

# Agentes Cognitivos Adaptativos

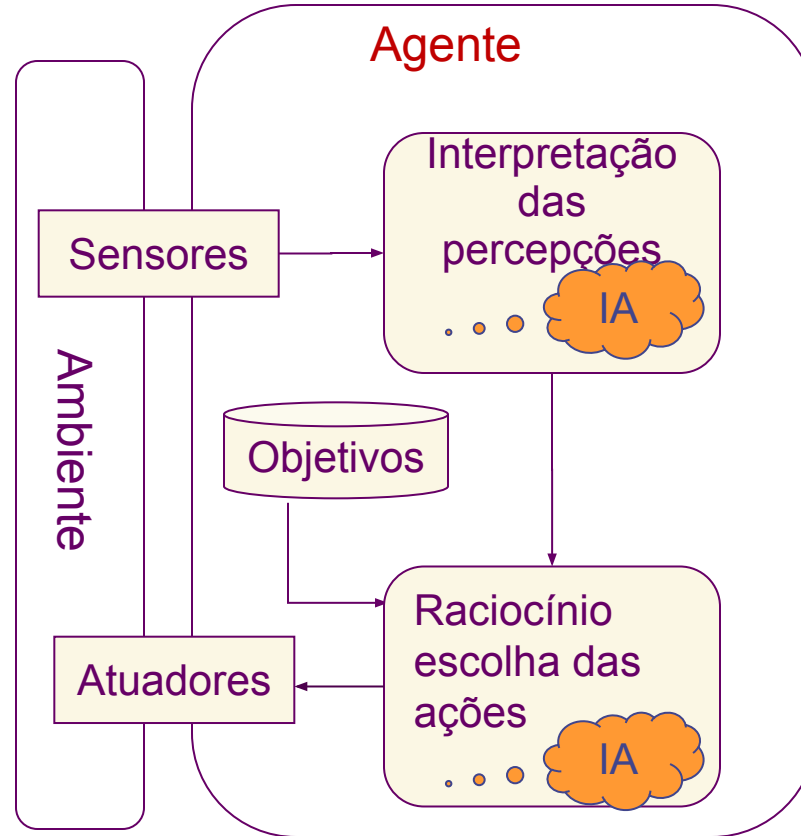
## Agentes Inteligentes Visão geral

Flávia Barros (fab@cin) &  
Ricardo Prudêncio (rbcp@cin)

# O que é um Agente?

- Qualquer entidade (humana ou artificial) que:
  - está imersa ou situada em um ambiente
    - ♦ físico, virtual/simulado
  - **percebe** seu ambiente através de sensores
    - ♦ ex. câmeras, microfone, teclado, ...
  - **age** sobre o ambiente através de atuadores
    - ♦ ex. vídeo, autofalante, impressora, braços, ...
  - possui **objetivos próprios**
    - ♦ explícitos ou implícitos
  - **escolhe** suas ações em função das suas **percepções** para atingir seus **objetivos**

# Agentes Inteligentes



# Agente Racional (Inteligente)

- Agente Racional
  - Toma a melhor decisão ou realiza a melhor ação *possível*
  - Segue o princípio da racionalidade:
    - ◆ dada uma *sequência perceptiva*
    - ◆ o agente escolhe, segundo seus conhecimentos
    - ◆ as *ações* que melhor satisfazem seu *objetivo*
- Racionalidade  $\neq$  Onisciência
  - Limitações de:
    - ◆ sensores
    - ◆ atuadores
    - ◆ Módulo de raciocínio (conhecimento, tempo, etc.)

# Outras propriedades associadas aos Agentes

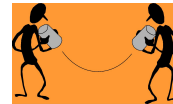
- Autonomia (IA)
  - raciocínio, comportamento guiado por objetivos
  - reatividade



- Adaptabilidade & aprendizagem (IA)



- Comunicação & Cooperação (IA)



- Personalidade (IA)



- Continuidade temporal



- Mobilidade



# Outras propriedades associadas aos Agentes

- Autonomia de raciocínio (IA):
  - Módulo de raciocínio
    - ♦ Regras (base de conhecimento) e máquina de inferência, ou
    - ♦ Aprendizagem de máquina
  - Essencial em sistemas especialistas, controle, robótica, jogos, agentes na internet ...

