



第2讲 体系论

尹慧琳, yinhuilin@tongji.edu.cn

同济大学 电子与信息工程学院



回顾

人工智能的概念

1.1 什么是人工智能

1.2 人工智能的基础

1.3 人工智能的历史

人工智能的发展

1.4 人工智能的判定

1.5 人工智能的层级与流派

1.6 人工智能的最新发展水平

人工智能的研究体系



第 2 讲 体系论

2.1 人工智能的研究路径

2.2 人工智能的主体论点

2.3 智能体的任务环境

2.4 智能体的结构

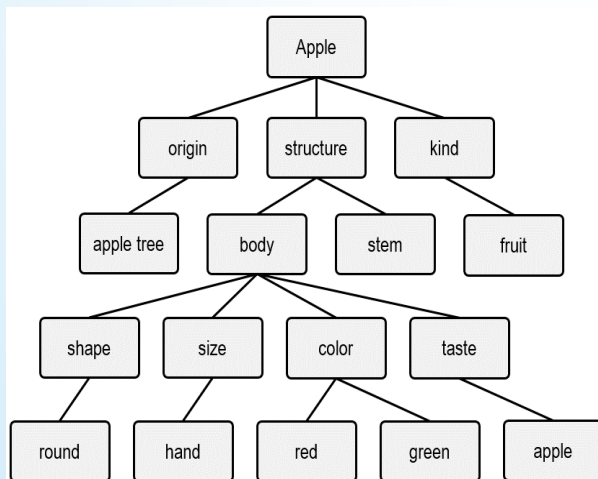
2.5 智能体的类型

第2讲 体系论

2.1 人工智能的研究路径：符号与亚符号

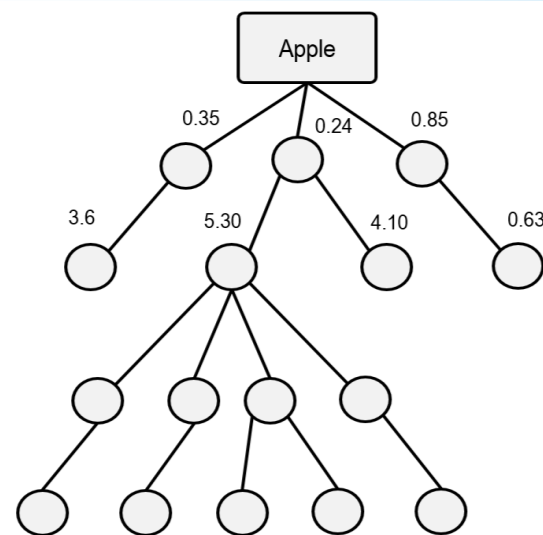
➡ 符号人工智能 (1950~1980)

- 认为：知识可以由**语义上有意义的符号**结构所组成，符号能够表征实体，智能的许多方面能够通过符号操作来获得



➡ 亚符号人工智能 (1980~)

- 认为：**不精确和不确定**是普遍存在的事实；智能行为可以通过**语义上无意义的亚符号**获得



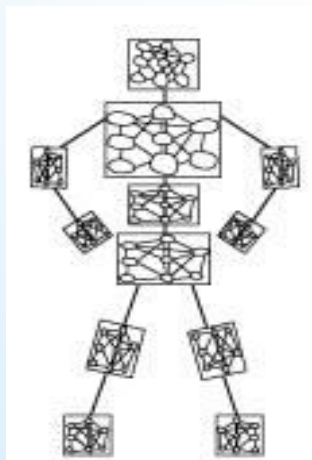


第2讲 体系论

2.1 人工智能的研究路径：整齐与不整齐

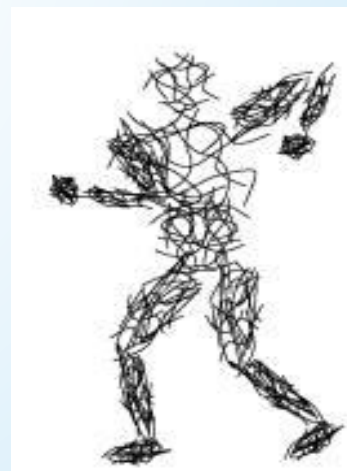
➡ 整齐派

- 认为：人工智能的解决方案应该是优雅的、清晰的、并且可证明是正确的
- 方法：基于逻辑和逻辑的**形式化**扩展



➡ 不整齐派

- 认为：智能太复杂，**无法用整齐方法解决复杂**的人工智能问题
- 方法：采用无定形的主义网络等形式





第2讲 体系论

2.1 人工智能的研究路径：逻辑与反逻辑

➡ 逻辑派 (Stanford)

