

智能代理/智能体

张文生

中国科学院自动化研究所
中科院大学人工智能学院

2020年10月09日

讲课内容

一、智能体和环境

二、智能体理性概念

三、环境的性质

四、智能体的结构

讲课内容

一、智能体和环境

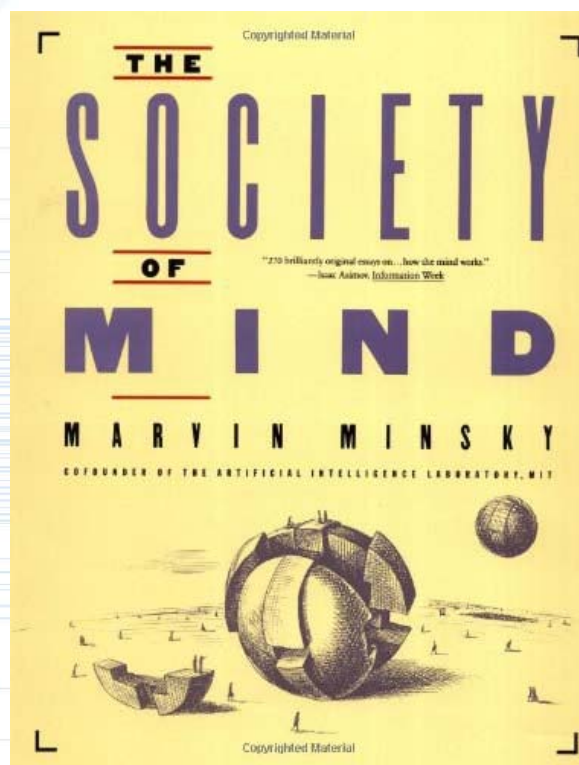
二、智能体理性的概念

三、环境的性质

四、智能体的结构

Agent概念

- 麻省理工学院的著名计算机学家和人工智能学科创始人Minsky的“Society of Mind/心智社会”一书
- 将社会与社会行为概念引入计算系统



作者: [Marvin Minsky](#)