# Outras propriedades associadas aos Agentes

- Adaptabilidade (IA):
  - Capacidade de adaptação a situações novas
    - para as quais não foi fornecido todo o conhecimento necessário com antecedência
  - Duas implementações
    - ♦ Aprendizagem de máquina e/ou regras+inferência
  - Essencial em agentes na internet, interfaces amigáveis, etc ...

# Outras propriedades associadas aos Agentes

- Comunicação & Cooperação (Sociabilidade) (IA):
  - IA + técnicas avançadas de sistemas distribuídos:
    - ◆ Protocolos padrões de comunicação, cooperação, negociação
    - ◆ Raciocínio autônomo sobre crenças e confiabilidade
    - ◆ Arquiteturas de interação social entre agentes
  - Essencial em sistemas multiagente
- Personalidade (IA):
  - IA + modelagem de traços de personalidade e emoções
  - Essencial em entretenimento digital, realidade virtual, interfaces amigáveis ...



# Outras propriedades associadas aos Agentes

- Continuidade temporal e persistência:
  - Requer interface com sistema operacional e banco de dados
  - Essencial em monitoramento, controle, ...
- Mobilidade:
  - Requer:
    - ♦ Suporte a código móvel
    - ♦ Interface com rede
    - ♦ Protocolos de segurança
  - Essencial em agentes de exploração da internet, ...

# Ambientes

# Ambientes

- Classes de ambientes
  - Ambiente físico: agentes robôs
  - Ambiente de Software: agentes softbots
  - Ambiente de Realidade virtual (simulação do ambiente físico): agentes softbots e avatares
- Propriedades de um ambiente
  - totalmente observável x parcialmente observável
  - determinista x estocástico
  - episódico x sequencial
  - estático x dinâmico
  - discreto x contínuo
  - um agente x multiagentes
  - **complexidade**: número de percepções, ações, objetivos,...

# Ambientes: propriedades

- Totalmente observável
  - quando os sensores do agente conseguem perceber o estado completo do ambiente.
- Determinista
  - o próximo estado do ambiente pode ser completamente determinado pelo estado atual e as ações selecionadas pelo agente.
- Episódico
  - A experiência do agente é dividida em episódios.
  - Cada episódio consiste em o agente perceber e então agir.
  - Cada episódio não depende das ações que ocorreram em episódios prévios.

# Ambientes: propriedades

- Estático
  - o ambiente não muda enquanto o agente está escolhendo a ação a realizar.
  - **Semi estático:** o ambiente não muda enquanto o agente delibera, mas o "score" do agente muda.
- Discreto
  - quando existe um número distinto e claramente definido de percepções e ações em cada turno.
- Contínuo
  - percepções e ações mudam em um espectro contínuo de valores.

# Exemplos de Ambientes

<b>Agente</b>	<b>observável</b>	<b>determinista</b>	<b>episódico</b>	<b>estático</b>	<b>discreto</b>
<b>xadrez sem relógio</b>	Sim	Sim	Não	Sim	Sim
<b>xadrez com relógio</b>	Sim	Sim	Não	Semi	sim
<b>gamão</b>	sim	não	não	sim	sim
<b>motorista de taxi</b>	Não	Não	Não	Não	Não
<b>médico</b>	Não	Não	Não	Não	Não
<b>tutor</b>	Não	Não	Não	Não	Sim
<b>Analizador de imagem</b>	Sim	Sim	Sim	Semi	Não
<b>Busca na web</b>	Não	Não	Sim	Não	Sim
<b>Filtrador de mail</b>	Sim	Não	Sim	Não	Sim

- A complexidade do ambiente é dada por:
  - número de percepções, ações e objetivos possíveis



# Arquiteturas dos Agentes

# Agentes: Arquiteturas

- Agente reativo
- Agente reativo com estado interno
- Agente cognitivo (baseado em objetivos)
- Agente otimizador
- Agente adaptativo