- 认为：机器无需仿真人类的思维，应该去发现抽象推理和问题求解的本质，而不必关注人类是否使用同样的算法
- 方法：
 - 形式逻辑的方法；
 - 基于逻辑的规则——知识系统
 - 专家系统——在AI应用中构建知识

➡ 反逻辑派 (MIT)

- 认为：没有简单和通用的逻辑能涵盖人工智能的所有方面，故而被称作反逻辑 (Anti-logic)
- 方法：常识知识库 (Commonsense knowledge bases)
 - 必须用手工构建
 - 每次一个复杂的概念



第2讲 体系论

2.1 人工智能的研究路径

对人工智能的争论

流派	理论之争	方法之争
符号主义	认为人的认知基元是 符号 ，认知过程即符号操作过程	认为AI的研究方法应为 功能模拟 方法。力图用数学逻辑方法来建立人工智能的统一理论体系。
连接主义	认为人的思维基元是 神经元 。提出连接主义的大脑工作模式，用于取代符号操作的计算机工作模式。	主张AI应着重于 结构模拟 ，即模拟人的生理神经网络结构，并认为功能、结构和智能行为是密切相关的。
行为主义	认为 智能取决于感知和行动 ，提出智能行为的感知-动作模式。	认为AI的研究方法应采用 行为模拟 方法，也认为功能、结构和智能行为是不可分的。



第 2 讲 体系论

2.1 人工智能的研究路径

2.2 人工智能的主体论点

2.3 智能体的任务环境

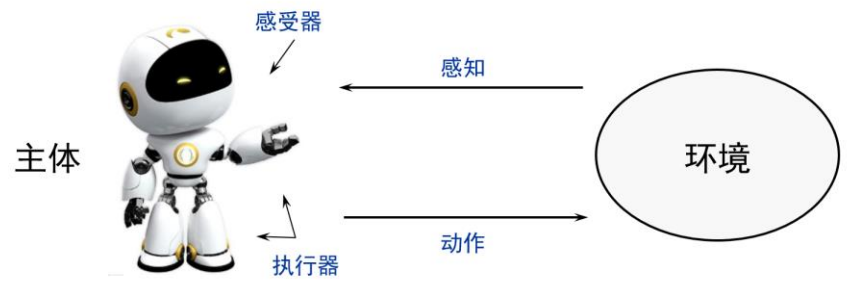
2.4 智能体的结构

2.5 智能体的类型

第2讲 体系论

2.2 人工智能的主体论点

- **主体 (Agent)**：通过**感受器**感知其环境并通过**执行器**对该环境起作用的任何事物。
- **智能主体**：能够感受外部环境，经过一定的**思维**，再作用于外部环境的事物。



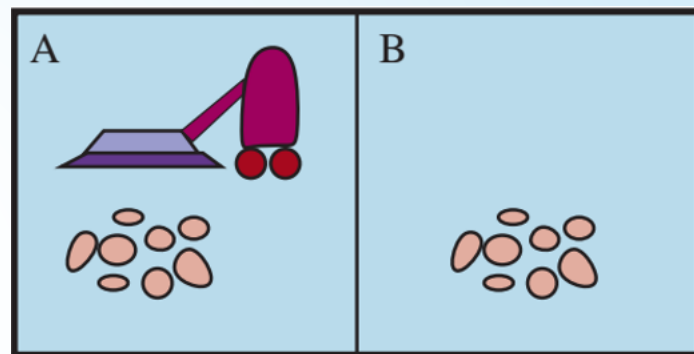
	类人地	理性地
思考	像人一样思考	合理地思考
行动	像人一样行动	合理地行动

第2讲 体系论

2.2 人工智能的主体论点

举例：机器人吸尘器

- 环境：
 - 两个地点，方形区域A和B
 - 每个地点可能有灰，可能无灰
- 感知：
 - 位置：A或B
 - 内容：有灰或无灰
- 动作：
 - 左行 Right
 - 右行 Left
 - 吸尘 Suck
 - ...



