动驾驶出租车也无法了解到别的司机在想什么。如果 Agent 根本没有传感器，环境则是无法观察的。人们也许会认为 Agent 陷入了毫无希望的困境，但就像我们在第 4 章所讨论的，有时仍能确定地说，Agent 可能达到目标，

任务环境的性质：单或多Agent (Agents)

单 Agent 与多 Agent: 单 Agent 与多 Agent 环境之间的区别看上去很简单。例如，独自玩字谜游戏的 Agent 显然处于单 Agent 环境中，下国际象棋的 Agent 处于双 Agent 环境中。然而这里有些微妙。首先，我们说明了实体怎样可以被视为 Agent，但我们并没有解释哪些 Agent 必须被视为 Agent。Agent A（例如出租车司机）是否要把对象 B（另外一辆车）当作 Agent 对待，还是仅仅把它当作一个随机行动的对象，就像是海滩上的波浪或者风中摇摆的树叶？关键的区别在于 B 的行为是否寻求让依赖于 Agent A 的行为的性能度量值最大化。例如，下国际象棋时，对手 B 试图最大化它的性能度量，而根据国际象棋的规则，也就是要最小化 Agent A 的性能度量。因此，国际象棋是竞争性的多 Agent 环境。另一方面，在出租车驾驶的环境中，避免发生冲撞使得所有 Agent 的性能度量都最大化，所以它是一个部分合作的多 Agent 环境。它同时也是部分竞争的，如一辆车只能占据一个停车位。多 Agent 环境中的 Agent 设计问题往往与单 Agent 环境的相差甚远；例如，**通讯**经常作为理性行为出现在多 Agent 环境中；在一些竞争环境中，**随机行为**是理性的，原因是这样可以避免预测中的缺陷。

任务环境的性质：确定或不确定 (Deterministic)

确定的与随机的：如果环境的下一个状态完全取决于当前状态和 Agent 执行的动作，那么我们说该环境是确定的；否则，它是随机的。原则上说，Agent 在完全可观察的、确定的环境中无需考虑不确定性（在我们定义中，在多 Agent 环境中我们忽略了纯粹由其他 Agent 行动导致的不确定性；这样，尽管每个 Agent 都不能预测其他 Agent 的行动决策，游戏依然是确定的）。然而，如果环境是部分可观察的，那么它可能表现为随机的。大多数现实环境相当复杂，以至于难以跟踪到所有未观察到的信息；从实践角度考虑，它们必须处理成随机的。出租车驾驶的环境显然是随机的，因为无人能够精确预测交通状况；而且，车辆爆胎或者引擎失灵都是不可能事先预告的。我们前面描述的真空吸尘器世界是确定的，但是这个世界的变型可以包含一些随机元素，如随机出现的尘土和不可靠的吸尘机制（习题 2.13）。我们说环境**不确定**是指它不是完全可观察的或不确定的。最后谈一点：我们使用单词“随机”是为了暗示后果是不确定的并且可以用概率来量化；而不确定的环境中行动后果有多种可能，但与概率无关。不确定的环境的描述通常与要求 Agent 成功达成所有可能行动结果的性能度量相关。

任务环境的性质：片段式与延续式 (Episodic)

片段式的与延续式的：在片段式的任务环境中，Agent 的经历被分成了一个一个原子片段。在每个片段中 Agent 感知信息并完成单个行动。关键的是，下一个片段不依赖于以前的片段中采取的行动。很多分类任务属于片段式的。例如，装配线上检验次品零件的机器人每次决策只需考虑当前零件，不用考虑以前的行动决策；而且，当前决策也不会影响到下一个零件是否合格。与之相反，在延续式环境中，当前的决策会影响到所有未来的决策。¹下棋和出租车驾驶都是延续式的：在这两种情况中，短期的行动会有长期的效果。片段式的环境要比延续式环境简单得多，因为 Agent 不需要前瞻。

任务环境的性质：静态或动态 (Episodic)

