

第1章 人工智能基础2_智能体

南京信息工程大学
计算机学院

应龙
2025年秋季

主要内容

1. Agents and Environments

2. Good Behavior: The Concept of Rationality

3. The Nature of Environments

4. The Structure of Agents

Bibliography:

- [1] Stuart J. Russel, Peter Norvig, “Artificial Intelligence: A Modern Approach (4th Ed.)”, Pearson, 2020; 中译版“人工智能 现代方法 (第4版) ”, 人民邮电出版社, 2022. Ch 2.
- [2] 王万良 编著, “人工智能导论 (第5版)”, 高等教育出版社, 2020. Ch 9.1

Acting rationally: The rational agent approach

Agent 就是能够行动的某种东西（英语的 agent 源于拉丁语的 *agere*, 意为“去做”）。期望计算机 Agent 做更多的事：自主的操作，感知环境，长期持续，适应变化，并能创建与追求目标。

理性 Agent (rational agent) 是一个实现最佳结果，或者当存在不确定性时，为了实现最佳期望结果而行动的 Agent。

在对AI的“思维法则”的途径中，重点在正确的推理。做出正确的推理有时是作为理性 Agent 的一部分。但正确的推理并不是合理性的全部；在某些环境中，不要做可证正确的事情，但是仍然必须做某些事情。还有一些合理行动的方法不能被说成涉及推理。

图灵测试需要的所有技能也允许一个 Agent 理性地行动。

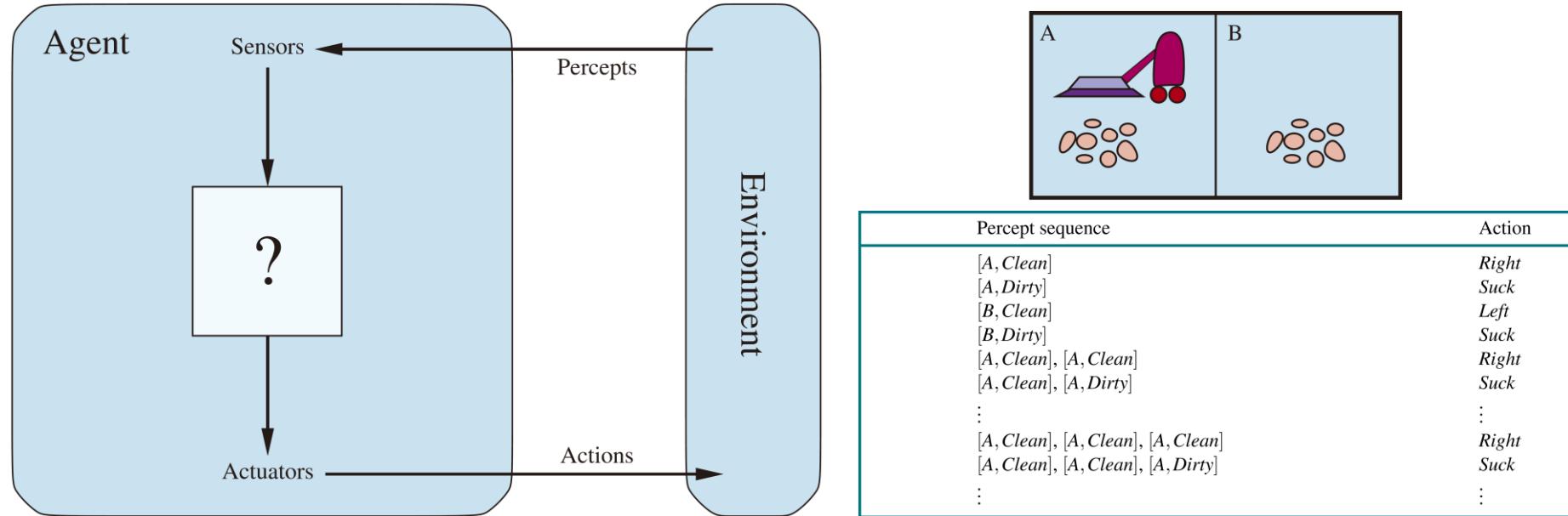
理性 Agent 的途径与其他途径相比有两个优点。

- ① 它比“思维法则”的途径更一般，因为正确的推理只是实现合理性的几种可能的机制之一。
- ② 它比其他基于人类行为或人类思维的途径更经得起科学发展的检验。

理性的标准在数学上定义明确且完全通用，并可被“解开并取出”来生成可证实实现了理性的Agent设计。另一方面，人类行为可以完全适应特定环境，并且可以很好地定义为人类做的所有事情的总和。

Agents and Environments

Agent 通过传感器感知环境并通过致动器对所处环境产生影响。



我们用 **感知 (percept)** 来表示任何给定时刻 Agent 的感知输入。Agent 的**感知序列 (percept sequence)** 是该 Agent 所收到的所有输入数据的完整历史。一般地，Agent 在任何给定时刻的行动选择依赖于到那个时刻为止该 Agent 的整个感知序列。

Agent 函数 (Agent function) 描述了 Agent 的行为，它将任意给定感知序列映射为行动。可以通过对 Agent 函数制表 (tabulating) 来描述任何给定的 Agent。

表格是从外部描述 Agent 的特性。从 Agent 内部来看，Agent 函数通过 **Agent 程序 (Agent program)** 实现。Agent 函数是抽象的数学描述；Agent 程序则是具体实现，它在某个物理系统内部运行。

Good Behavior: The Concept of Rationality

一个**理性 Agent (rational agent)** 是做正确的事的 Agent。从概念上讲，Agent函数表格的每一项都填写正确。

什么是做正确的事 (to do **the right thing**) ? consequentialism 结果主义
考虑 Agent 行为的后果。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。如果该系列状态变化正是希望的，那么这个Agent 性能良好。这里的希望，通过性能度量 (performance measure) 表述，它对环境状态的任何给定序列进行评估。

Humans have **desires and preferences of their own**, so the notion of rationality as applied to humans has to do with their success in choosing actions that produce sequences of environment states that are **desirable from their point of view**.
Machines, on the other hand, do not have desires and preferences of their own; the performance measure is, initially at least, **in the mind of the designer of the machine**, or **in the mind of the users the machine is designed for**.