出版社: Simon & Schuster

出版年: 1988-3-15

页数: 336

定价: USD 20.00

装帧: Paperback

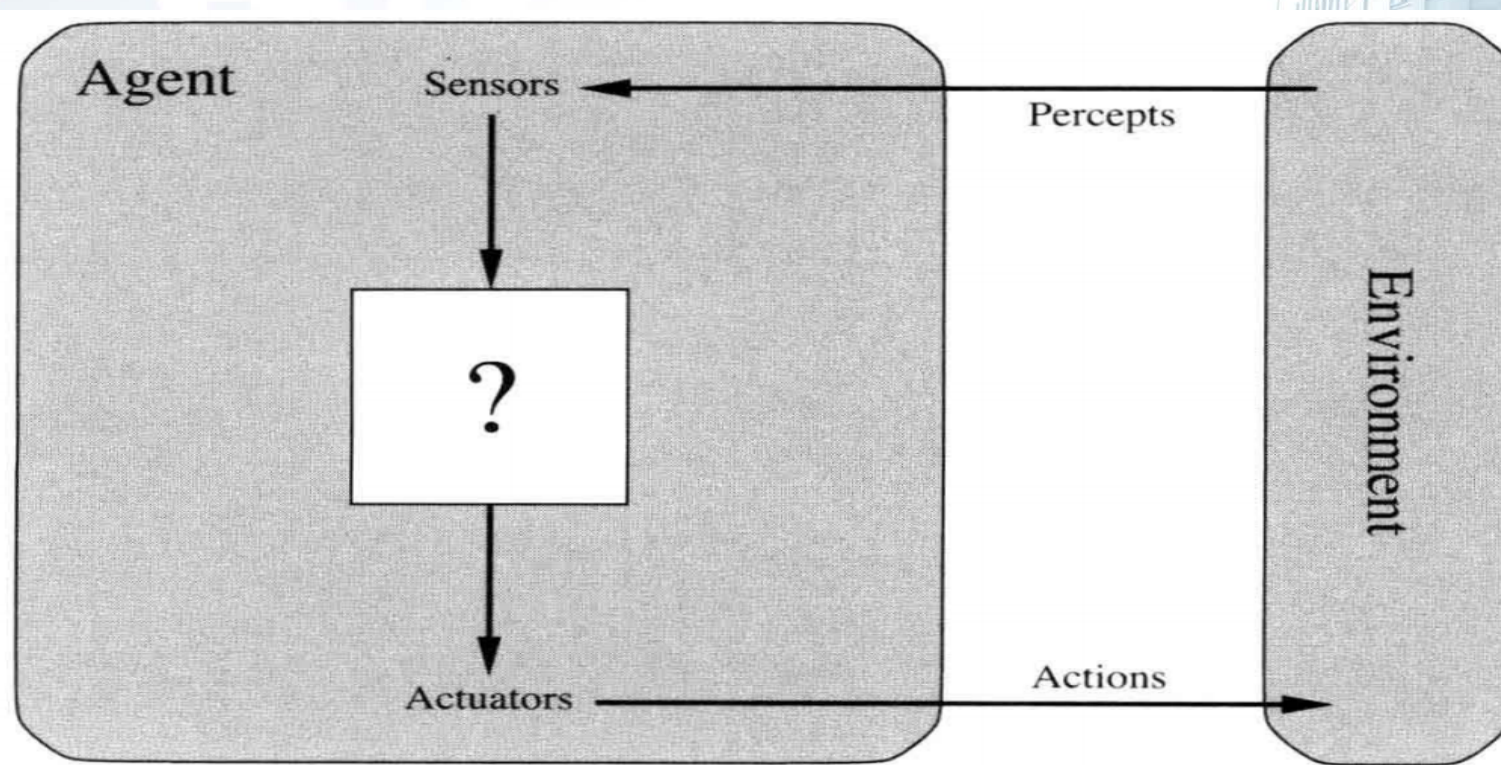
ISBN: 9780671657130

智能体的典型定义

- **Agent**是驻留于环境中的实体，可以解释从环境中获得的反映环境中所发生事件的数据，并执行对环境产生影响的行动
 - Agent技术标准化的组织：FIPA (Foundation for Intelligent Physical Agent)
- **Agent**是一个处于环境之中并且作为这个环境一部分的系统，它随时可以感测环境并且执行相应的动作，同时逐渐建立自己的活动规划以应付未来可能感测到的环境变化
 - Franklin, S. and Graesser, A.
 - 智能**Agent**能够持续执行三项功能：感知环境中的动态条件；执行动作影响环境条件；进行推理以解释感知信息、求解问题、产生推断和决定动作
 - 著名人工智能学者、美国斯坦福大学的Hayes-Roth