简单吸尘器智能体功能的部分列表

感 知	动 作
[A, Clean]	Right
[A, Dirty]	Suck
[B, Clean]	Left
[B, Dirty]	Suck
.....



第2讲 体系论

2.2 人工智能的主体论点

智能体Agent举例：

人

- 感受器：眼、耳、其它感官
- 执行器：手、脚、声道、等等

机器人

- 感受器：摄像头、红外测距仪、其它检测器
- 执行器：各种马达

软件

- 感受器：击键、鼠标、文件内容、网络包等
- 执行器：屏幕显示、写文



第 2 讲 体系论

2.1 人工智能的研究路径

2.2 人工智能的主体论点

2.3 智能体的任务环境

2.4 智能体的结构

2.5 智能体的类型



第2讲 体系论

2.3 智能体的任务环境

任务环境的规范描述：包括性能度量、环境、智能体的执行器和传感器。

➡ PEAS描述

- ➡ 性能 (**P**erformance)
- ➡ 环境 (**E**nvironment)
- ➡ 执行器 (**A**ctuators)
- ➡ 传感器 (**S**ensors)

智能体	性能	环境	执行器	传感器
出租司机	安全、快速、守法、舒适、收益	道路、交通、行人、顾客	方向盘、油门、刹车器、信号、喇叭	摄像头、声纳、速度仪、GPS、里程表、加速度计、引擎传感器、操作盘



第2讲 体系论

2.3 智能体的任务环境

➡ 完全可观测与部分可观测

- ➡ 若智能体的传感器在每个时间点上可访问环境的完整（部分）状态，则该任务环境是完全（部分）可观测。

➡ 单智能体与多智能体

- ➡ 若是一个智能体在一个环境内自运行，则为一个单智能体；如果考虑环境中的其他智能体（合作或竞争），则为多智能体。

➡ 确定的与随机的

- ➡ 环境的下一个状态完全由当前的状态和由该智能体执行的动作所决定，则该环境是确定的；否则，它是随机的。



第2讲 体系论

2.3 智能体的任务环境

► 片段式（阵发性）与延续式（连续性）

- 片段式任务环境：智能体的动作过程被分为原子的片段，并且每个片段的动作选择仅仅依赖于片段本身；
- 延续式任务环境：当前的决策会影响到所有未来的决策。

► 静态的与动态的

- 如果环境随智能体的动作而改变，则该智能体的环境是动态的；否则是静态的；
- 如果环境本身不随时间的推移而改变，但该智能体的性能发生变化，则为半动态的。



第2讲 体系论

2.3 智能体的任务环境

➡ 离散型与连续型

- ➡ 离散与连续的区别在于环境的状态、时间处理的方式、智能体的感知信息、智能体的动作多方面。

➡ 已知的与未知的

- ➡ 在一个已知的环境下，所有动作的结果是给定的。
- ➡ 如果环境是未知的，则该智能体将需要学习如何动作，以便做出正确的决策。



第 2 讲 体系论

2.1 人工智能的研究路径

2.2 人工智能的主体论点

2.3 智能体的任务环境

2.4 智能体的结构

2.5 智能体的类型



第2讲 体系论

2.4 智能体的结构

智能体的主体结构

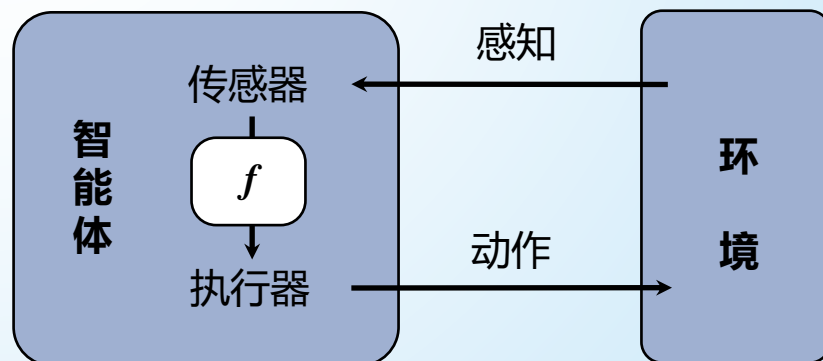
- **硬件平台**: 具备物理传感器和执行器的计算装置, 能够运行Agent程序
- **主体函数**: 控制智能体实现其功能的主体算法