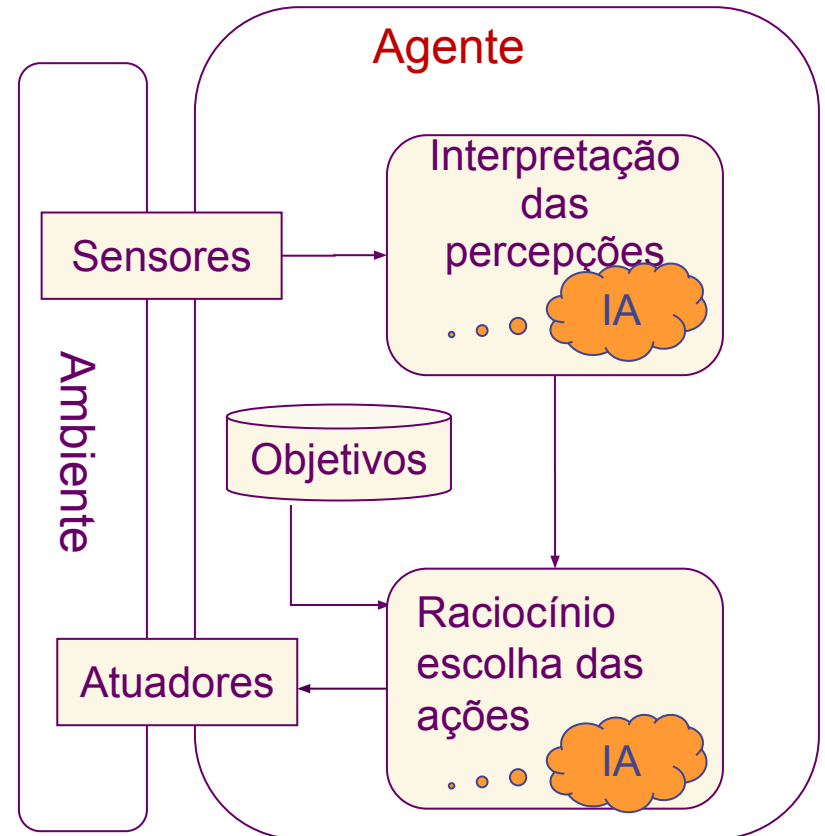


autonomia  
complexidade



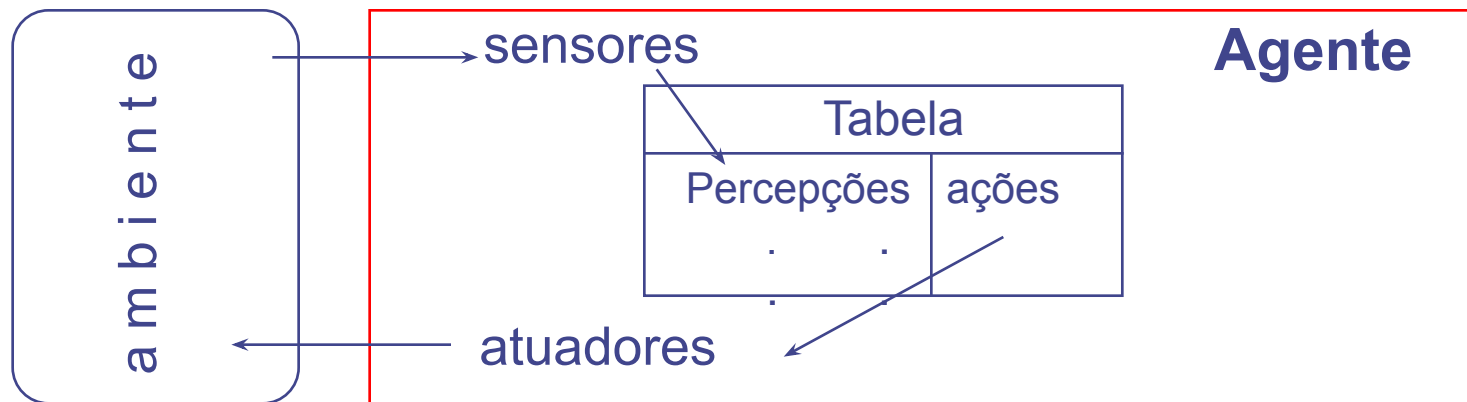
# Agentes: Arquiteturas

- De forma simplificada, um agente pode ser visto como um **mapeamento**:
  - sequência perceptiva => ação



