

人工智能技术

Artificial Intelligence

——人工智能: 经典智能+计算智能+机器学习

AI: Classical Intelligence + Computing Intelligence + Machine Learning

王鸿鹏、杜月、王润花、许丽

南开大学人工智能学院



概论

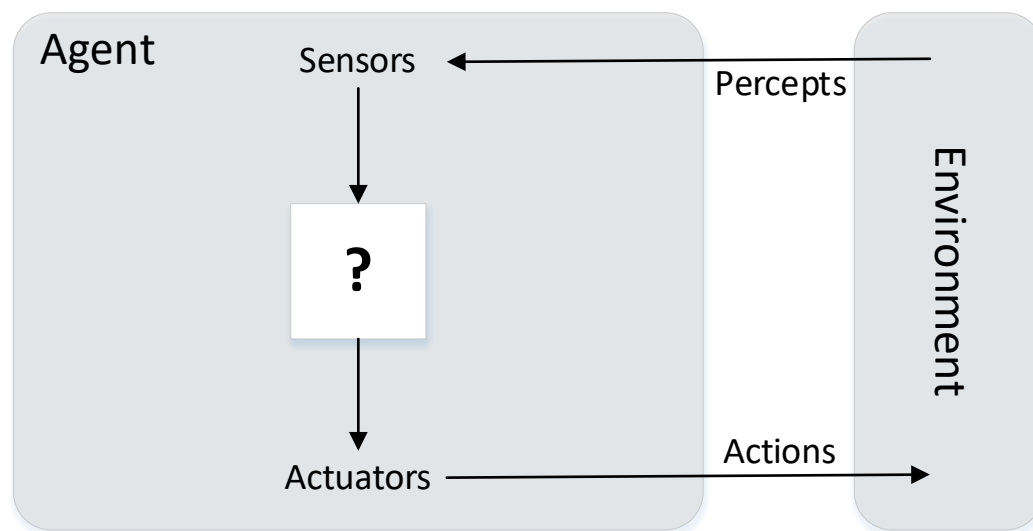
Introduction

——智能化智能体

Chapter 2: Intelligent Agent

Agents

- Agent（智能体）：具有智能的实体，是驻留在某一环境下，能持续自主地发挥作用，具备驻留性、反应性、社会性、主动性等特征的计算实体。
- 智能体是人工智能领域中一个很重要的概念。任何独立的能够思想并可以同环境交互的实体都可以抽象为智能体。



图：Agent通过传感器和执行器与环境进行交互

Agent的特性

- **自治性**(Autonomy):智能体能根据外界环境的变化,而自动地对自己的行为和状态进行调整,而不是仅仅被动地接受外界的刺激,具有自我管理自我调节的能力。
- **反应性**(Reactive):能对外界的刺激作出反应的能力。
- **主动性**(Proactive):对于外界环境的改变,智能体能主动采取活动的能力。
- **社会性**(Social):智能体具有与其它智能体或人进行合作的能力,不同的智能体可根据各自的意图与其它智能体进行交互,以达到解决问题的目的。
- **进化性**:智能体能积累或学习经验和知识,并修改自己的行为以适应新环境。

Agent概念的提出

- 现在IT界的智能体概念是由麻省理工学院的著名计算机学家和人工智能学科创始人之一的Minsky提出来的，他在“Society of Mind”一书中将社会与社会行为概念引入计算系统。
- 传统的计算系统是封闭的，要满足一致性的要求，然而社会机制是开放的，不能满足一致性条件，这种机制下的部分个体在矛盾的情况下，需要通过某种协商机制达成一个可接受的解。Minsky将计算社会中的这种个体称为智能体。这些个体的有机组合则构成计算社会——多智能体系统。
- Simon的有限性理论是多智能体系统形成的另一个重要的理论基础，Simon认为一个大的结构把许多个体组织起来可以弥补个体工作能力的有限；每个个体负责一项专门的任务可以弥补个体学习新任务的能力的有限；社会机构间有组织的信息流动可以弥补个体知识的有限；精确的社会机构和明确的个体任务可以弥补个体处理信息和应用信息的能力的有限。

Intelligent Agents

- Intelligent Agent 是这样一种智能体，给定它所感知到的和它拥有的先验知识，以一种被期望最大化其性能指标的方式运行（性能指标定义了智能体成功的标准）。此类智能体也成为理性智能体(Rational Agent)。智能体的合理性是通过其性能指标，其拥有的先验知识，它可以感知的环境及其可执行的操作来衡量的。这个概念是人工智能的核心。
- **PEAS**: Performance, Environment, Actuators & Sensors.