智能体和环境

- Agent通过传感器感知环境并通过执行器对所处环境产生影响



通过传感器和执行器与环境进行交互

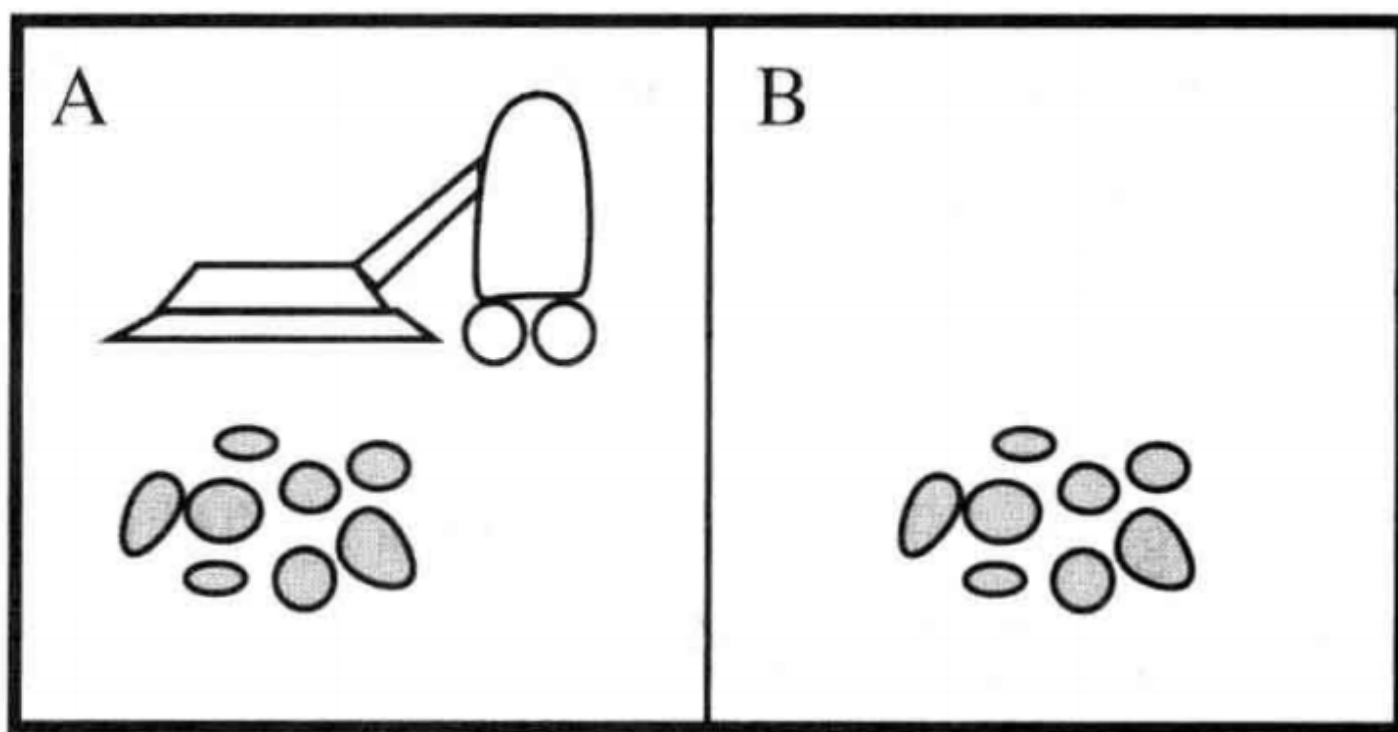
智能体和环境

- 人类Agent（眼睛、耳朵等**传感器**，手、腿、声道等**执行器**）
- 机器人Agent（摄像头、红外测距仪作为**传感器**、马达作为**执行器**）
- 软件Agent（键盘敲击、文件内容和网络数据包作为**传感器输入**，屏幕显示、写文件和发送网络数据包为**执行器**）

智能体和环境

- **感知**表示任何给定时刻Agent的感知输入
- **感知序列**是该Agent所收到的所有输入数据的完整历史
- **Agent函数**描述了Agent的行为，它将任意给定感知序列映射为行动

智能体和环境



A、B地点的真空吸尘器模型

智能体和环境

Percept sequence	Action
<i>[A, Clean]</i>	<i>Right</i>
<i>[A, Dirty]</i>	<i>Suck</i>
<i>[B, Clean]</i>	<i>Left</i>
<i>[B, Dirty]</i>	<i>Suck</i>
<i>[A, Clean], [A, Clean]</i>	<i>Right</i>
<i>[A, Clean], [A, Dirty]</i>	<i>Suck</i>
<i>⋮</i>	<i>⋮</i>
<i>[A, Clean], [A, Clean], [A, Clean]</i>	<i>Right</i>
<i>[A, Clean], [A, Clean], [A, Dirty]</i>	<i>Suck</i>
<i>⋮</i>	<i>⋮</i>

吸尘器模型的简单Agent函数的列表

讲课内容

一、智能体和环境

二、智能体理性的概念

三、环境的性质

四、智能体的结构

好的行为：理性的概念

- 理性Agent是做事正确，即Agent函数表格的每一项都填写正确
- 理性依据：任何指定的时刻
 - 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

1、理性

- **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
- Agent可以完成的行动
- Agent截止到此时的感知序列

- **理性Agent假设条件:**

- 性能度量在每个时间步对每块清洁的方格奖励1分，整个“生命”周期考虑1000个时间步
- 环境的“地形”作为先验知识是已知的，但灰尘的分布和Agent的初始位置未知
- 行动只有左、右和吸尘
- Agent能正确地感知位置及所在方格是否有灰尘

在上述条件下称Agent是理性的

2、全知者、学习和自主性

- **全知Agent**：明确地知道它的行动产生的实际结果并且做出相应的动作。
 - 有一天沿着大道散步，这时看到了对面的老朋友，附近没有车辆，我没有别的事情，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路过的飞机上掉了下来，在我到达马路对面之前拍扁了我
- 理性不等于完美
 - 理性是使期望的性能最大化，完美是使实际的性能最大化

- 理性Agent应该是自主的——它应该学习，弥补不完整的或者不正确的先验知识
- 实践中，很少要求Agent从一开始就完全自主
 - 当Agent没有或者只有很少的经验时，行为往往随机
 - 给人工智能Agent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Agent的行为才能独立于先验知识有效行动

3、环境的性质

- 任务环境包括性能度量 (Performance)、环境 (Environment) 以及Agent的执行器 (Actuators) 和传感器 (Sensors)
 - 自动驾驶出租车

Agent Type	Performance Measure	Environment	Actuators	Sensors
Taxi driver	Safe, fast, legal, comfortable trip, maximize profits	Roads, other traffic, pedestrians, customers	Steering, accelerator, brake, signal, horn, display	Cameras, sonar, speedometer, GPS, odometer, accelerometer, engine sensors, keyboard