对所有的任务和 Agent，没有一成不变的固定的 performance measure；典型地，设计人员会具体问题具体分析。

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，而不是根据人们认为Agent应该如何做出行为。

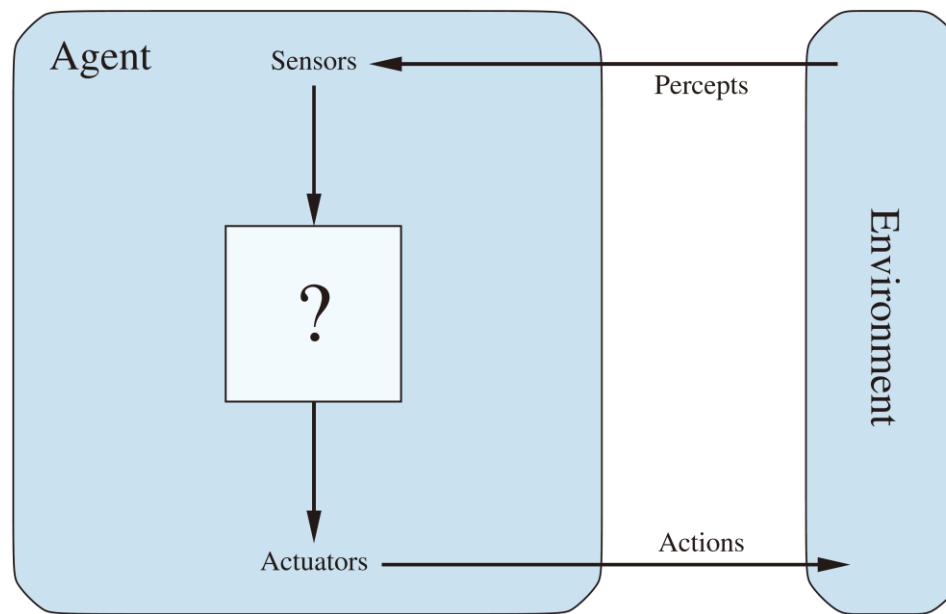
Good Behavior: The Concept of Rationality

任何给的时刻，理性决策的内容依赖于以下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。

Rational agent: 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。

同样的Agent在不同的环境下会变成非理性。



Good Behavior: The Concept of Rationality

对理性的定义并不要求全知 (omniscience)，因为理性的选择只依赖于到当时为止的感知序列。

Rationality maximizes expected performance, while perfection maximizes actual performance.

观察有助于最大化期望性能。为了改变未来的感知 (future percepts) 而采取行动，有时称为信息收集 (information gathering)，这是理性的重要部分

理性的定义不仅要求理性 Agent 收集信息，而且要求Agent从它所感知的信息中尽可能多的学习 (to learn)。Agent 最初的设定可能反映的是环境的先验知识，但随着 Agent 经验的丰富，这些知识会被改变或者增加。

在一些极端情况下，环境是先验已知和完全可预测的。智能体不需要感知或学习，只需正确地运行。但这样的智能体是脆弱的。

Agent 依赖于设计人员的先验知识而不是它自身的感知信息，这种情况我们会说该Agent 缺乏自主性 (autonomy)。理性 Agent 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。

主要内容

1. Agents and Environments

2. Good Behavior: The Concept of Rationality

3. The Nature of Environments

4. The Structure of Agents

Bibliography:

- [1] Stuart J. Russel, Peter Norvig, “Artificial Intelligence: A Modern Approach (4th Ed.)”, Pearson, 2020; 中译版“人工智能 现代方法 (第4版) ”, 人民邮电出版社, 2022. Ch 2.
- [2] 王万良 编著, “人工智能导论 (第5版)”, 高等教育出版社, 2020. Ch 9.1

The Nature of Environments

任务环境 (task environment): 理性 Agent 要 “求解” 的 “问题” 。

在简单的吸尘器 Agent 的讨论中，必须把性能度量、环境以及 Agent 的执行器和传感器。把所有这些归在一起，都属于任务环境。

任务环境的规范包括性能度量(Performance)、外部环境(Environment)、执行器(Actuators) 和传感器(Sensors)。根据首字母缩写，称为 PEAS 描述。设计 Agent 时，第一步就是尽可能完整地详细说明任务环境。

Agent Type	Performance Measure	Environment	Actuators	Sensors
Taxi driver	Safe, fast, legal, comfortable trip, maximize profits, minimize impact on other road users	Roads, other traffic, police, pedestrians, customers, weather	Steering, accelerator, brake, signal, horn, display, speech	Cameras, radar, speedometer, GPS, engine sensors, accelerometer, microphones, touchscreen