静态的与动态的：如果环境在 Agent 计算的时候会变化，那么我们称该 Agent 的环境是动态的；否则环境则是静态的。静态环境相对容易处理，因为 Agent 在决策的时候不需要观察世界，也不需要顾虑时间的流逝。而动态的环境会持续地要求 Agent 做决策；如果 Agent 没有做出决策，Agent 则认为它决定不做任何事情。如果环境本身不随时间变化而变化，但是 Agent 的性能评价随时间变化，我们则称这样的环境是**半动态的**。出租车自动驾驶明显是动态的：即使驾驶算法对下一步行动犹豫不决，其他车辆和出租车自身仍然是不断运动的。国际象棋比赛的时候要计时，是半动态的。填字谜游戏是静态的。

任务环境的性质：离散的和连续的 (Discrete)

离散的和连续的： 环境的状态、时间的处理方式以及 Agent 的感知信息和行动，都有离散/连续之分。例如，国际象棋环境中的状态是有限的。国际象棋的感知信息和行动同时也是离散的。出租车驾驶是一个连续状态和连续时间问题：出租车和其他车辆的速度和位置都在连续空间变化，并且随时间的流逝而变化。出租车驾驶行动也是连续的（转弯角度等）。虽然严格来说，来自数字摄像头的输入信号是离散的，但处理时它表示的是连续变化的亮度和位置。

任务环境的性质：已知或未知

已知的与未知的：严格地说，这种区分指的不是环境本身，指的是 Agent（或设计人员）的知识状态，这里的知识则是指环境的“物理法则”。在已知环境中，所有行动的后果（如果环境是随机的，则是指后果的概率）是给定的。显然，如果环境是未知的，Agent 需要学习环境是如何工作的，以便做出好的决策。要注意的是已知环境和未知环境的区别，与完全可观察环境和部分可观察环境的区别有所不同。很可能已知的环境是部分可观察的——例如在翻牌游戏中，我知道所有的规则但仍然不知道未翻出的牌是什么。相反，未知的环境可能是完全可观察的——在玩新的视频游戏时，显示器上会给出所有的游戏状态，但我仍然不知道按钮的作用直到我真正试过。

任务环境类型

	智能出租车Agent	智能卫星图像分析Agent	智能购物Agent
完全可观察或部分可观察			
单Agent或多Agent			
确定或随机			
片段或延续			
静态或动态			
离散或连续			
已知或未知			

2.4 Agent的结构

- AI的任务是设计Agent程序，实现把感知信息映射到行动的Agent函数。

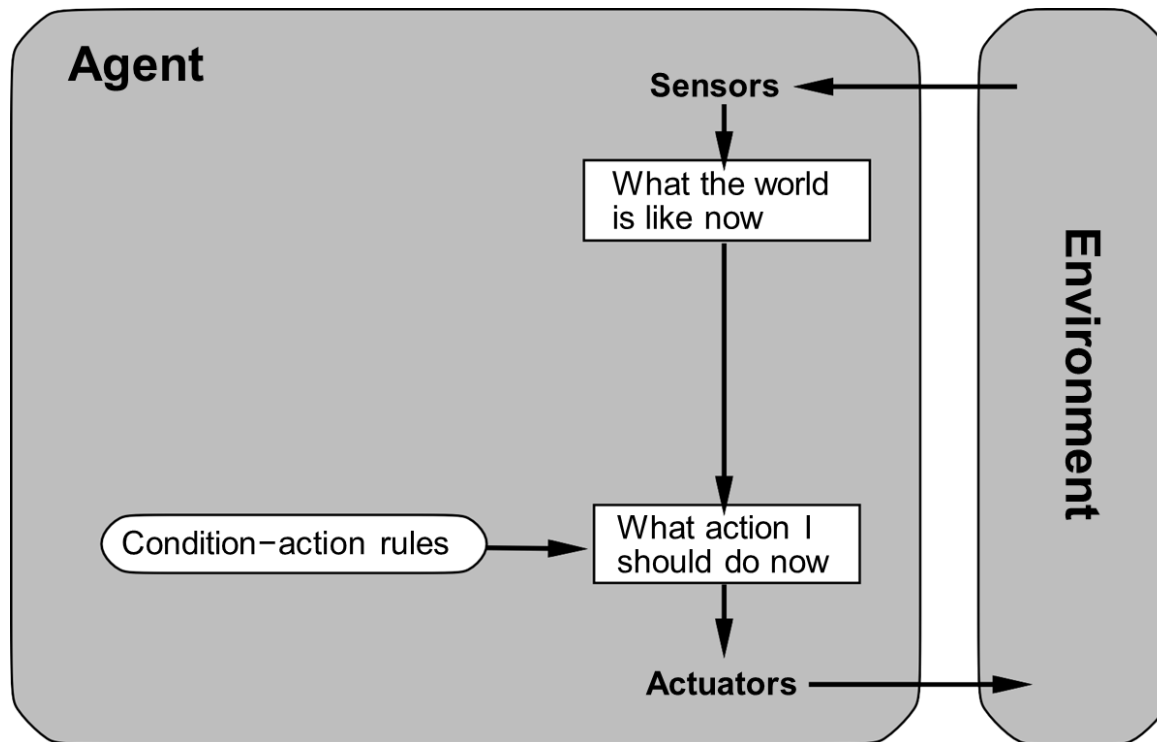
$$\text{Agent} = \text{体系结构} + \text{程序}$$