# Agente Tabela?

## Não é um agente inteligente...



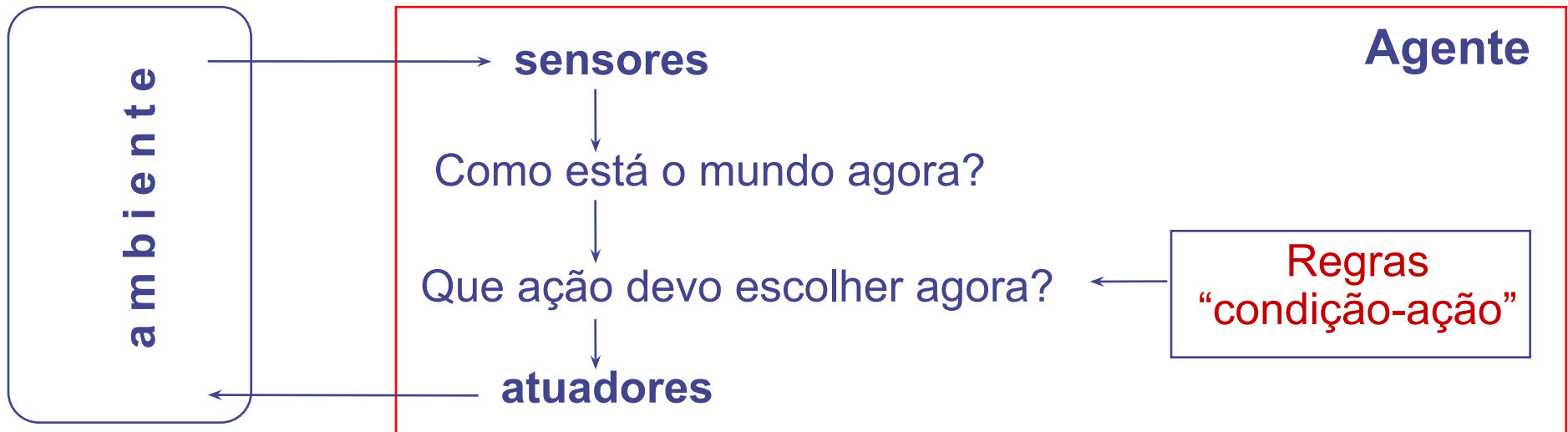
- **Limitações**

- Mesmo problemas simples requerem tabelas muito grandes
  - ♦ ex. **xadrez**  $30^{100}$
- Nem sempre é possível, por ignorância ou questão de tempo, construir a tabela
- Não tem autonomia nem flexibilidade

- **Ambiente**

- Totalmente observável, determinista, episódico, estático, discreto e minúsculo!