Agent Type	Performance Measure	Environment	Actuators	Sensors
Taxi driver	Safe, fast, legal, comfortable trip, maximize profits, minimize impact on other road users	Roads, other traffic, police, pedestrians, customers, weather	Steering, accelerator, brake, signal, horn, display, speech	Cameras, radar, speedometer, GPS, engine sensors, accelerometer, microphones, touchscreen
Medical diagnosis system	Healthy patient, reduced costs	Patient, hospital, staff	Display of questions, tests, diagnoses, treatments	Touchscreen/voice entry of symptoms and findings
Satellite image analysis system	Correct categorization of objects, terrain	Orbiting satellite, downlink, weather	Display of scene categorization	High-resolution digital camera
Part-picking robot	Percentage of parts in correct bins	Conveyor belt with parts; bins	Jointed arm and hand	Camera, tactile and joint angle sensors
Refinery controller	Purity, yield, safety	Refinery, raw materials, operators	Valves, pumps, heaters, stirrers, displays	Temperature, pressure, flow, chemical sensors
Interactive English tutor	Student's score on test	Set of students, testing agency	Display of exercises, feedback, speech	Keyboard entry, voice

The Nature of Environments

◆ 任务环境的性质

完全可观察的与部分可观察的（Fully observable vs. partially observable）：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。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。

单智能体与多智能体（Single-agent vs. multiagent）：关键的区别在于某个Agent的行为是否寻求让依赖于其他Agent的行为的性能度量值最大化。竞争(competitive)与合作(cooperative) - 博弈论

确定的与随机的（Deterministic vs. nondeterministic）：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。

片段式的（Episodic）与延续式的（sequential）：在片段式的任务环境中，Agent的经历被分成了一个个原子片段。在每个片段中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。很多分类任务属于片段式的。在延续式环境中，当前的决策会影响到所有未来的决策。

The Nature of Environments

静态的与动态的 (Static vs. dynamic) : 如果环境在 Agent 思考的时候会变化 (If the environment can change while an agent is deliberating) , 那么我们称该 Agent 的环境是动态的; 否则环境则是静态的。

离散的与连续的 (Discrete vs. continuous) : 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。

已知的与未知的 (Known vs. unknown) : 严格地说，这种区分指的是 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。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。

最难处理的情况就是部分可观测的 (partially observable) 、多智能体的 (multiagent) 、随机的 (nondeterministic) 、延续的 (sequential) 、动态的 (dynamic) 、连续的 (continuous) 和未知的 (unknown) 环境。

The Nature of Environments

Task Environment	Observable	Agents	Deterministic	Episodic	Static	Discrete
Crossword puzzle	Fully	Single	Deterministic	Sequential	Static	Discrete
Chess with a clock	Fully	Multi	Deterministic	Sequential	Semi	Discrete
Poker	Partially	Multi	Stochastic	Sequential	Static	Discrete
Backgammon	Fully	Multi	Stochastic	Sequential	Static	Discrete
Taxi driving	Partially	Multi	Stochastic	Sequential	Dynamic	Continuous
Medical diagnosis	Partially	Single	Stochastic	Sequential	Dynamic	Continuous
Image analysis	Fully	Single	Deterministic	Episodic	Semi	Continuous
Part-picking robot	Partially	Single	Stochastic	Episodic	Dynamic	Continuous
Refinery controller	Partially	Single	Stochastic	Sequential	Dynamic	Continuous
English tutor	Partially	Multi	Stochastic	Sequential	Dynamic	Discrete

环境的实现和通用的环境仿真器 (general-purpose environment simulator)

Experiments are often carried out not for a single environment but for many environments drawn from an environment class. We are then interested in the agent's average performance over the environment class.

主要内容

1. Agents and Environments

2. Good Behavior: The Concept of Rationality

3. The Nature of Environments

4. The Structure of Agents

Bibliography:

- [1] Stuart J. Russel, Peter Norvig, “Artificial Intelligence: A Modern Approach (4th Ed.)”, Pearson, 2020; 中译版“人工智能 现代方法 (第4版) ”, 人民邮电出版社, 2022. Ch 2.
- [2] 王万良 编著, “人工智能导论 (第5版)”, 高等教育出版社, 2020. Ch 9.1

The Structure of Agents

AI 的任务是设计 Agent 程序， 它实现的是把感知信息映射到行动的 Agent 函数。假设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，我们称为体系结构 (architecture):

$$\text{agent} = \text{architecture} + \text{program}$$

我们选择的程序必须适合体系结构。

◆ Agent 程序

输入为从传感器得到的当前感知信息， 返回的是制动器的行动抉择。
Agent 程序以当前感知为输入， 这与Agent 函数不同， Agent 函数是以整个感知历史作为输入的。Agent程序只把当前感知作为输入是因为从环境无法得到更多信息；如果Agent的行动要依赖于整个感知序列， 那么该Agent 必须要记住感知信息。