显然，我们选择的程序必须适合体系结构。如果程序要能够进行诸如行走这样的行动，那么体系结构最好有腿。体系结构可能只是普通的个人计算机，或者一辆自动驾驶汽车，车上有车载计算机、摄像头和其他传感器。一般而言，体系架构为程序提供来自传感器的感知信息，运行程序，并把程序计算出的行动决策送达执行器。

2.4.1 Agent程序

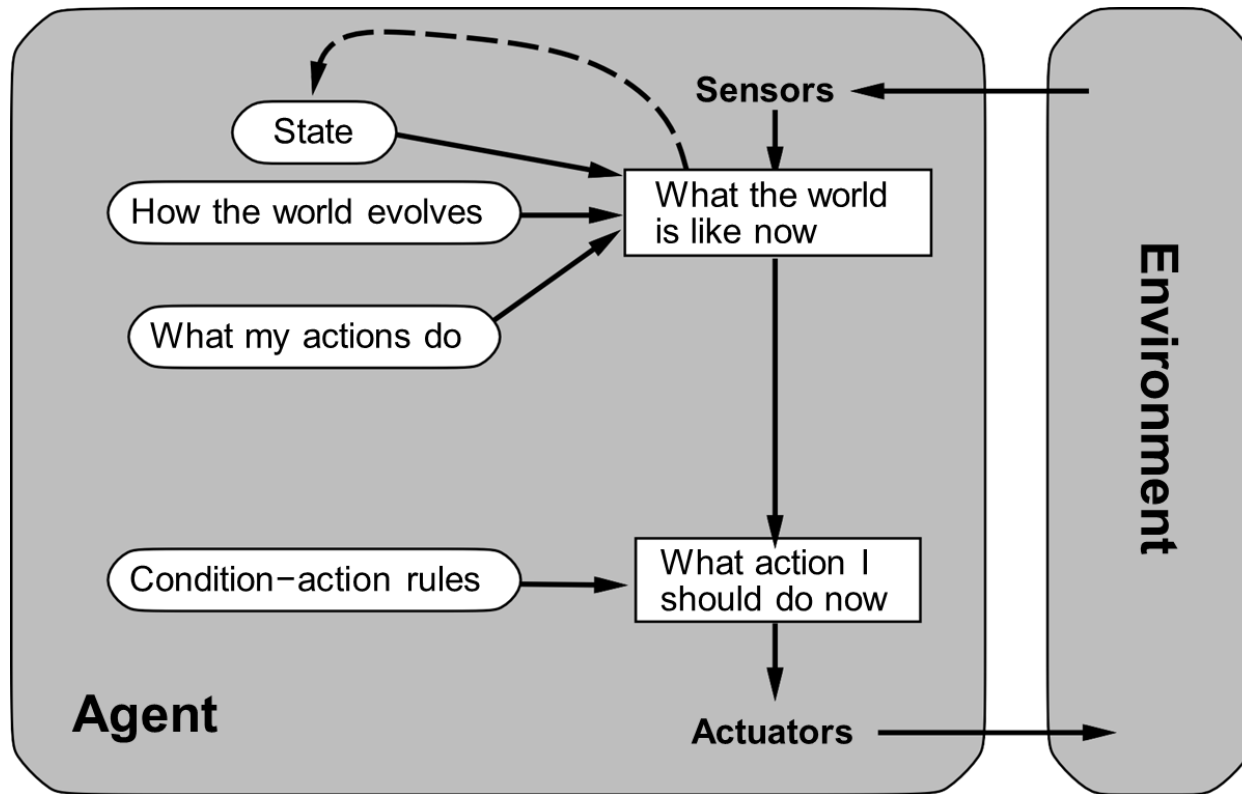
- 以下四种Agent程序能够涵盖几乎全部智能系统的基本原则：
 - 简单反射Agent
 - 基于模型的反射Agent
 - 基于目标的Agent
 - 基于效用的Agent
 - 学习Agent