# Agente Reativo Simples



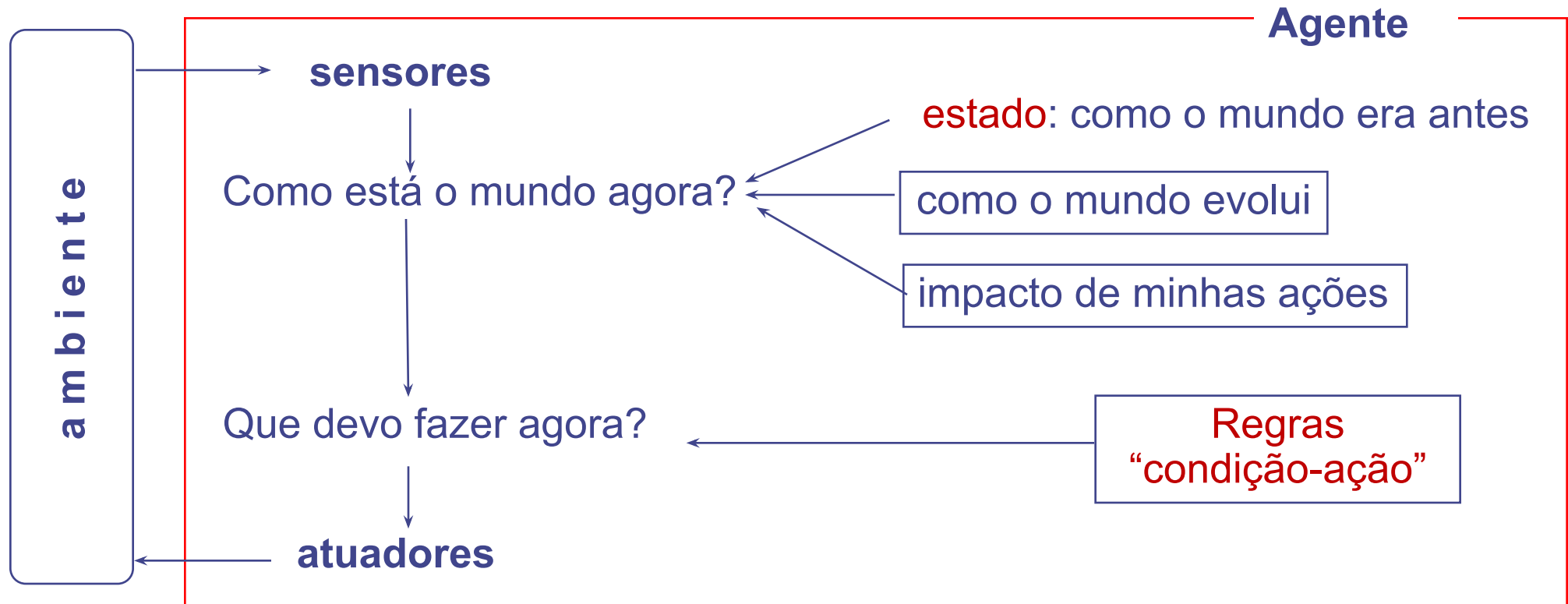
- **Vantagens e desvantagens**

- Regras condição-ação - representação inteligível, modular e eficiente
  - ♦ ex. **Se** velocidade > 60 **então** multar
- Não pode armazenar uma sequência perceptiva, tem pouca autonomia

- **Ambiente**

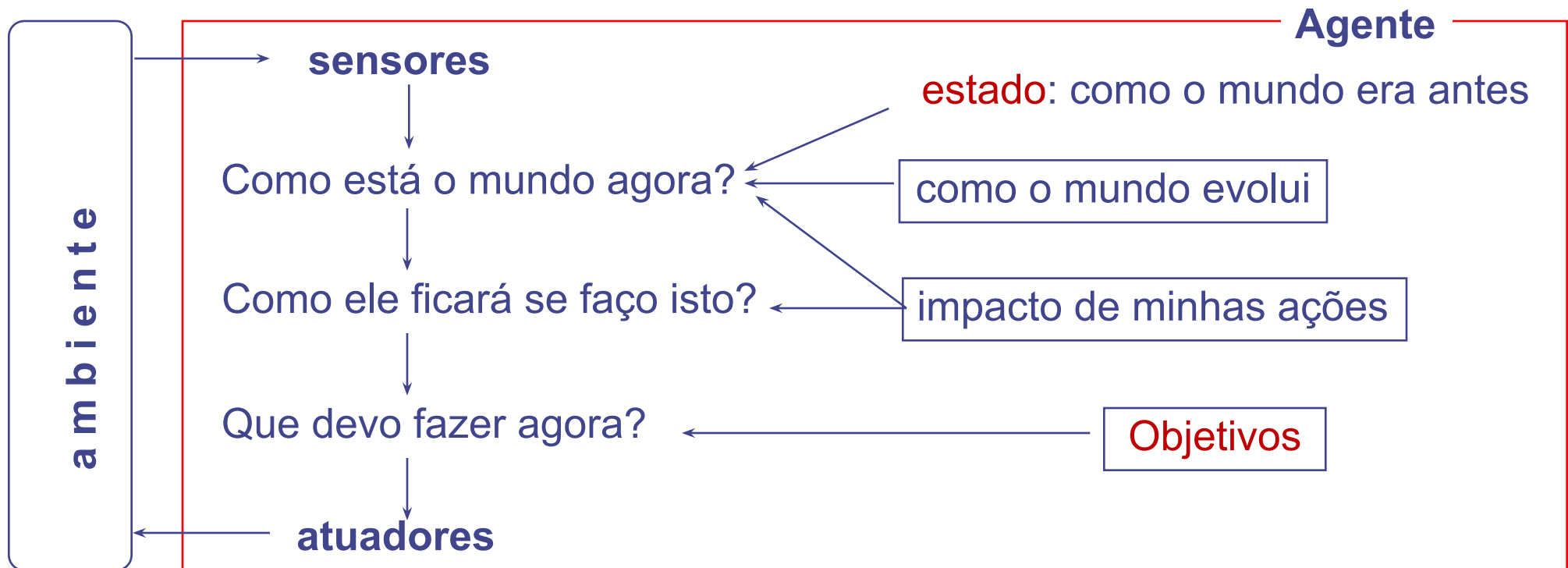
- Totalmente observável, episódico, pequeno
- **Reflexo é imprescindível em ambientes dinâmicos**