The table-driven approach: 构建出包含针对每个可能感知序列的适当行动的函数表。

穷举 

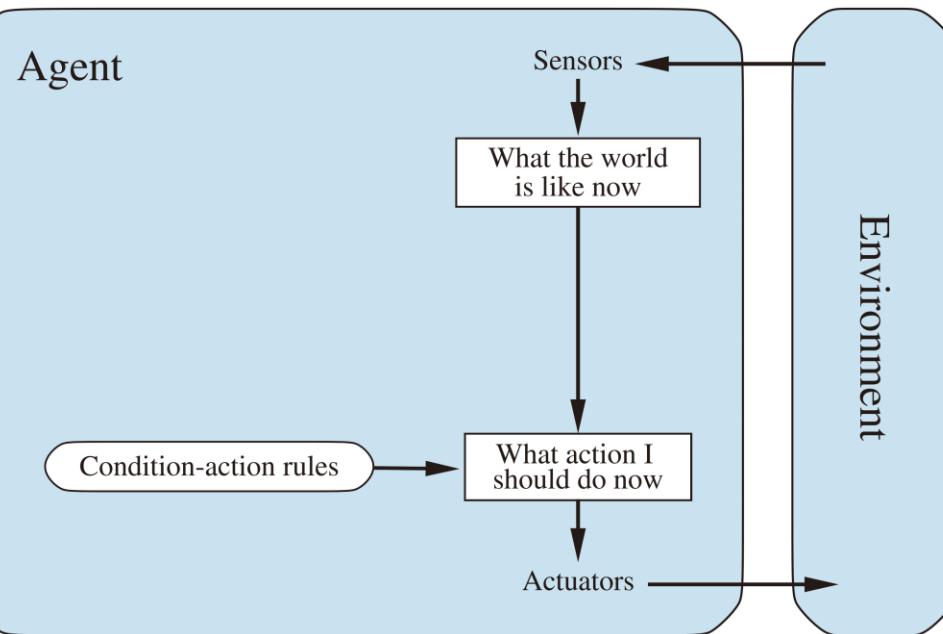
- 简单反射 Simple reflex agents
- 基于模型的反射 Model-based reflex agents
- 基于目标的 Goal-based agents
- 基于效用的 Utility-based agents

The key challenge for AI is to find out how to write programs that, to the extent possible, produce rational behavior from a smallish program rather than from a vast table.

The Structure of Agents

● Simple reflex agents

基于当前的感知选择行动，不关注感知历史。



```
function SIMPLE-REFLEX-AGENT(percept) returns an action
  persistent: rules, a set of condition–action rules
    state  $\leftarrow$  INTERPRET-INPUT(percept)
    rule  $\leftarrow$  RULE-MATCH(state, rules)
    action  $\leftarrow$  rule.ACTION
  return action
```

Figure 2.10 A simple reflex agent. It acts according to a rule whose condition matches the current state, as defined by the percept.

环境要是完全可观察的。即使有少量不可观测的情况也会带来麻烦。在部分可观察环境中运转的简单反射 Agent 经常不可避免地陷入无限循环中。如果 Agent 的行动能够随机化，则可能避免无限循环。

The Structure of Agents

● Model-based reflex agents

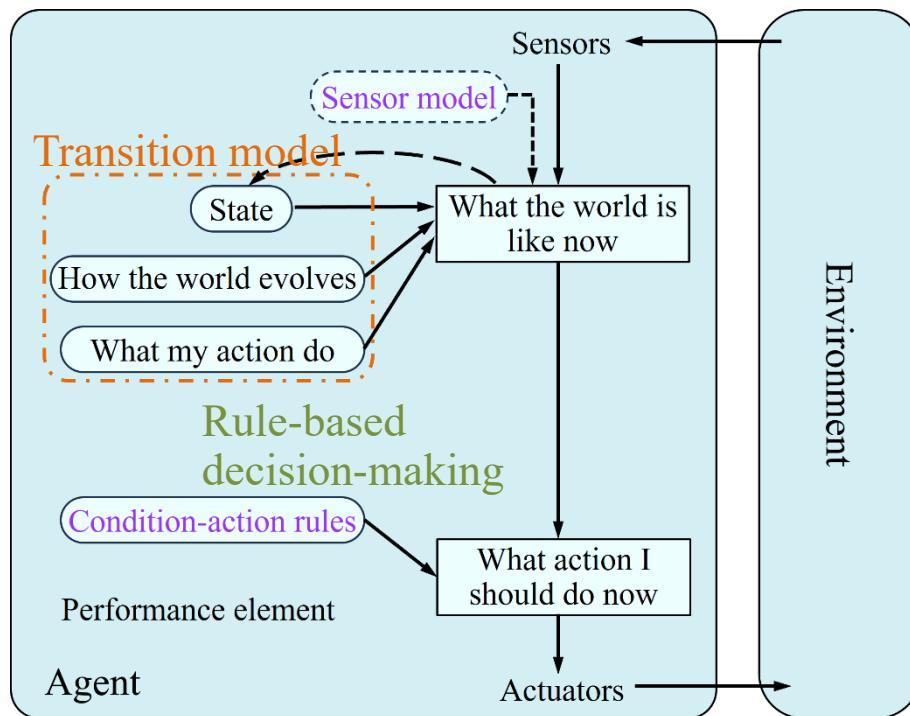
处理部分可观测环境的最有效途径是让Agent跟踪记录现在看不到的那部分世界。Agent应该根据感知历史维持内部状态，从而至少反映出当前状态看不到的信息。随时更新内部状态信息要求在Agent程序中加入两种类型的知识：