自动驾驶出租车系统的PEAS描述

- **性能度量/目标**：到达正确的目的地；油量消耗和磨损最小化；到达目的地的时间或费用最小化；对交通法规的触犯和对其他司机的干扰最少化；安全性和乘客舒适度最高化；利润最高化
- **驾驶环境**：各种各样的道路（乡间小路、城市街巷，12车道的高速公路）、其他车辆、行人、游荡的动物、道路施工、警车、石头和坑洞等、潜在的和实际的乘客

- 自动出租车的**执行器**：方向盘、加速油门、刹车转向控制发动机、显示输出或者语音合成器与乘客交谈、同其他车辆交流
- **传感器**：一个或多个可控制的视频摄像头、红外或声呐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、全球卫星定位系统（GPS）、键盘或者麦克风

不同类型的智能体与PEAS描述

Agent Type	Performance Measure	Environment	Actuators	Sensors
Medical diagnosis system	Healthy patient, reduced costs	Patient, hospital, staff	Display of questions, tests, diagnoses, treatments, referrals	Keyboard entry of symptoms, findings, patient's answers
Satellite image analysis system	Correct image categorization	Downlink from orbiting satellite	Display of scene categorization	Color pixel arrays
Part-picking robot	Percentage of parts in correct bins	Conveyor belt with parts; bins	Jointed arm and hand	Camera, joint angle sensors
Refinery controller	Purity, yield, safety	Refinery, operators	Valves, pumps, heaters, displays	Temperature, pressure, chemical sensors
Interactive English tutor	Student's score on test	Set of students, testing agency	Display of exercises, suggestions, corrections	Keyboard entry

不同类型的Agent及PEAS描述

- 完全可观察的与部分可观察的
 - Agent的传感器在每个时间点上都能获取环境的完整状态，若传感器能够检测所有与行动决策相关的信息，称为有效完全可观察的
 - 噪音、不精确的传感器，或者传感器丢失了部分状态数据
- 单Agent与多Agent
 - 不同智能体是否寻求各自性能度量最大化

• 确定的与随机的

- **确定的**：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，该环境是确定的

- **随机的**：环境是部分可观察

- 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

- **不确定**是指它不是完全可观察的或不确定的

注：随机暗示后果是不确定的并且可以用概率量化，不确定的环境后果也有多种可能，但与概率无关

- 片段式的与延续式的

- 在片段式任务环境中，Agent的经历被分成一个个原子片段
- 每个片段Agent感知信息并完成单个行动，下一个片段不依赖于以前的片段中采取的行动（装配线上检验次品零件的机器人）
- 在延续式环境中，当前的决策会影响到所有未来的决策（下棋、出租车驾驶），短期行动会有长期效果

• 静态的与动态的

- **动态的**：如果环境在Agent计算的时候会变化（出租车自动驾驶），动态的环境会持续要求Agent做决策
- **静态的**：相对容易处理，在决策时不需要观察世界，不需要顾虑时间流逝（填字谜游戏）
- **半动态的**：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（国际象棋比赛）

• 离散的与连续的

- 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
 - 国际象棋中的状态有限，感知信息和行动同时也是离散
 - 出租车驾驶状态和时间皆连续：出租车和其他车辆的速度和位置在连续空间变化，并随时间流逝而变化，出租车驾驶行动也连续（转弯角度等）

• 已知的与未知的

- 指的不是环境本身，而是Agent/设计人员的知识状态，知识是指环境的“物理法则”
 - 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
 - 未知环境中，需要学习环境是如何工作的，以便做出好的决策
 - 已知的环境是部分可观察的（翻牌游戏，已知所有规则），未知的环境可能完全可观察（玩新的视频游戏，不知道按钮作用）

4、智能体的结构

- 前面的讨论是通过描述行为——在任何给定的感知序列下所采取的行动进行的，接下来讨论Agent内部是如何工作的
- AI的任务是设计Agent程序，把感知信息映射到行动的Agent函数
- 该程序要在某个具备物理传感器和执行器的计算装置上运行，称为体系结构