# Agente Reativo baseado em Modelo do Mundo



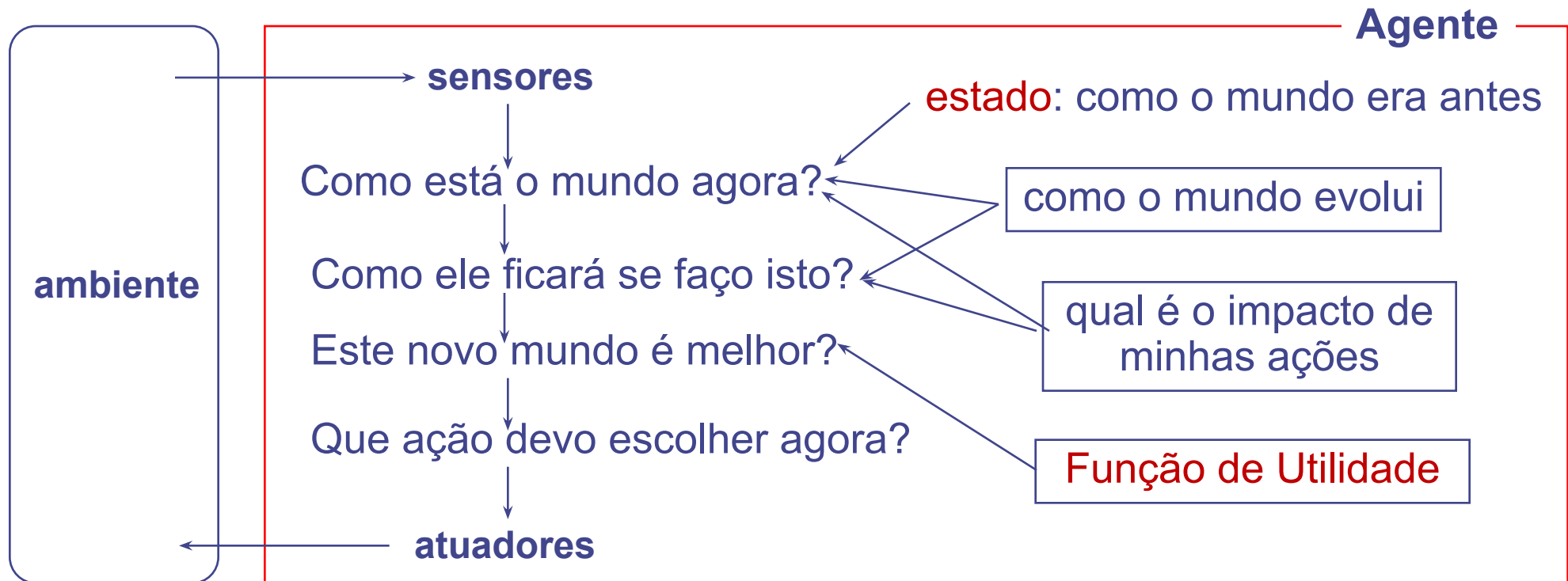
- **Desvantagem:** pouca autonomia
  - não tem objetivo, não encadeia regras
- **Ambiente:** determinista e pequeno
  - Ex. Tamagotchi