1) 世界是如何独立于Agent而发展的信息；

2) Agent自身的行动如何影响世界的信息。世界模型

transition modal

how the state of the world is reflected in the agent's percepts. sensor modal



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

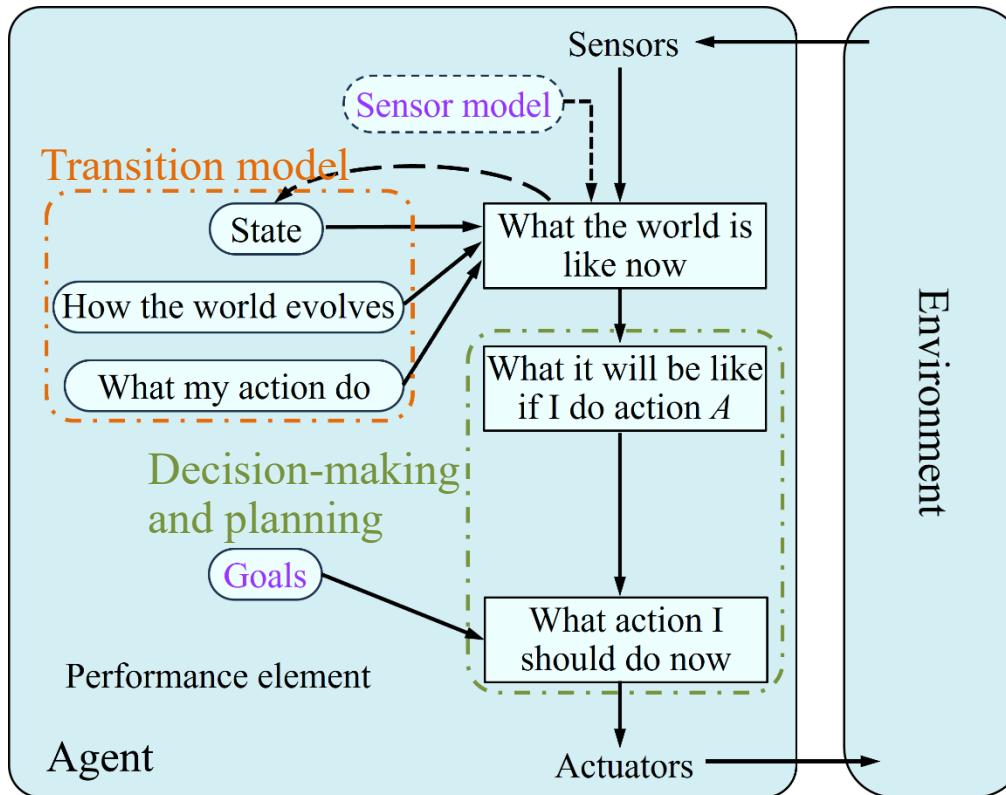
  state ← UPDATE-STATE(state, action, percept, transition_model, sensor_model)
  rule ← RULE-MATCH(state, rules)
  action ← rule.ACTION
  return action
```

Figure 2.12 A model-based reflex agent. It keeps track of the current state of the world, using an internal model. It then chooses an action in the same way as the reflex agent.

The Structure of Agents

● Goal-based agents

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。Agent 程序可以把这种信息和模型相结合，以选择能达到目标的行动。

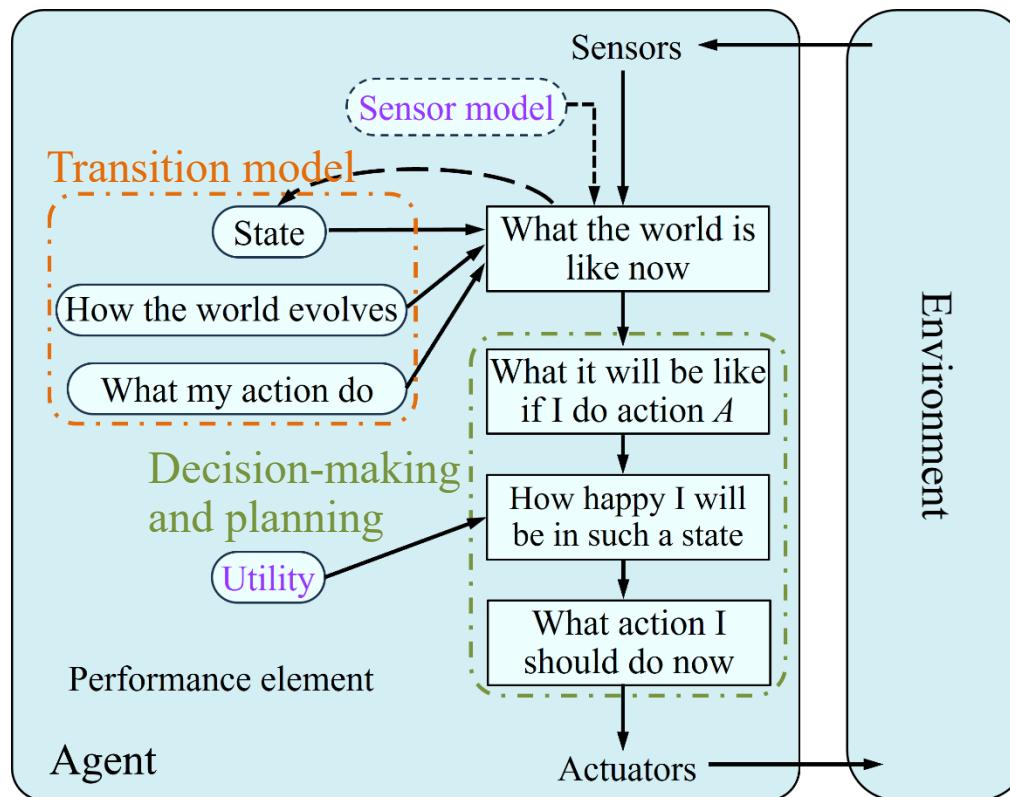


此类决策与前面描述的条件-行动规则有根本的不同，它考虑了未来。基于目标的Agent原则上会推理。支持它决策的知识被显式表示出来，并且可以修改。

The Structure of Agents

● Utility-based agents