以无人驾驶汽车为例：

- 性能**P**：安全性、时间、合法驾驶、舒适性。
- 环境**E**：道路、其他汽车、行人、路标。
- 执行器**A**：转向、加速器、制动器、信号、喇叭。
- 传感器**S**：相机、声纳、GPS、速度计、里程计、加速度计、发动机传感器、键盘。

好的行为表现：理性的概念

- **理性智能体**是做事正确的智能体：对的行动就是使得智能体更加成功的行动。

- **性能度量**：

我们需要某种方法来度量一个智能体的成功与否。连同对智能体所处的环境、智能体的传感器和执行器的描述，这将为智能体面临的任务提供一个完整的规范说明。

性能度量是智能体成功程度标准的具体化。

- **理性**：

对于每个可能的感知序列，根据已知的感知序列提供的证据和智能体内建的先验知识、理性智能体应该选择期望能使其性能度量最大化的行动。

环境类型

- **完全可观察和部分可观察**：如果是完全可观察的，智能体的传感器可以在每个时间点访问环境的完整状态，否则不能。例如，国际象棋是一个完全可观察的环境，而扑克则不是。
- **确定性和随机性**：环境的下一个状态完全由当前状态和由智能体接下来所执行的操作决定的。（如果环境是确定性的，而其他智能体的行为不确定，那么环境是随机性的）。随机环境在本质上是随机的，不能完全确定。例如，8数码难题（8-puzzle）这个在线拼图游戏有一个确定性的环境，但无人驾驶的汽车没有。
- **静态和动态**：当智能体在进行协商（deliberate）时，静态环境没有任何变化。（环境是半动态的，环境本身并没有随着时间的流逝而变化，但智能体的性能得分则是会发生相应变化的）。另一方面，动态环境却改变了。西洋双陆棋具有静态环境，而扫地机器人roomba具有动态环境。
- **离散和连续**：有限数量的明确定义的感知和行为，构成了一个离散的环境。例如，跳棋就是离散环境的一个范例，而自动驾驶汽车则需要在连续环境下运行。
- **单一智能体和多智能体**：仅有自身操作的智能体本身就有单一智能体环境。但是如果还有其他智能体包含在内，那么它就是一个多智能体环境。自动驾驶汽车就具有多智能体环境。