2.4.2 简单反射Agent



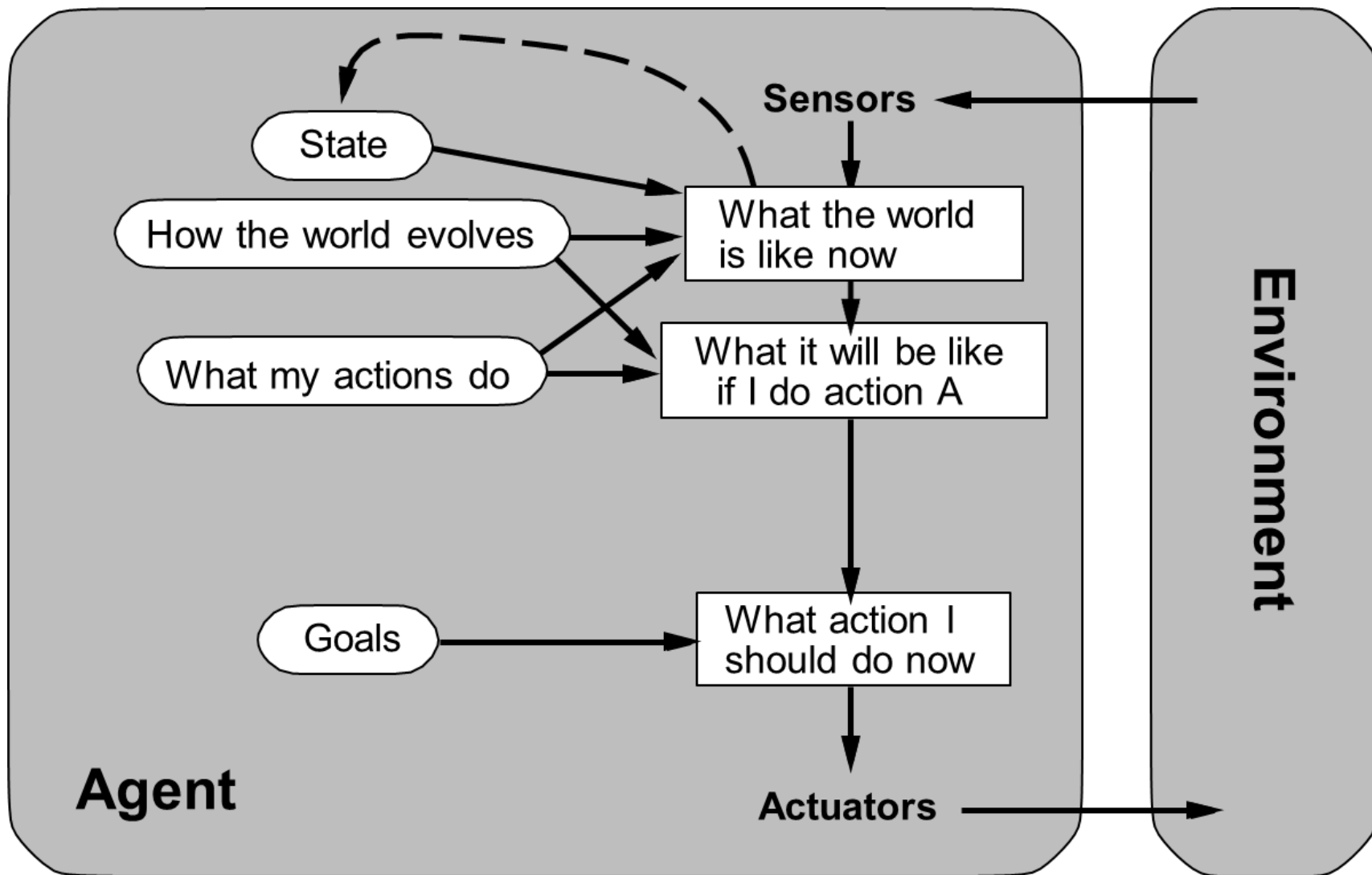
```
function Reflex-Vacuum-Agent([location,status]) returns an action
  if status = Dirty then return Suck
  else if location = A then return Right
  else if location = B then return Left
```