仅靠目标在很多环境中不足以生成高品质的行为。更通用的性能度量应该允许不同的世界状态之间进行效用比较。性能度量给环境状态的任何给定序列赋了一个值。Agent 的效用函数是性能度量的内在化。



拥有显式效用函数的 Agent 因此可以做出理性决策，它可以通过通用算法做到，且此算法并不依赖于要最大化的特定效用函数。

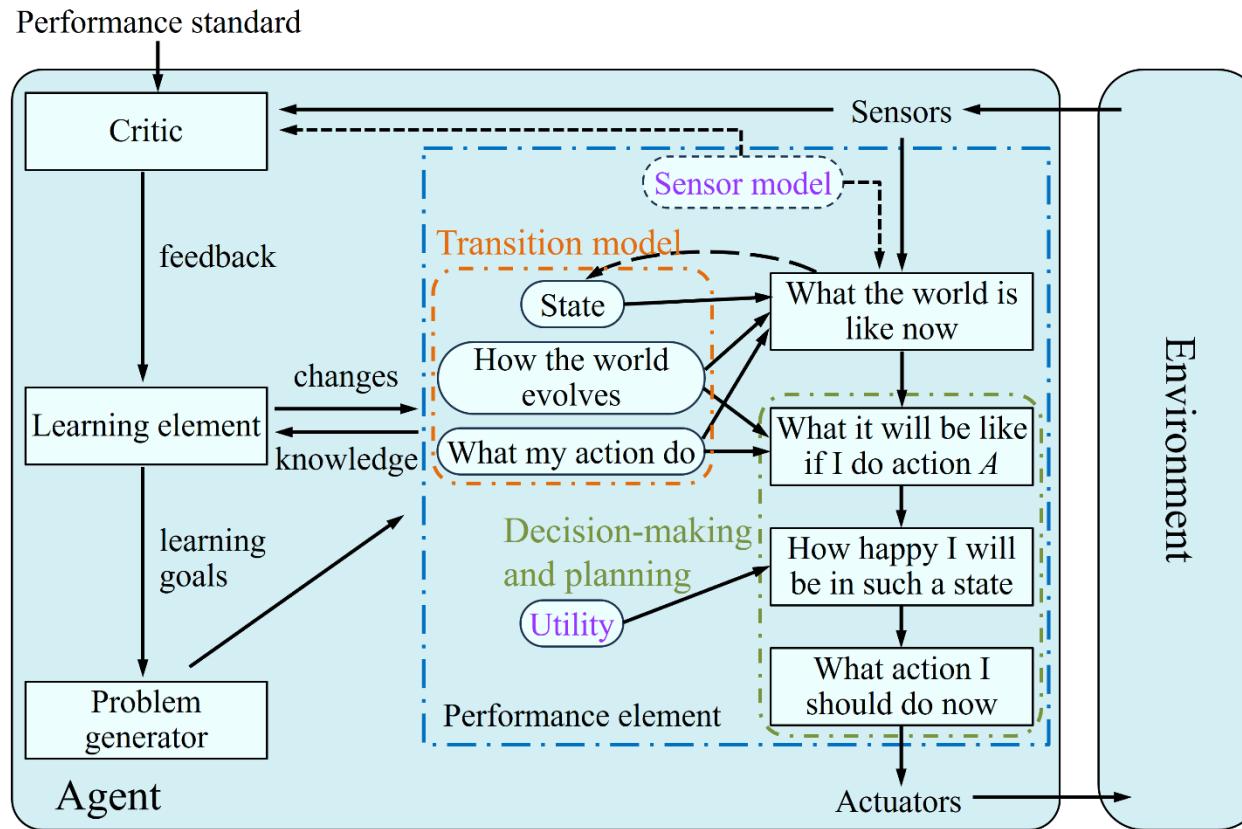
基于效用的Agent要对环境建模并跟踪，任务会涉及很多感知、表示、推理和学习方面的研究。

The Structure of Agents

● Learning agents

已经给出了使用各种方法选择下一步行动的 Agent 程序。还没有说明这些 Agent 程序是如何形成的。构建学习型智能体

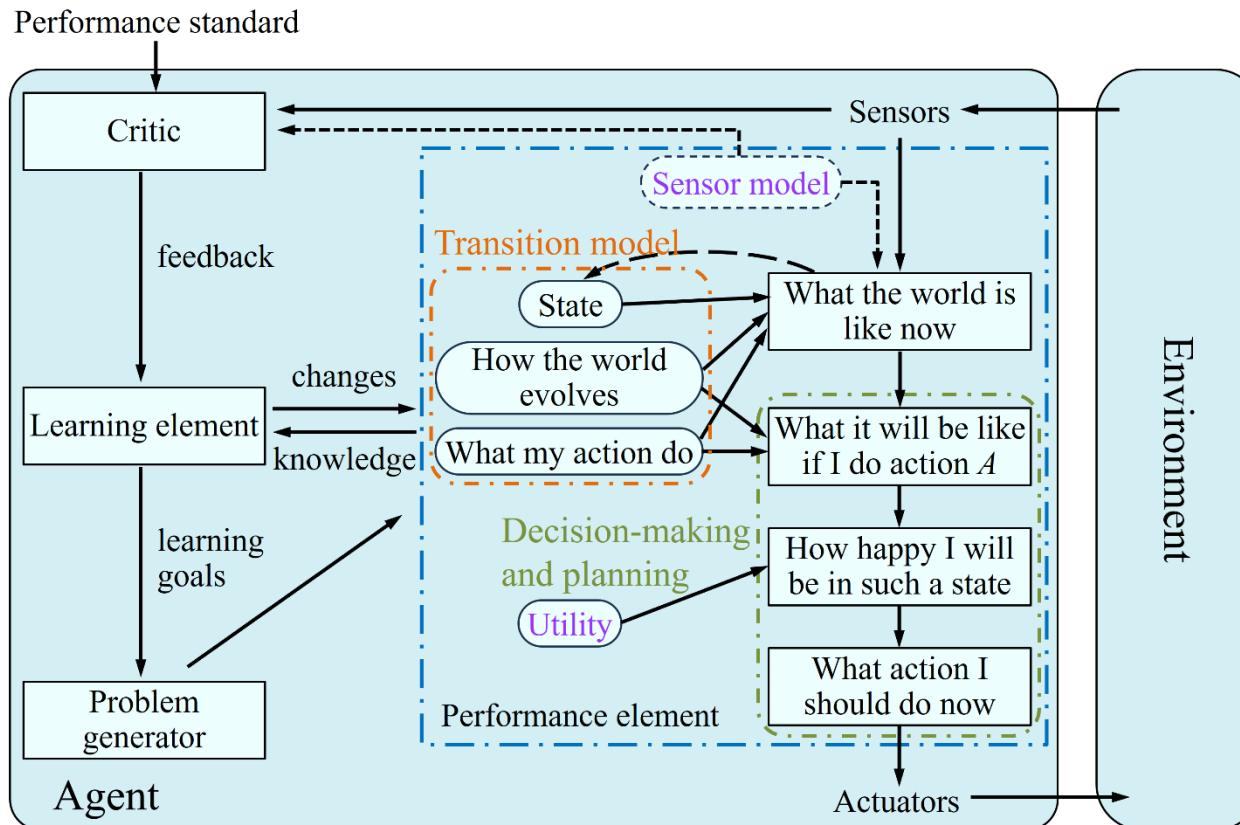
学习元件负责改进提高，而性能元件负责选择外部行动。性能元件是前面考虑的整个 Agent，它接受感知信息并决策。学习元件利用来自评判元件的反馈评价Agent做得如何，并确定应该如何修改执行元件以便将来做得更好。



学习元件的设计很大程度上依赖于性能元件的设计。对于给定的Agent设计，可以构造学习机制来改进Agent的各个部分。

The Structure of Agents

● Learning agents



In a sense, the performance standard distinguishes part of the incoming percept as a reward (or penalty) that provides direct feedback on the quality of the agent's behavior.

评判元件根据固定性能标准告诉学习元件 Agent 的表现如何。

问题产生器负责建议探索性行动，得到新的和有信息的经验。

学习元件可以更改 Agent 结构图中的任何“知识”组件。最简单的情况包括直接从感知序列学习。观察环境的后继状态可以让 Agent 学到“世界是如何发展的”，而观察行动的结果可以让 Agent 学到“我的行动做了什么”。