主体函数

- 将感受器的输入作为当前的感知, 然后返回一个动作给执行器。
- 数学上被描述为一个映射 $f: P^* \rightarrow A$ where $P^* = \sum_{t=1}^T |P|^t$

主体函数的方法:

- 单个选项的效用计算
- 基于逻辑规则的推论
- 神经网络
- 查找表
-

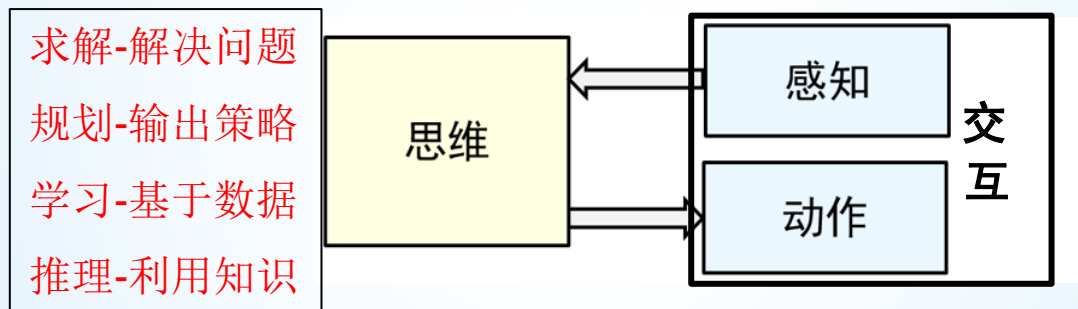


第2讲 体系论

2.4 智能体的结构

➤ 人工智能的内涵：机器能否思维？

- 思维：主体将感知的信息加以处理的过程。
- 交互：主体从环境得到感知并作用于外部环境的过程。



➤ 人工智能的外延：智能科学和技术的应用

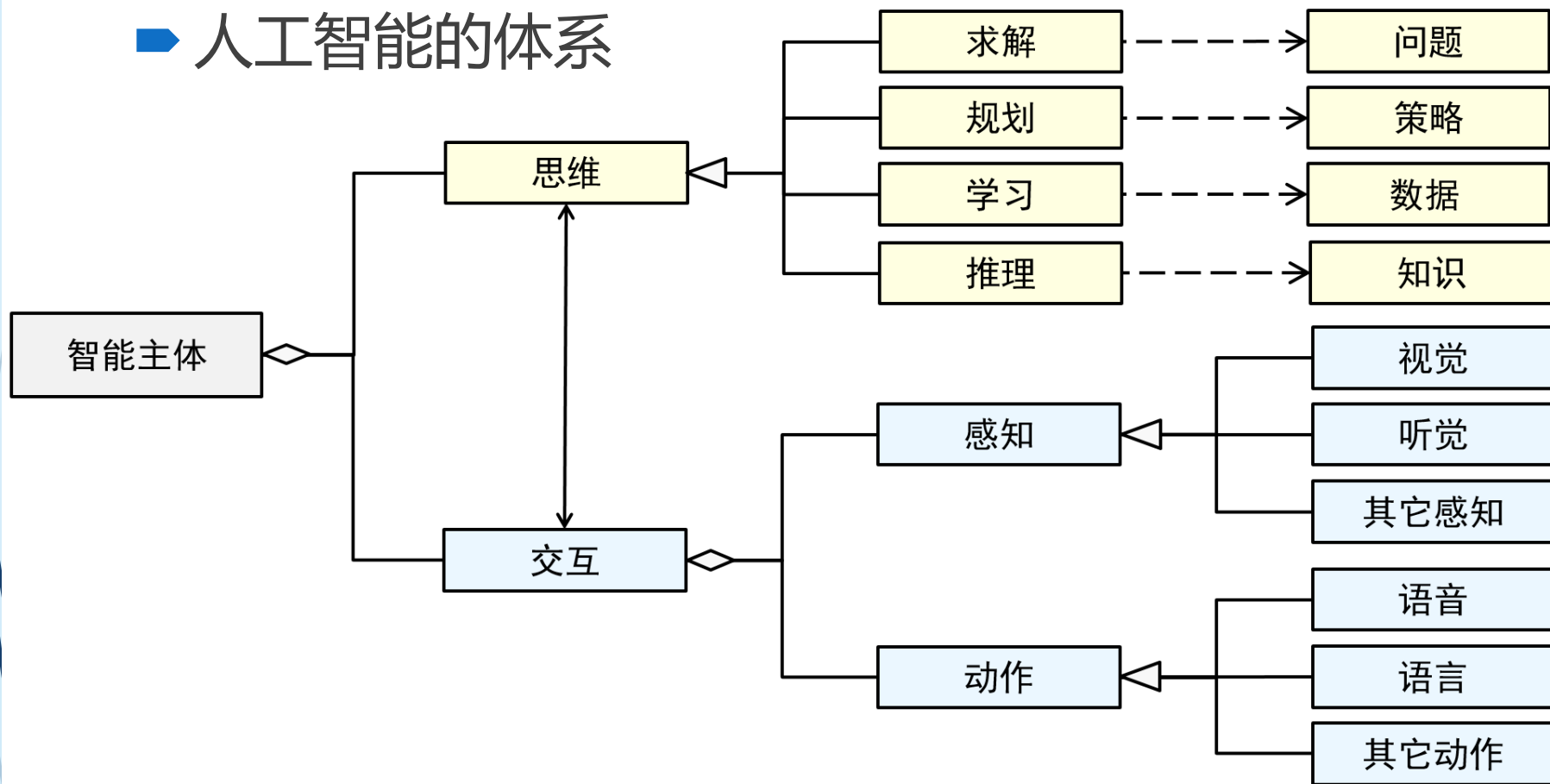
- 工业机器人、农业机器人、服务机器人等各类机器人
- 智能交通
- 智能制造
- 智能驾驶和无人机
- 智慧医疗、
- 智慧城市
-