# Agente cognitivo - Baseado em Objetivo



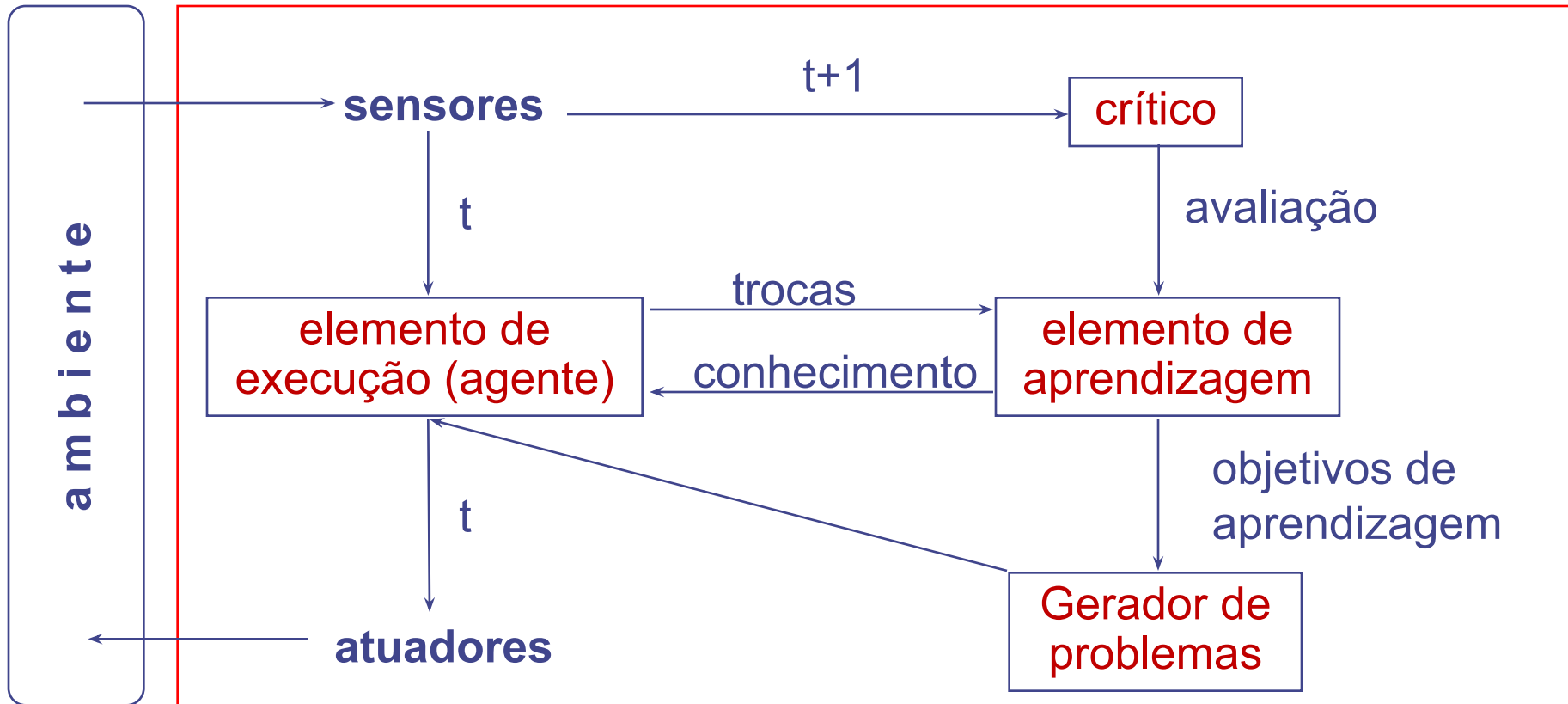
- **Vantagens e desvantagens**
  - Mais **difícil de construir**, porém mais **flexível** e **autônomo**
  - Não trata objetivos conflitantes
- **Ambiente**: determinista
  - Ex. de objetivo: xeque-mate no xadrez

# Agente otimizador - Baseado em Utilidade




- **Vantagem** = Ambiente sem restrição
- Ex. motorista de táxi
  - Segurança e velocidade – **conflito!**
- **Desvantagem:** não tem adaptabilidade

# Agente que aprende



- **Vantagens:** tem adaptabilidade (aprende), ambiente sem restrição
  - **Contudo**, não necessariamente trata objetivos conflitantes !!
- Ex. motorista sem o mapa da cidade

# Afinal, por que usar a “metáfora” de agentes?

- 
1. Oferece uma visão unificadora das várias subáreas da IA
  2. Fornece metodologias de desenvolvimento de sistemas inteligentes estendendo as de engenharia de software
  3. Ajuda a embutir a IA em sistemas computacionais tradicionais
  4. Permite tratar melhor a interação com o ambiente
  5. Permite tratamento natural da IA distribuída



