

# COMP 9414 WEEK 1 - Artificial Intelligence and Agents

Author: Junji Duan

Last Update: 2025/09/20

## 1 W1A: Foundations

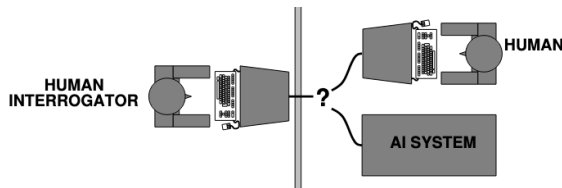
### 1.1 What is Artificial Intelligence

**Def:** Artificial intelligence, or AI, is the field that studies the synthesis and analysis of computational agents that act intelligently. Consider each part of this definition.

**Def:** An agent is something that acts in an environment; it does something.

**Def:** Intelligence is a matter of degree.

#### 1.1.1 Turing Test



Aim is for the computer to imitate(模拟) a human well enough to fool a human interrogator

### 1.2 Physical Symbol System Hypothesis

**Def:** A symbol is meaningful physical pattern that can be manipulated

**Def:** A symbol system combines and manipulates symbols

### 1.3 Knowledge Level and Symbol Level

**Def:** The knowledge level is in terms of what an agent knows and what its goals are (external theory of the agent)

**Def:** The symbol level is a level of description in terms of what reasoning the agent is doing (internal description of the agent's operation)

## 2 W1B: Agent

### 2.1 Agents

**Def: Agent:** An entity that perceives its environment through sensors and acts on its environment through effectors.

- Agent as actor; Agent as helper, representative
- An entity(实体)
  - situated(环境嵌入性): operates in a dynamically changing environment (运行在一个动态变化的环境, 实时感知)
  - reactive(反应性): responds to changes in a timely manner (及时对环境的变化做出反应)
  - autonomous(自主性): can control its own behaviour (自己决定如何行动, 而不是完全依赖外部指令)
  - proactive(主动性/前瞻性): exhibits goal-oriented behaviour (不仅仅是被动反应, 还能主动追求目标)
  - communicating(沟通性): coordinate with other agents (多个 Agent 可以相互交流、协调合作)
- Specifying Agents
  - percepts(感知): inputs to the agent via sensors
  - actions(行动): outputs available to the agent via effectors
  - goals(目标): objectives or performance measure of the agent
  - environment(环境): world in which the agent operates

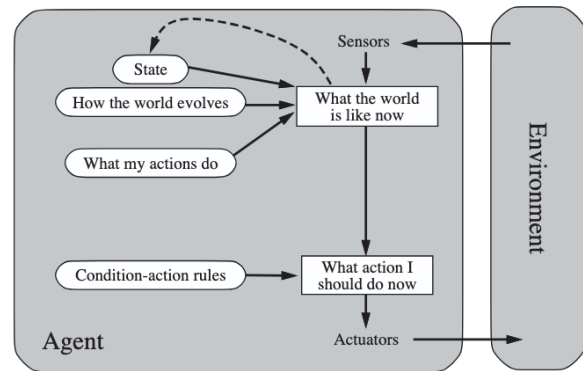
An agent is an entity that can autonomously perceive, decide, and act within an environment. It can respond reactively to changes or proactively pursue its goals. Ideally, an agent is rational, but in practice, agents often have to make approximately optimal decisions under limited resources.

## 2.2 Agent Architectures and Programs

Agent = Architecture + Program

Architecture –hardware and software components, and their organization, on which agent program executes

Program –function implementing mapping from percept sequence to actions, using an internal representation of the percept history



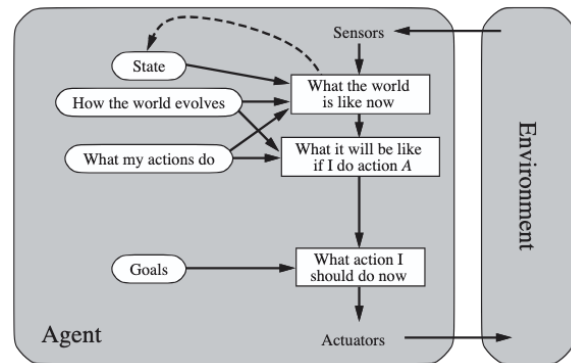
### 2.2.1 Representation

In Artificial Intelligence, representation refers to how we formally describe the world or a problem so that an agent can reason and make decisions.

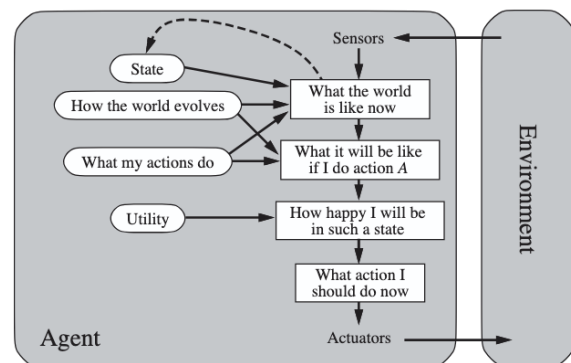
What makes a good representation?