第2讲 体系论

2.4 智能体的结构

► 人工智能的体系





第 2 讲 体系论

2.1 人工智能的研究路径

2.2 人工智能的主体论点

2.3 智能体的任务环境

2.4 智能体的结构

2.5 智能体的类型



第2讲 体系论

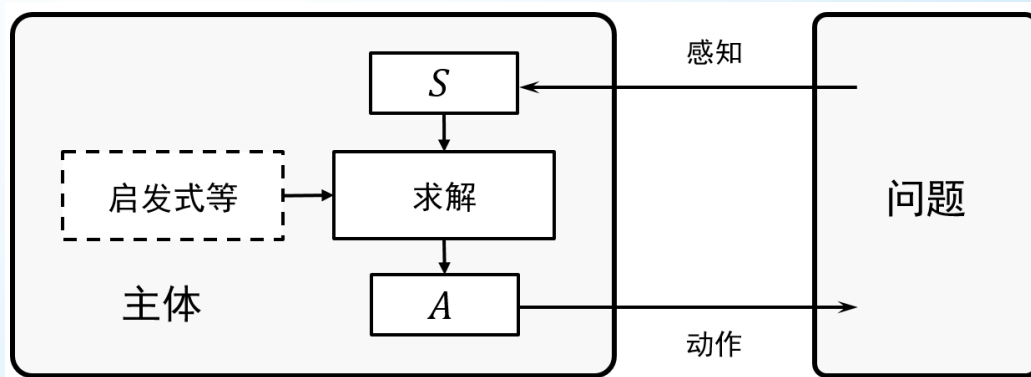
2.5 智能体的类型

■ 求解型主体：

- 环境是要求解的**问题**
- 感知的是问题的状态
- 求解后生成相应的动作A

■ 代表性的求解问题：

- 搜索问题
- 优化问题
- 博弈问题
- 约束问题





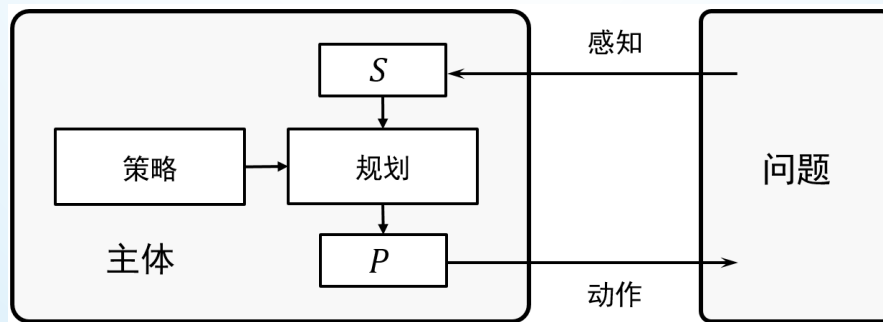
第2讲 体系论

2.5 智能体的类型

► 规划型主体：

针对问题的初始状态 S 、预期的目标状态、以及可能的动作，制定从初始状态到达目标状态的**动作计划 P** 。

- 规划依赖于策略，包括**时空关联规划**（时间、空间、调度策略），以及**决策理论规划**（基于马尔科夫决策过程）。