# Pra não dizer que não falei de.... LLMs

- LLMs armazenam conhecimento do mundo que podem ser integrados em arquiteturas de agentes em diversas tarefas
- LLMs podem ser usadas como interfaces de comunicação

# Pra não dizer que não falei de.... LLMs

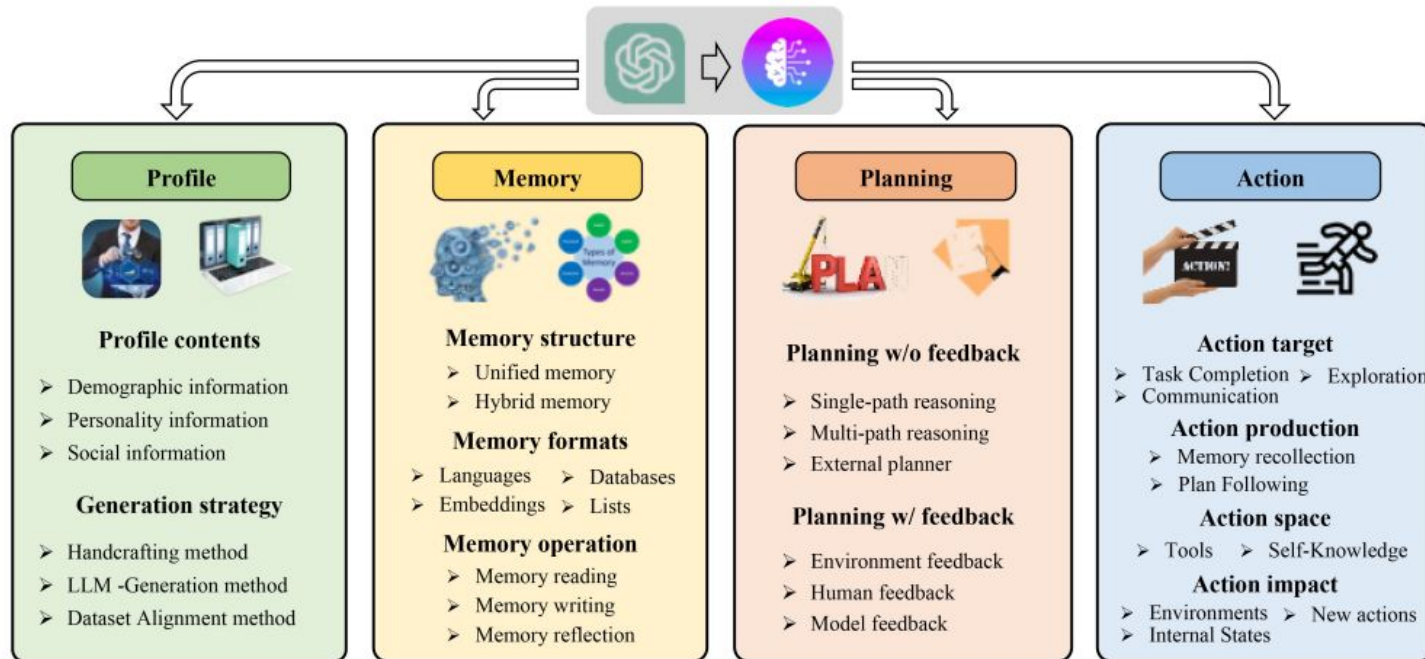


Fig. 2 A unified framework for the architecture design of LLM-based autonomous agent

- Wang et al. (2024). A survey on large language model based autonomous agents - <https://link.springer.com/content/pdf/10.1007/s11704-024-40231-1.pdf>

# Disciplina de Agentes Cognitivos e Adaptativos (ACA)

- Focada nas tarefas de tomada de decisão e planejamento de ações (raciocínio)
- Módulos:
  - Tomada de decisão e tratamento de incerteza
  - Busca e otimização (planejamento)
  - Decisões sequenciais e incerteza (AM por reforço)

# Disciplina de Agentes Cognitivos e Adaptativos (ACA)

- Material básico: Artificial Intelligence - A Modern Approach (Russell and Norvig)
- Avaliação:
  - 4 tarefas individuais (ver cronograma)