- Rich enough: It must capture all the knowledge needed to solve the problem.
- Compact, natural, and maintainable: As close to the problem as possible, not unnecessarily complicated.
- Efficient computation:
  - Should highlight exploitable features for computational gain.
  - Should allow tradeoffs between accuracy and time/space resources.
- Acquirable: Should be learnable from people, data, or past experiences.

Goal-based (teleological) agent(基于目标) –state description often not sufficient for agent to decide what to do so it needs to consider its goals (may involve searching and planning)

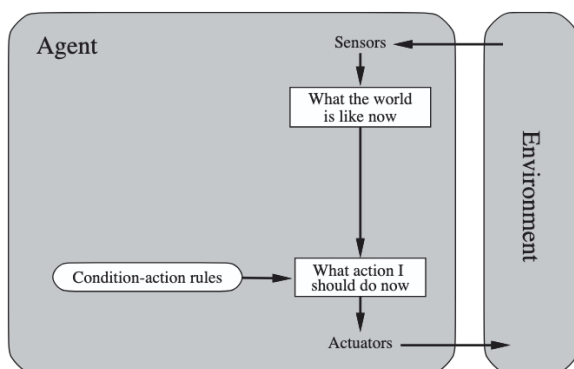


Utility-based agent(基于效用) –considers preference for certain world states over others



### 2.2.2 A Taxonomy of Agent Programs

Reflex (reactive) agent(条件反射) –applies condition-action rules to each percept



Agent with internal state(带内部状态) –keeps track of world

### 2.2.3 Environment Types

- Fully Observable vs Partially Observable
  - Agent's sensors give access to complete (relevant) state of environment (no internal state required)

- Fully: 传感器能获取完整的环境状态 → 不需要内部记忆
- Partially: 只能看到部分环境 → 需要内部状态
- Deterministic vs Stochastic
  - Next state of environment determined only by current state and agent's choice of action
  - Deterministic: 下一状态完全由当前状态 + 动作决定
  - Stochastic: 具有不确定性, 同样动作可能导致不同结果
- Episodic vs Sequential
  - Agent's experience divided into "episodes"; agent doesn't need to think ahead in episodic environment
  - Episodic: 每个决策互不影响, 环境分成独立 "片段"
  - Sequential: 当前决策会影响未来 → 需要前瞻性思考
- Static vs Dynamic
  - Environment changes while agent deliberates
  - Static: 环境在 agent 思考时不变
  - Dynamic: 环境会随时间变化
- Discrete vs Continuous
  - Limited number of distinct, clearly defined percepts and actions
  - Discrete: 感知和动作是有限的、可数的
  - Continuous: 感知和动作是连续的

## 2.3 Layered Architectures and Program

An agent is not a "single-layer brain" but is composed of multiple controllers, each with different functions and levels of abstraction.

Layered Architecture = the higher layers are responsible for thinking (planning, reasoning), while the lower layers are responsible for execution (acting, controlling). In emergency situations, the lower layers can override higher-level commands.

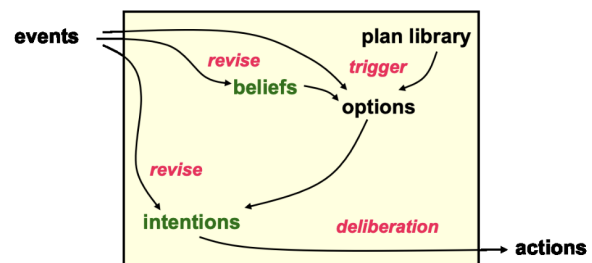
## 2.4 Rational Agents

### 2.4.1 BDI Agents

- Beliefs: Explicit representation of the world

- Desires: Preferred states of the environment
- Goals: Desires the agent has chosen to pursue (must be consistent)
- Intentions: Actions the agent has chosen and committed to
  - Pose problems for deliberation (how to fulfil them)
  - Constrain further choices (must be compatible)
  - Control conduct (lead to future action)

All defined functionally –no X factor  
 "Intentions, Plans, and Practical Reason"  
 (Bratman 1987)



### 2.4.2 PRS (Procedural Reasoning System)

#### Abstract PRS Interpreter:

```

initialize-state()
do
  options := option-generator(event-queue, B, G, I)
  selected-options := deliberate(options, B, G, I)
  update-intentions(selected-options, I)
  execute(I)
  get-new-external-events()
  drop-successful-attitudes(B, G, I)
  drop-impossible-attitudes(B, G, I)
until quit
  
```

### 2.4.3 PRS (Procedural Reasoning System)

- Useful in dynamic environments where
  - Reasonable plans can be formed in advance
  - Agent needs continuity of commitment
  - Agent needs to respond rapidly to situation
  - Agent's computational resources are limited
  - Agent can keep up with changes in the world

## 参考文献

- [1] UNSW COMP9414 Lecture slides –Prof. W.Wobcke
- [2] Other online resources (eg.graph)