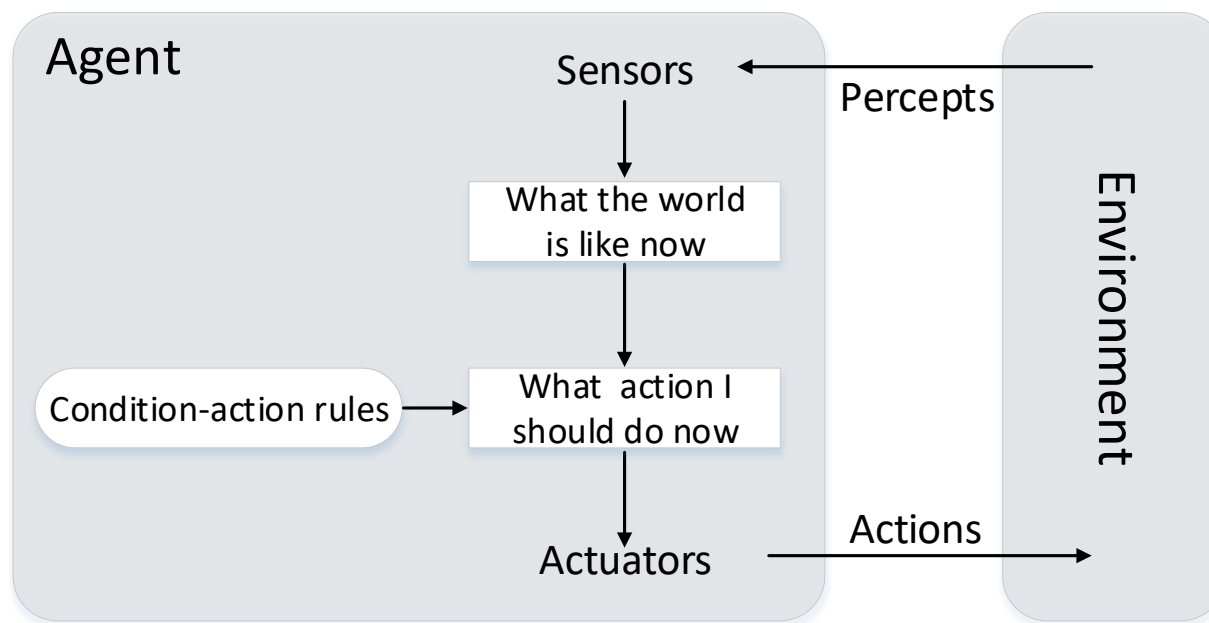
环境类型

任务环境 Task Environment	可观察的 Observable	智能体 Agents	确定性 Deterministic	片段/延续式 Episodic	静态/动态 Static	离散/连续式 Discrete
填字游戏 Crossword puzzle 计时国际象棋 Chess with a clock	Fully Fully	Single Multi	Deterministic Deterministic	Sequential Sequential	Static Semi	Discrete Discrete
扑克 Poker 双陆棋 Backgammon	Partially Fully	Multi Multi	Stochastic Stochastic	Sequential Sequential	Static Static	Discrete Discrete
出租车 Taxi driving 医疗诊断 Medical diagnosis	Partially Partially	Multi Single	Stochastic Stochastic	Sequential Sequential	Dynamic Dynamic	Continuous Continuous
图像分析 Image analysis 分拣机器人 Part-picking robot	Fully Partially	Single Single	Deterministic Stochastic	Episodic Episodic	Semi Dynamic	Continuous Continuous
炼油厂控制器 Refinery controller 互动英语导师 Interactive English tutor	Partially Partially	Single Multi	Stochastic Stochastic	Sequential Sequential	Dynamic Dynamic	Continuous Discrete

智能体的类型-1

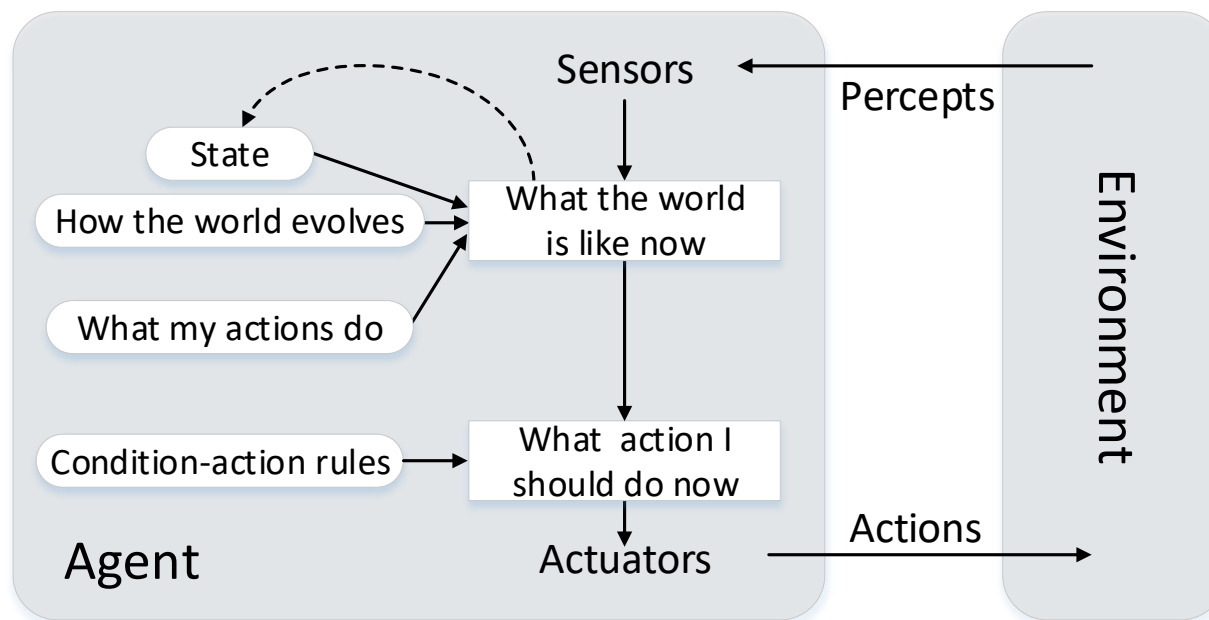
- 一般有4种类型的智能体，根据智能水平或其能够执行任务的复杂性不同而区分。
- 所有类型都可以随着时间的推移改进性能并产生更好的操作。这些可以概括为学习智能体 (learning agents) 。