Agent=体系结构+程序

1、智能体程序

- Agent程序具有同样的框架：输入为从传感器得到的**当前**感知信息，返回的是执行器的行动抉择
- Agent函数是以整个感知历史作为输入的

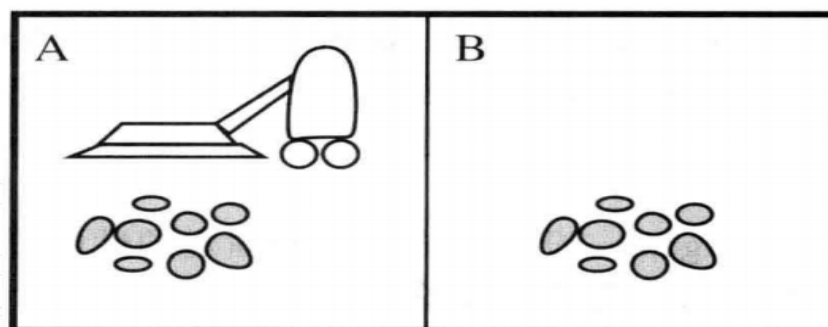
- 表驱动方法：以跟踪感知序列为索引，到行动表里查询以做出决策（国际象棋 10^{150} ）
 - 宇宙中没有实际Agent能够保存该表
 - 设计人员没有时间来创建该表
 - 没有Agent能够从经验中学习所有正确的表条目
 - 即使环境足够简单，表的容量可以接受，设计人员仍然没有向导来填写表中条目

```
function TABLE-DRIVEN-AGENT(percept) returns an action
  persistent: percepts, a sequence, initially empty
               table, a table of actions, indexed by percept sequences, initially fully specified

  append percept to the end of percepts
  action  $\leftarrow$  LOOKUP(percepts, table)
  return action
```

2、简单反射的智能体

- Agent基于当前的感知选择行动，不关注感知历史

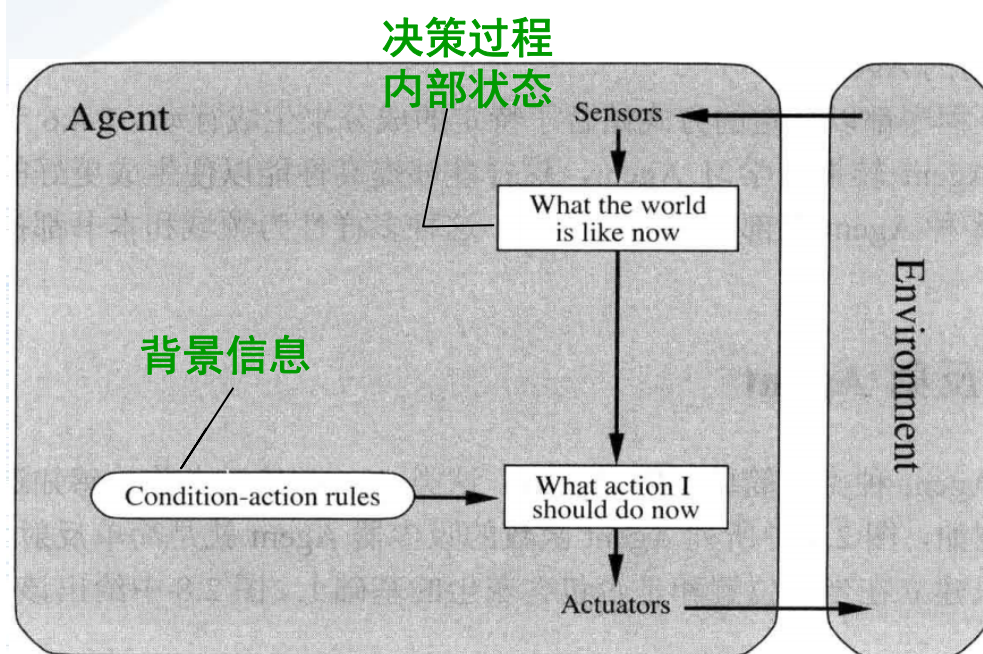


```
function REFLEX-VACUUM-AGENT([location,status]) returns an action
```