How the components of agent programs work

◆ Three ways to represent states

组件使用不同方式表示Agent所处环境的区别

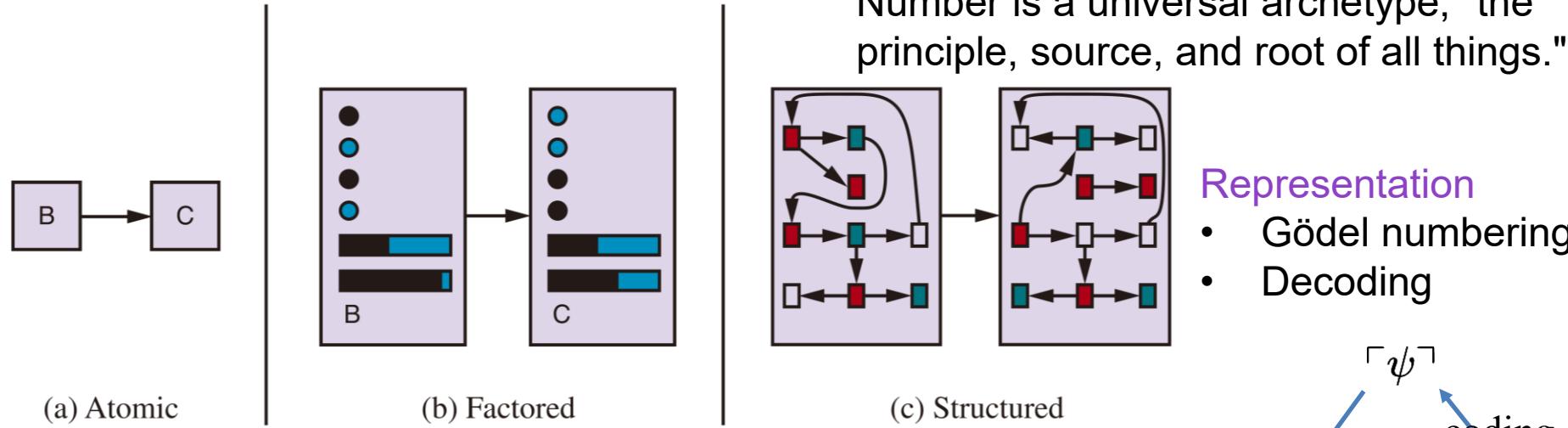


Figure 2.16 Three ways to represent states and the transitions between them. (a) Atomic representation: a state (such as B or C) is a black box with no internal structure; (b) Factored representation: a state consists of a vector of attribute values; values can be Boolean, real-valued, or one of a fixed set of symbols. (c) Structured representation: a state includes objects, each of which may have attributes of its own as well as relationships to other objects.

原子、要素化和结构化表示的轴线是表达能力增长的轴线。大概来说，表达力强的表示可以获取表达力弱的表示的所有信息，至少会更简洁。

随着表达力的增强，推理和学习也越来越复杂。为了找到更具表达力的表示，避免它们的缺点，现实世界的智能系统可能需要同时在轴线上的所有点上操作。

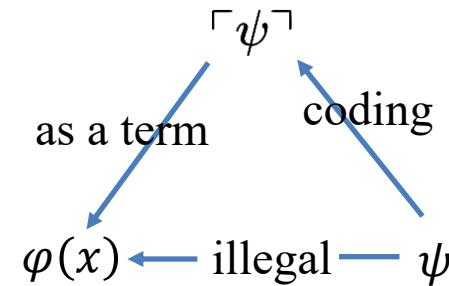
Pythagoras 570BC - 495BC

All things are numbers

Number is a universal archetype, "the principle, source, and root of all things."