2.4.3 基于模型的反射Agent

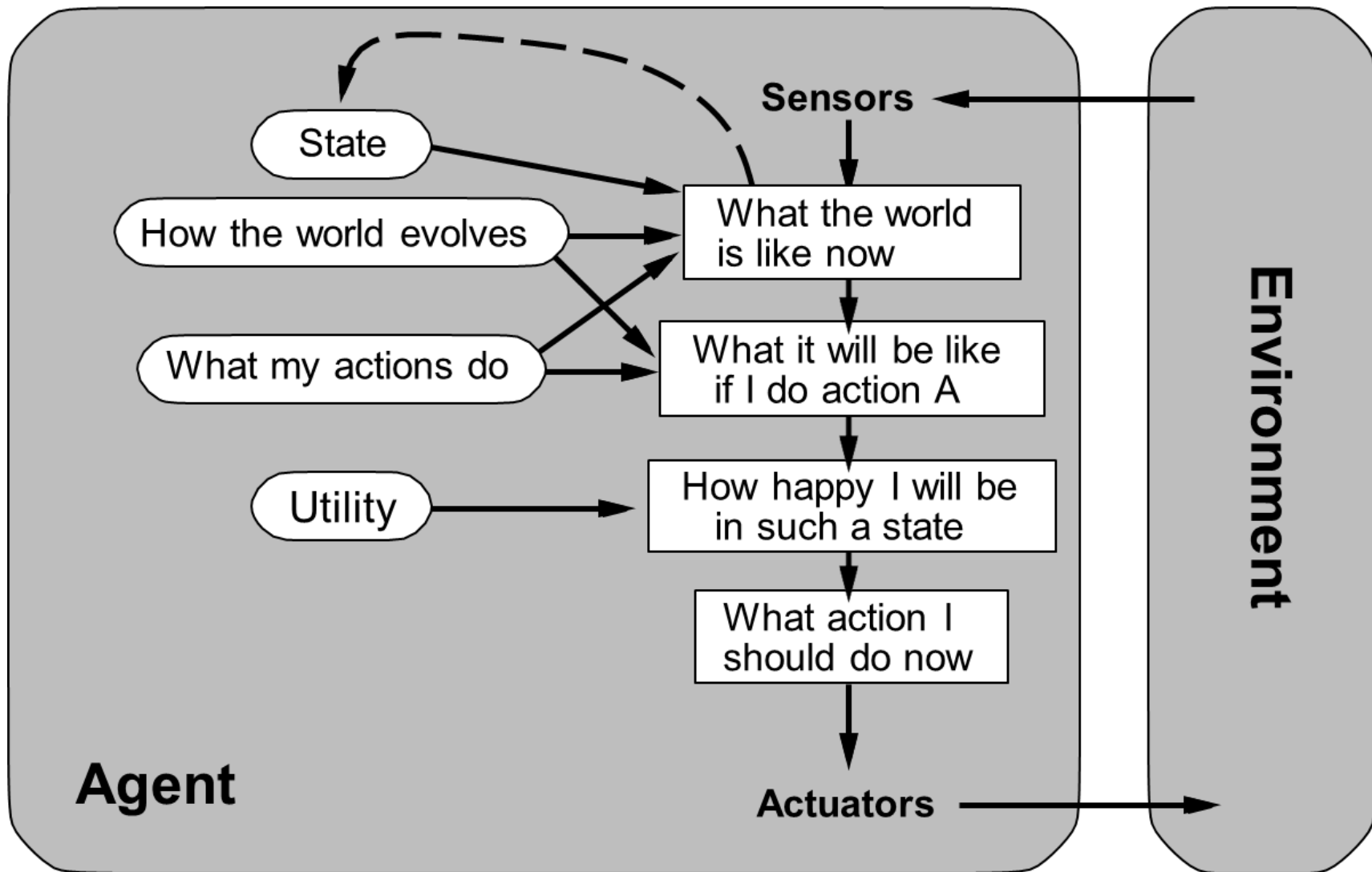


```
function Reflex-Vacuum-Agent( [location,status]) returns an action
static: last A, last B, numbers, initially  $\infty$ 
if status = Dirty then . . .
```