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

两状态真空吸尘器环境中的简单反射Agent的Agent程序

- 简单反射行为也会发生在更加复杂的环境中
 - 如果前方的车辆在刹车，那么开始刹车（这种联接称为条件—行为规则）



简单反射Agent示意图

```
function SIMPLE-REFLEX-AGENT(percept) returns an action
  persistent: rules, a set of condition-action rules

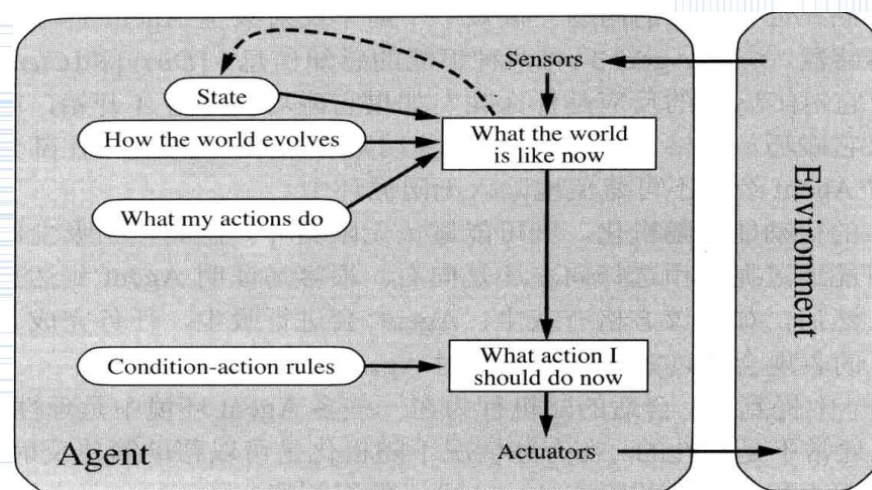
  state ← INTERPRET-INPUT(percept)
  rule ← RULE-MATCH(state, rules)
  action ← rule.ACTION
  return action
```

简单反射型**Agent**具有极好的简洁性，
但是智能有限，要求环境是完全可观察

3、模型反射的智能体

- 处理部分可观测环境的最有效途径——让Agent根据感知历史维持内部状态，跟踪记录现在看不到那部分世界
- 更新内部状态信息要求Agent程序中加入两种类型知识
 - 世界是如何独立于Agent而发展的信息
 - 正在超车的汽车在下一时刻更靠近本车
 - Agent自身的行动如何影响世界的信息
 - Agent顺时针转动方向盘，汽车会右转

- 这种关于“世界如何运转”的知识——用简单布尔电路还是用完备的科学理论实现——都被称为世界**模型**，使用这种模型的Agent被称为基于模型的Agent



模型反射Agent示意图

负责创建描述新的内部状态

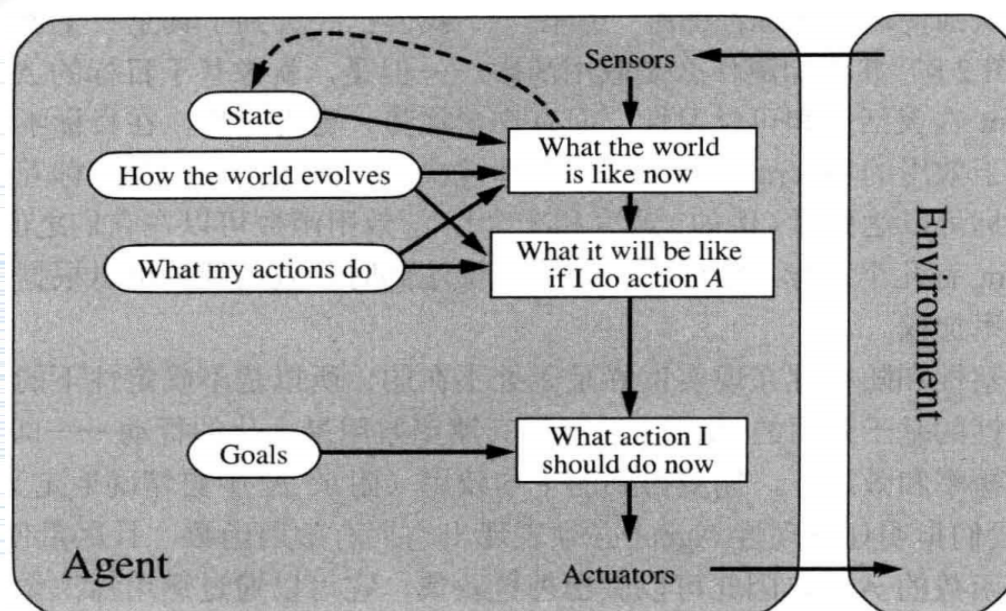
```
function MODEL-BASED-REFLEX-AGENT(percept) returns an action
  persistent: state, the agent's current conception of the world state
               model, a description of how the next state depends on current state and action
               rules, a set of condition-action rules
               action, the most recent action, initially none

  state ← UPDATE-STATE(state, action, percept, model)
  rule ← RULE-MATCH(state, rules)
  action ← rule.ACTION
  return action
```