Representation

- Gödel numbering
- Decoding



How the components of agent programs work

Another axis for representation involves **the mapping of concepts to locations in physical memory**, whether in a computer or in a brain.

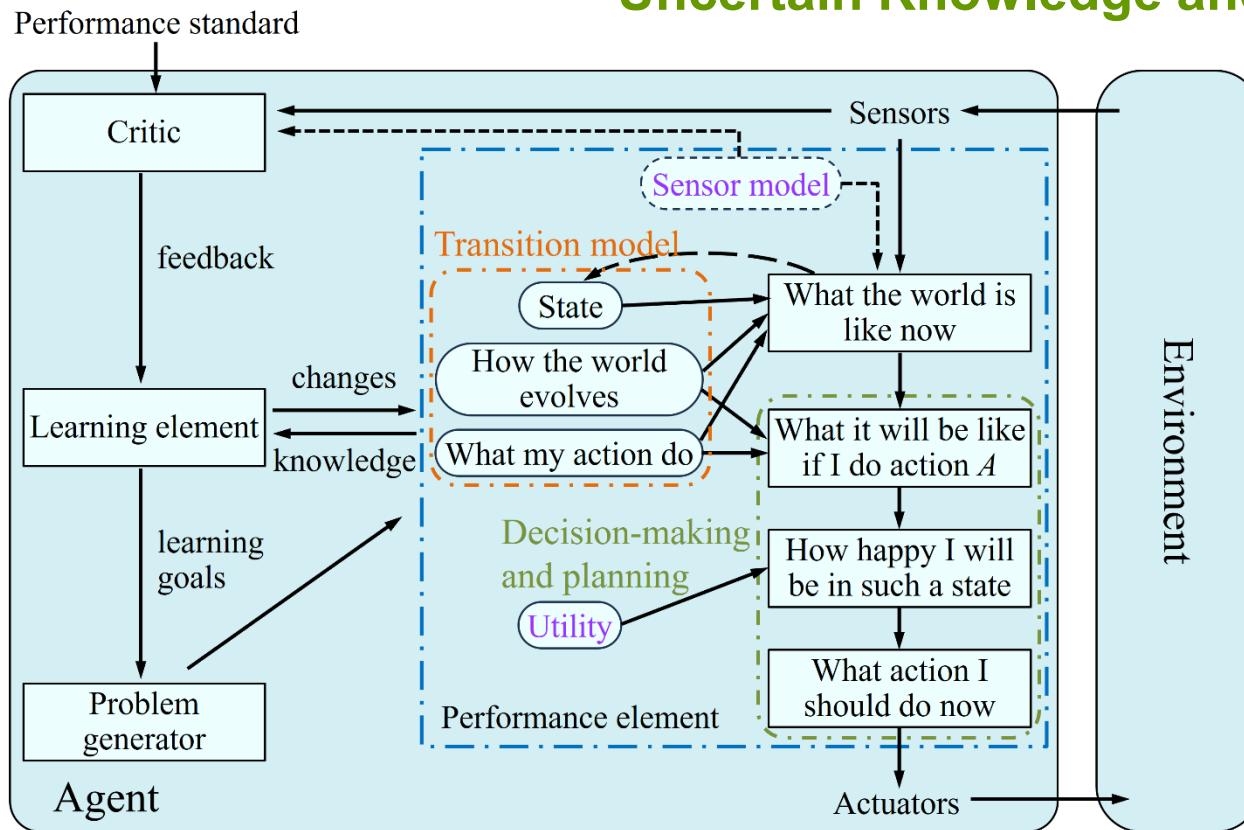
Localist representation: If there is a one-to-one mapping between concepts and memory locations.

Distributed representation: if the representation of a concept is spread over many memory locations, and each memory location is employed as part of the representation of multiple different concepts.

Distributed representations are more robust against noise and information loss.

课程总体框架

Problem-solving
Knowledge, reasoning and planning
Uncertain Knowledge and reasoning



Communicating,
perceive, and acting

Machine Learning

智能计算方法

- 以符号主义为核心的逻辑推理
- 以问题求解为核心的搜索
- 以数据驱动为核心的机器学习
- 以行为主义为核心的强化学习
- 以对抗博弈为核心的决策智能