单反射性智能体 (Simple reflex agents)



智能体的类型-2

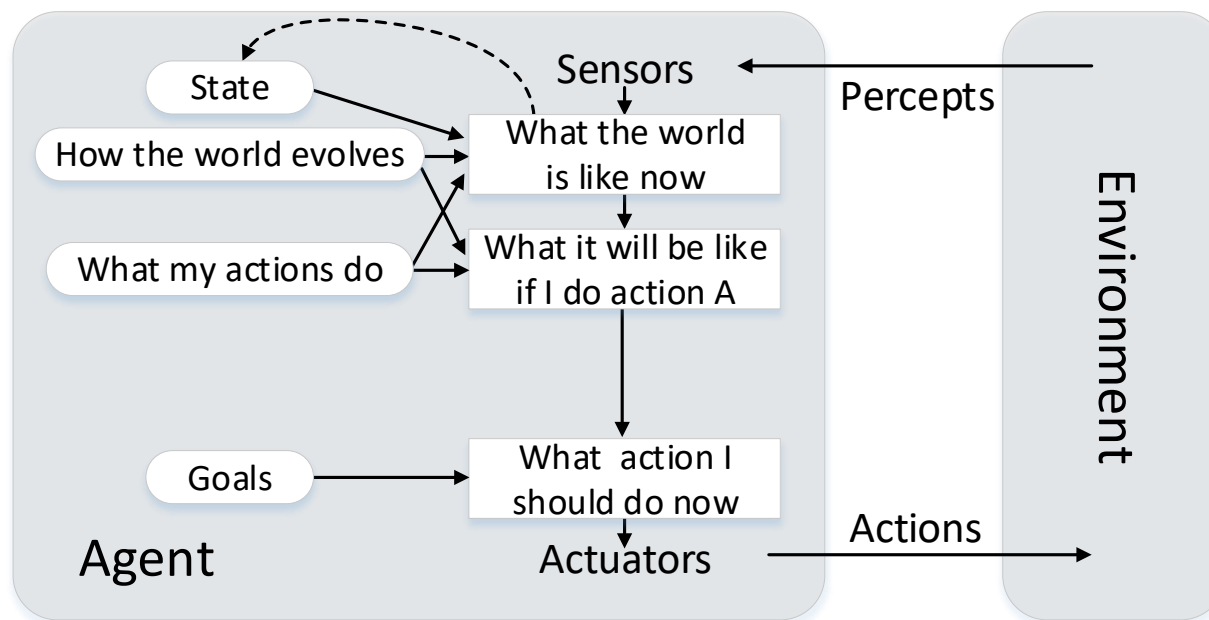
基于模型的反射性智能体(Model-based reflex agents)