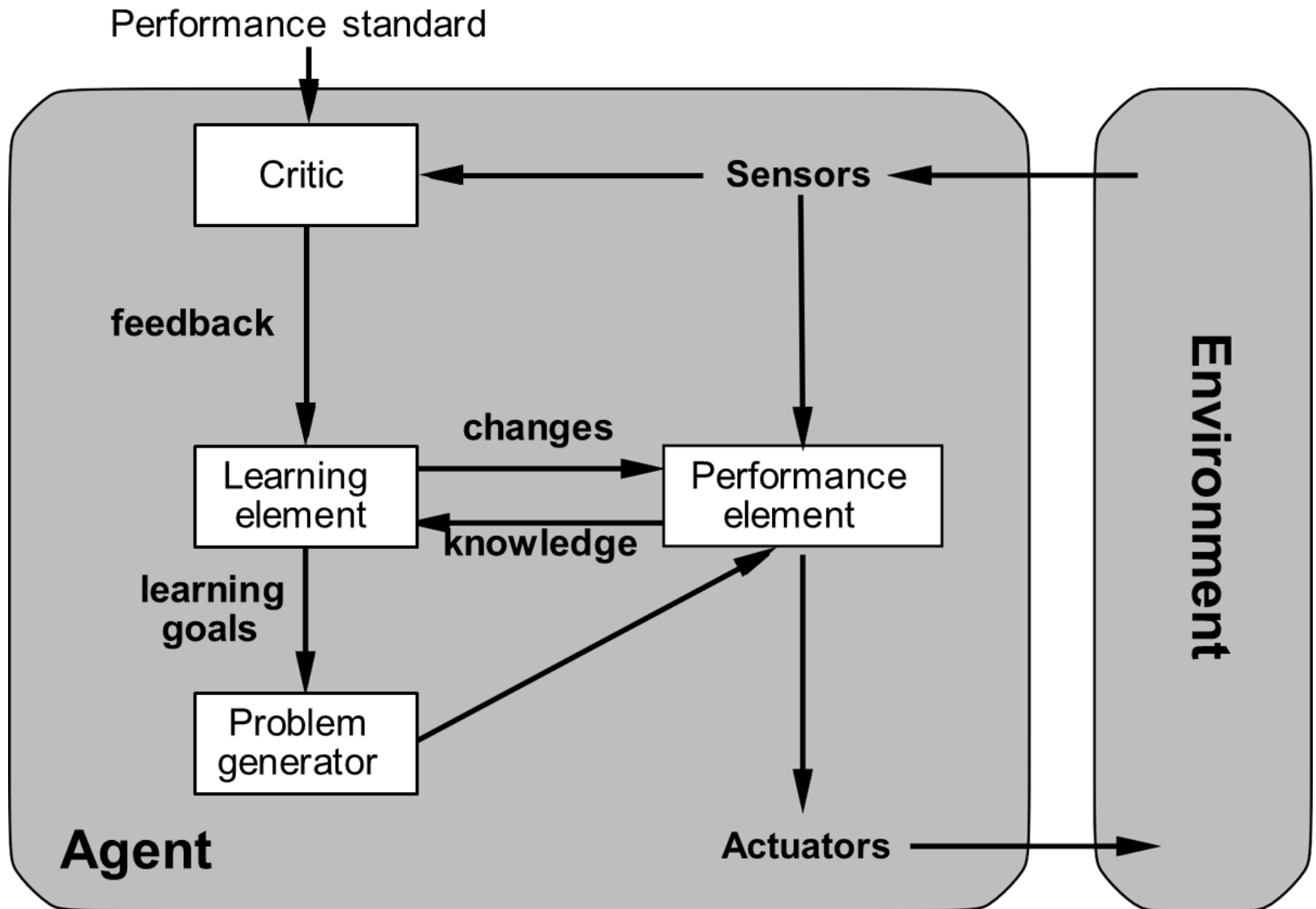
2.4.4 基于目标的Agent



2.4.5 基于效用的Agent



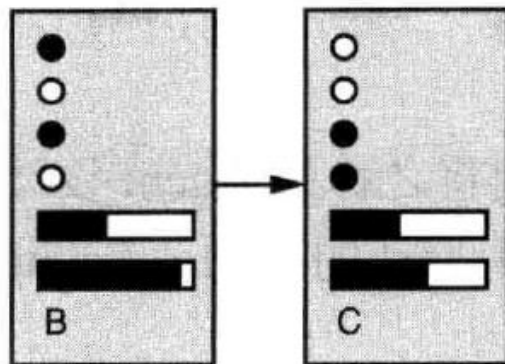
2.4.6 学习Agent



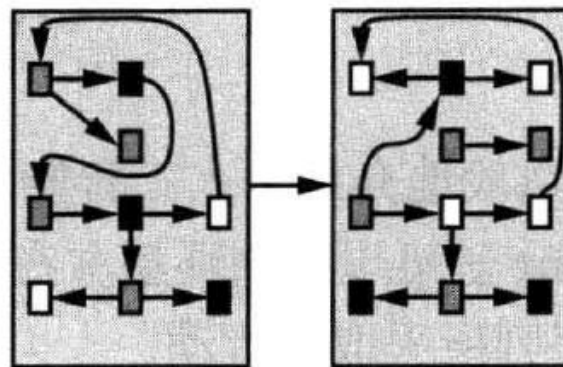
2.4.7 Agent程序的各组件如何工作



(a) Atomic

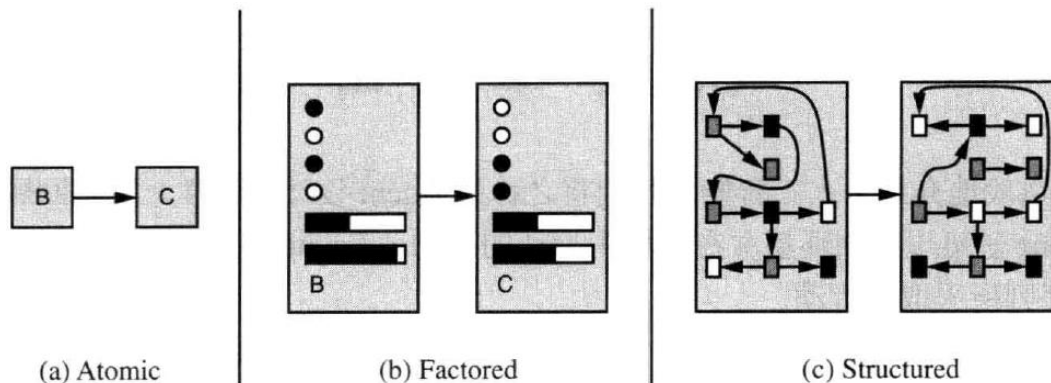


(b) Factored



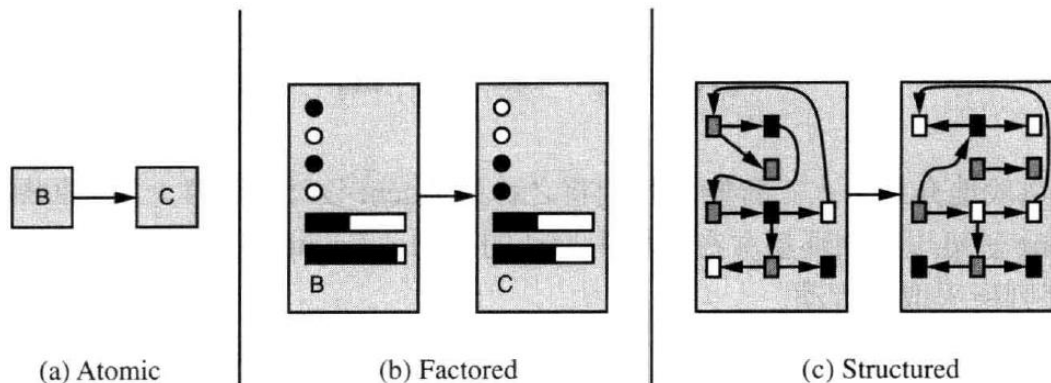
(c) Structured

原子表示



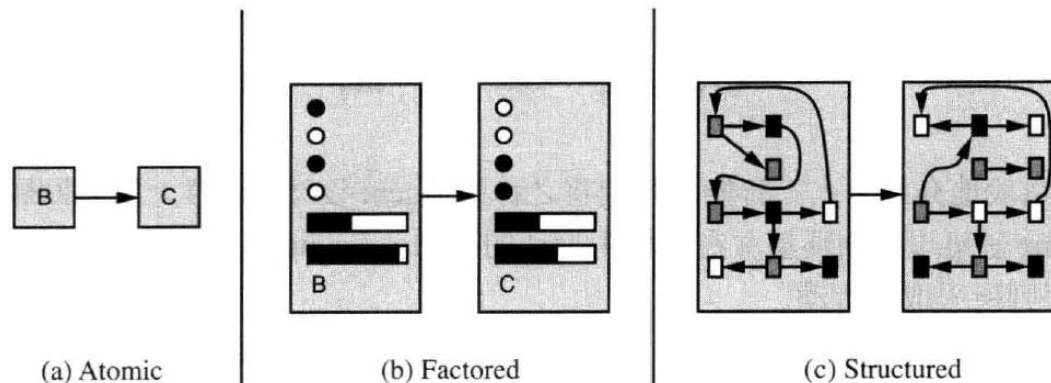
在原子表示中，世界的每个状态是不可见的——它没有内部结构。考虑从美国的一个边远的城市出来到达其他城市的路径问题（此问题可参见图 3.2）。为了求解这个问题，它可能将世界状态简化地表示为只有城市的名字——知道单个原子；“黑盒子”最明显的性质是与其他黑盒子相同或是不同。搜索和博弈论（第 3~5 章），隐马尔可夫模型（第 15 章），马尔可夫决策过程（第 17 章）中的算法都用的是原子表示——或者，至少把表示当作是原子的。

要素化表示