使用内部模型跟踪世界的当前状态，按照与反射Agent相同的方式选择行动

4、目标的智能体

- 除了当前状态的描述，Agent还需要目标信息来描述想要达到的状况
 - 更灵活，支持决策的知识被显式表示，可以修改

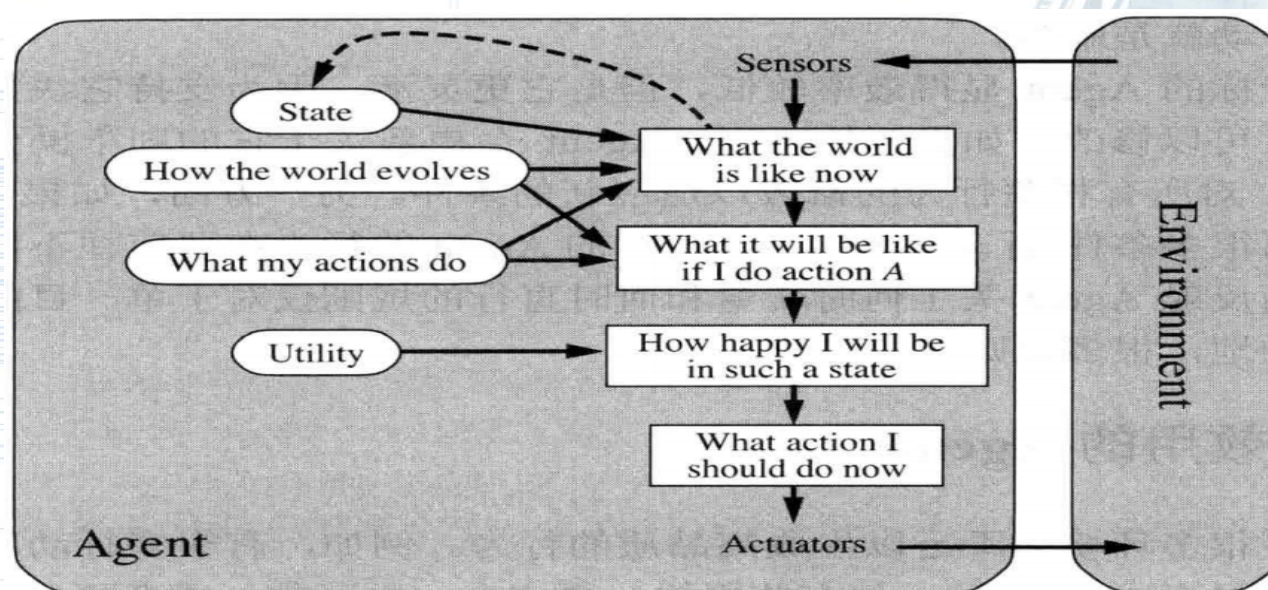


仅靠目标在很多环境中不足以生成高品质的行为，只提供粗略的二值区分，更快、更安全、更可靠、或者更便宜

基于模型和目标的Agent。既跟踪记录世界的状态，也记录要达到的目标集合，并选择能导致目标达到的行动

5、效用的智能体

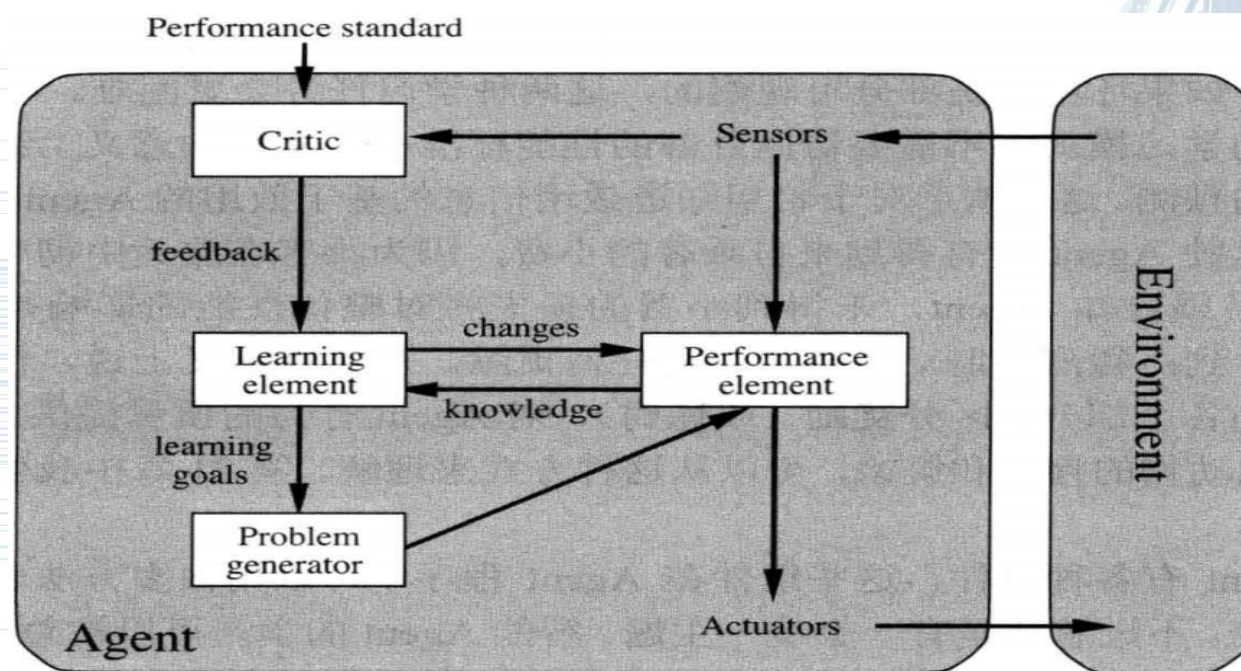
- 性能度量给环境状态的任何给定序列赋值
- 效用函数是性能度量的内在化，体现对各世界状态偏好程度
 - 当多个目标互相冲突时，可以在它们之间适当折中
 - 当有几个目标，根据目标的重要性对成功的似然率加权