智能体跟踪部分可观察的环境。这些内部状态取决于感知历史。环境/世界的建模是基于它如何从智能体中独立演化，以及智能体行为如何影响世界。

智能体的类型-3

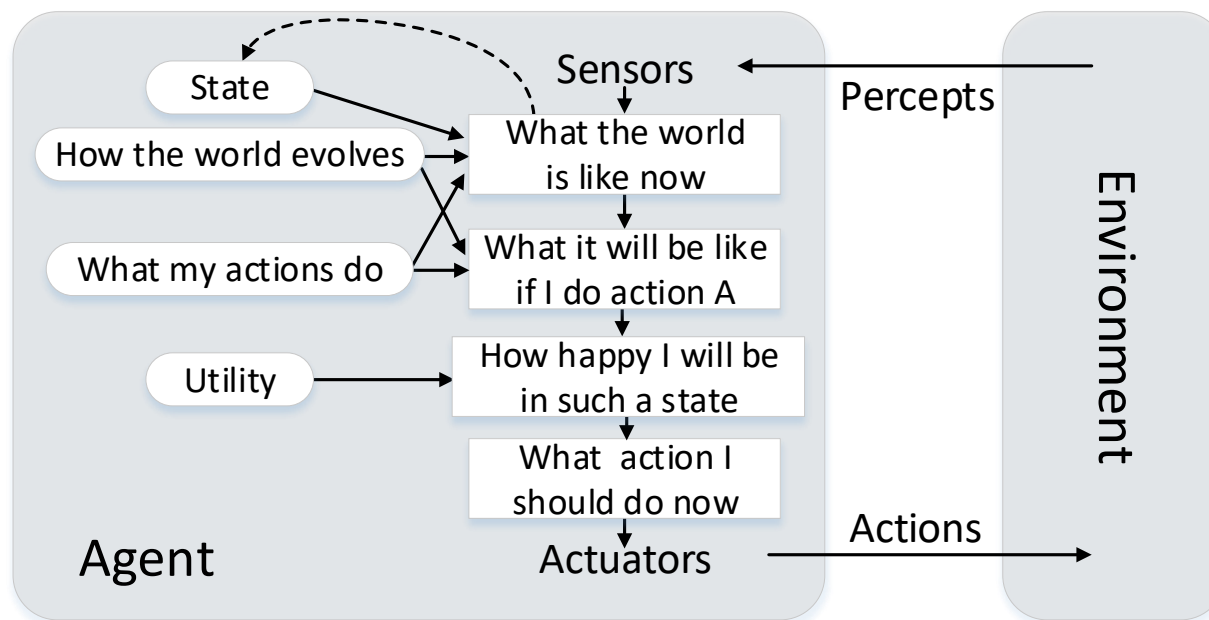
基于目标的智能体 (Goal-based agents)



这是对基于模型的智能体的改进，并且在知道当前环境状态不足的情况下使用。智能体将提供的目标信息与环境模型相结合，选择实现该目标的行动。

智能体的类型-4

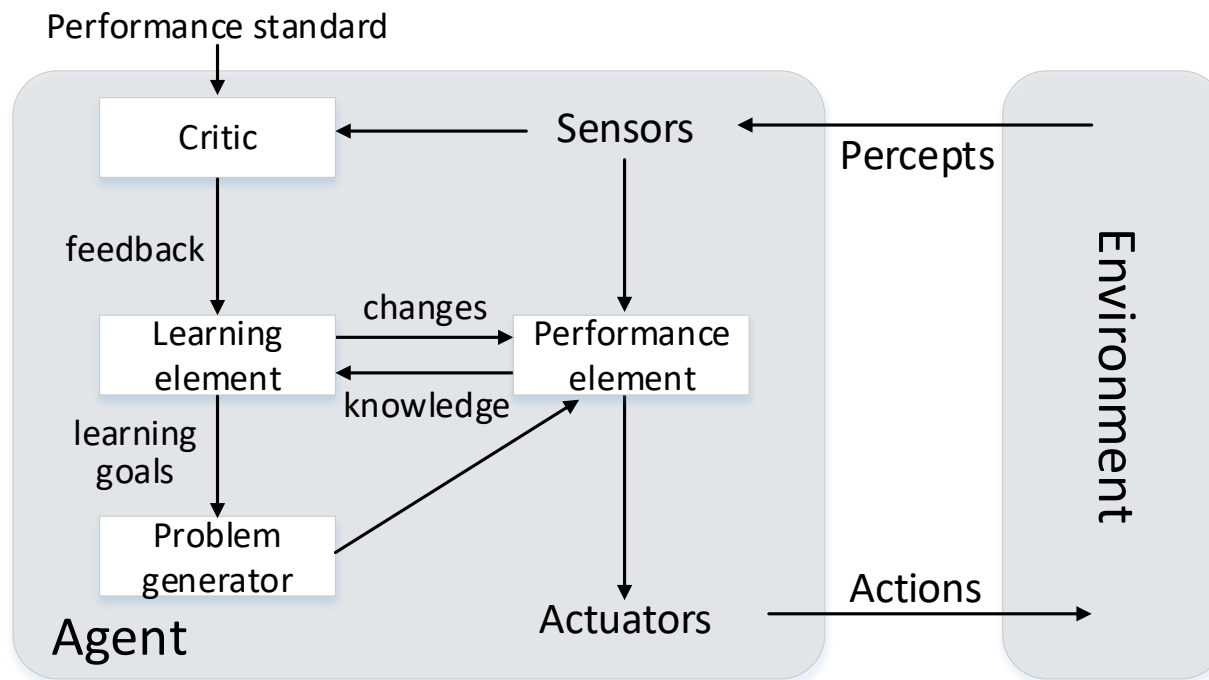
基于效用的智能体 (Utility-based agents)