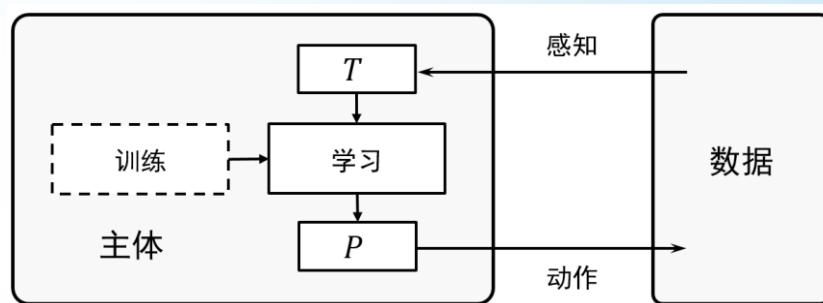
第2讲 体系论

2.5 智能体的类型

► 学习型主体：

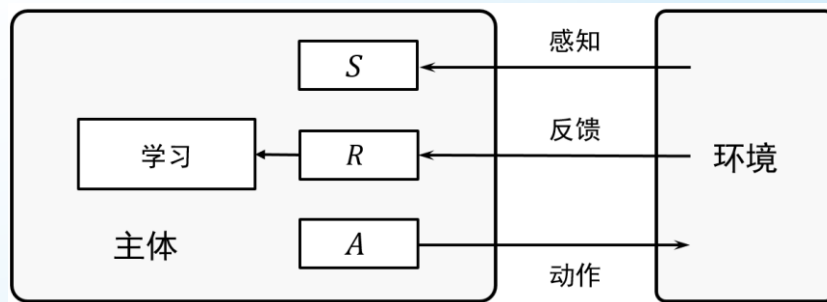
► 基于数据的学习

- 监督学习：有已标注训练数据
- 无监督学习：无已标注训练数据



► 基于环境的学习

- 强化学习
- 根据当前的状态 S 、和环境的回报值 R (reward) 来决定下一步的动作





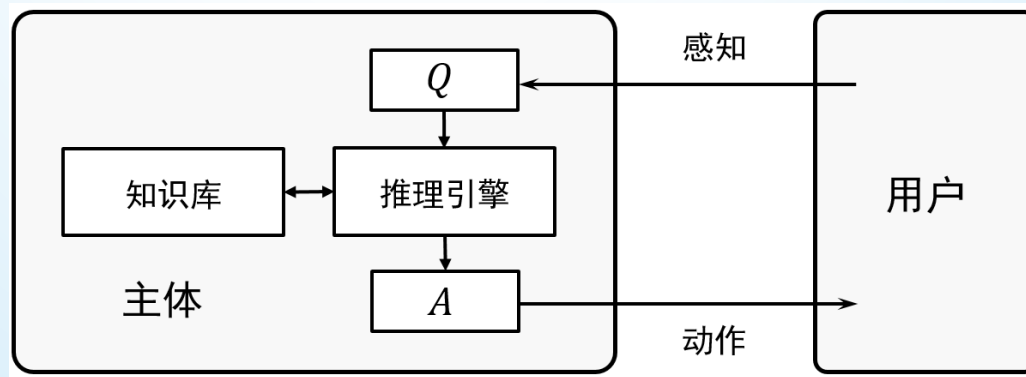
第2讲 体系论

2.5 智能体的类型

推理型主体

知识库 (Knowledge base) 与推理引擎 (Inference engine)

- 知识基于知识表示 (Knowledge representation) , 包括: 语义网络、框架、逻辑、本体、知识图谱、贝叶斯网络等。
- 推理引擎基于推理机制 (Inference mechanism) , 包括: 逻辑推理、定性推理、语义推理、概率推理、以及机器推理。





作业:

描述人工智能的研究体系