基于模型和效用的Agent示意图

6、学习的智能体

- 学习使得Agent可以在初始未知的环境中运转，并逐渐变得比只具有初始知识的时候更有竞争力



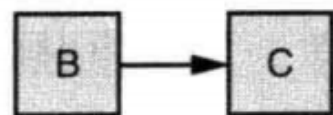
学习Agent的通用模型

• 智能出租车

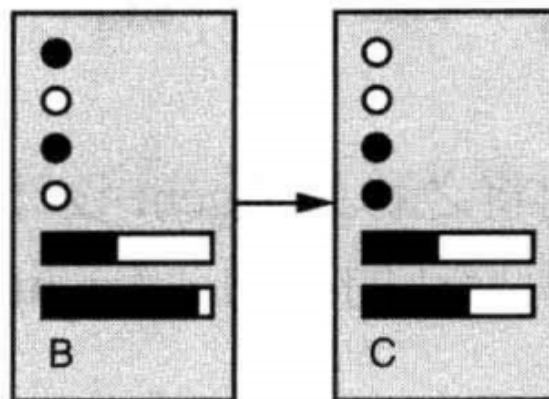
- 性能元件包括用来选择驾驶行动的全部知识和过程集合，出租车使用性能元件在公路上行驶
- 评判元件观察世界并把信息传递给学习元件（比如，横穿三条车道，其他司机惊呼）
- 学习元件制定出一条规则表示行动优劣
- 问题产生器为了改进的需要而确定需要修改一定范围内的行为，并提议进行试验（不同路面不同条件下试验刹车）

7、智能体程序如何工作

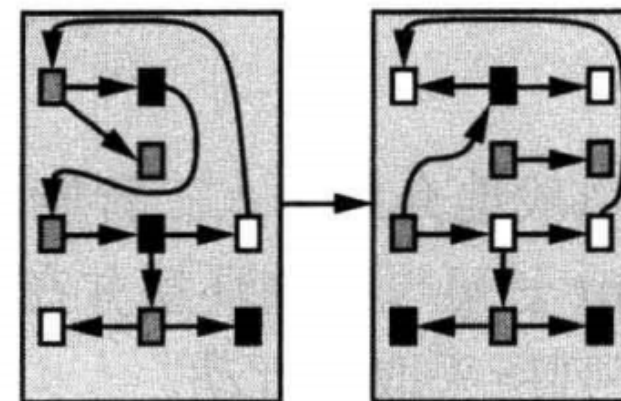
- 将表示放置在不断增长的复杂度和表达能力的轴线上——原子、要素和结构



(a) Atomic



(b) Factored



(c) Structured

- **原子表示：**搜索和博弈论（3-5章）、隐马尔可夫模型（15章）、马尔可夫决策过程（17章）
- **要素化表示：**约束满足算法（6章）、命题逻辑（7章）、规划（10、11章）、Bayesian网（13-16章）、机器学习算法（18、20、21章）
- **结构化表示：**关系数据库和一阶逻辑（8、9、12章）、一阶概率模型（14章）、基于知识的学习（19章）和自然语言理解（22、23章）

本次课程作业

- 教科书 P54—55

- 2.4

- 2.6

- 2.8

- 2.9

**感谢同学们听课
欢迎讨论与交流**