对基于目标的智能体进行改进，在实现预期目标方面有所帮助是不够的。我们可能需要考虑成本。例如，我们可能会寻找更快、更安全、更便宜的旅程到达目的地。这由一个效用函数标记。效用智能体将选择使期望效用最大化的操作。

智能体的类型-5

学习智能体 (Learning Agents)



学习要素：负责改进

性能要素：负责选择外部行为，这是截至目前我们通常认为的智能体。

评论：关于确定的性能标准，智能体做得如何？

问题生成器：允许智能体探索。

多智能体(Multi-Agent System)

- 多智能体系统是多个智能体组成的集合，它的目标是将大而复杂的系统建设成小的、彼此互相通信和协调的，易于管理的系统。
- 多智能体系统是分布式人工智能（DAI, Distributed Artificial Intelligence）的一个重要分支，是20世纪末至21世纪初国际上人工智能的前沿学科。研究的目的在于解决大型、复杂的现实问题，而解决这类问题已超出了单个智能体的能力。
- 多智能体系统研究领域，主要包括：多智能体规划、学习、推理、协商、交互机制等等理论，及其实际应用。

多智能体(Multi-Agent System)

- MAS的研究涉及智能体的知识、目标、技能、规划以及如何使智能体采取协调行动解决问题等。研究者主要研究智能体之间的交互通信、协调合作、冲突消解等方面，强调多个智能体之间的紧密群体合作，而非个体能力的自治和发挥，主要说明如何分析、设计和集成多个智能体构成相互协作的系统。
- 同时，人们也意识到，人类智能的本质是一种社会性智能，人类绝大部分活动都涉及多个人构成的社会团体，大型复杂问题的求解需要多个专业人员或组织协调完成。要对社会性的智能进行研究，构成社会的基本构件物——人的对应物——智能体理所当然成为人工智能研究的基本对象，而社会的对应物——多智能体系统，也成为人工智能研究的基本对象，从而促进了对多智能体系统的行为理论、体系结构和通信语言的深入研究，这极大的繁荣了智能体技术的研究与开发。

多智能体(Multi-Agent System)的优势、特点

- 多智能体系统在表达实际系统时，通过各智能体间的通讯、合作、互解、协调、调度、管理及控制来表达系统的结构、功能及行为特性。
- 多智能体系统具有自主性、分布性、协调性，并具有自组织能力、学习能力和推理能力。采用多智能体系统解决实际问题，具有很强的鲁棒性和可靠性，并具有较高的问题求解效率。
- 在多智能体系统中，每个智能体具有独立性和自主性，能够解决给定的子问题，自主地推理和规划并选择适当的策略，并以特定的方式影响环境。
- 多智能体系统支持分布式应用，所以具有良好的模块性、易于扩展性和设计灵活简单，克服了建设一个庞大的系统所造成的管理和扩展的困难，能有效降低系统的总成本；
- 在多智能体系统的实现过程中，不追求单个庞大复杂的体系，而是按面向对象的方法构造多层次，多元化的智能体，其结果降低了系统的复杂性，也降低了各个智能体问题求解的复杂性；

多智能体(Multi-Agent System)的应用领域

- 在智能机器人中，信息集成和协调是一项关键性技术，它直接关系到机器人的性能和智能化程度。一个智能机器人应包括多种信息处理子系统，如二维或三维视觉处理、信息融合、规划决策以及自动驾驶等。各子系统是相互依赖、互为条件的，它们需要共享信息、相互协调，才能有效地完成总体任务，其目标是用来结合、协调、集成智能机器人系统的各种关键技术及功能子系统，使之成为一个整体以执行各种自主任务。利用多智能体系统，将每个机器人作为一个智能体，建立多智能体机器人协调系统，可实现多个机器人的相互协调与合作，完成复杂的并行作业任务。

Q&A

THANKS!