现在我们考虑同一个问题的高保真度描述，要比仅考虑城市的原子位置关注更多；我们可能需要注意油箱里的油还有多少，过收费站的零钱还有多少，广播里是哪个电台等等。要素化表示将上述状态表示为变量或特征的集合，每个变量或特征都可能值。当两个不同的原子表示没有任何共同点时——它们只是不同的黑盒子——两个不同的要素化表示可以共享一些特征（如一些特定的 GPS 位置），而不是其他的（有很多油或是没有油），这使得状态之间的转换工作变得更加容易。有了原子表示，我们还可以表示不确定性——例如，忽视油箱中的油量可表示为将该特征值置空。AI 的一些重要领域是基于原子表示的，包括约束满足算法（第 6 章），命题逻辑（第 7 章），规划（第 10 章和第 11 章），Bayesian 网（第 13~16 章）和第 18、20、21 章的机器学习算法。

结构化表示



我们为须理解世界中有事物，事物间互相关联，它不仅仅是有值的变量。例如，我们可能注意到前面有一辆大卡车正在调头进入奶牛场，但一头奶牛挡住了卡车的路。要素化表示可能不会准备好属性 *TruckAheadBackingIntoDairyFarmDrivewayBlockedByLooseCow* 的值真或假。因此，我们需要结构化表示，这样才能显式描述像奶牛和卡车这样的对象之间的关系。（见图 2.16 (c)）结构表示是关系数据库和一阶逻辑（第 8、9 和 12 章），一阶概率模型（第 14 章），基于知识的学习（第 19 章）和自然语言理解（第 22 章和第 23 章）的基础。事实上，人类用自然语言表述的信息几乎都与对象及其关系相关。

作业（第2次）

- 参考任务环境类型的格式，对下面几个Agent的环境进行分析：
 - 12306抢票Agent
 - AlphaGo围棋Agent
 - siri聊天Agent
 - 无人战斗机Agent
- 作业以word形式提交

推荐阅读

- Newell, A., & Simon, H. A. (1972). *Human problem solving* (Vol. 104, No. 9). Englewood Cliffs, NJ: Prentice-Hall. 《人类问题解决》
- Rao, A. S., & Wooldridge, M. (1999). Foundations of rational agency. In *Foundations of rational agency* (pp. 1-10). Springer, Dordrecht. 《理性Agent基础》
- 李志昂. (2010). *可控的Agent理性模型研究*. (Doctoral dissertation, 湖南大学).
- 刘勇. (2003). *多Agent系统理论和应用研究*. (Doctoral dissertation, 重庆大学).
- 郭亮. . *基于Agent的分布式计算研究*. (Doctoral dissertation, 重庆大学).

