

Universidade de Aveiro

Inteligência Artificial (LEI, LECI)

Tópicos de Inteligência Artificial: Conceitos Gerais

Ano lectivo 2025/2026

Regente: Luís Seabra Lopes

Definição de “Inteligência” - I

- Segundo www.dictionary.com, “inteligência” é:
 - Capacidade de adquirir e aplicar conhecimento
 - Capacidade de pensar e raciocinar
 - O conjunto de capacidades superiores da mente.

Definição de “Inteligência” - II

- O estudo da inteligência envolve [Albus,1995]:
 - Como adquirir, representar e armazenar conhecimento
 - Como gerar comportamento inteligente
 - Como surgem e são utilizadas as motivações, emoções e prioridades
 - Como as percepções (sinais) dão origem a entidades simbólicas
 - Como raciocinar sobre o passado
 - Como planejar acções no futuro
 - Como surgem fenómenos como a ilusão, crença, esperança, amor, etc.

Definição de “Inteligência Artificial”

- “Inteligência Artificial” é a disciplina que estuda as teorias e técnicas necessárias ao desenvolvimento de “artefactos” inteligentes. [Nilsson, 1998]
- Direcções seguidas [Russell & Norvig, 1995]:

Pensar como o ser humano	Pensar racionalmente
Agir como o ser humano	Agir racionalmente

Agir como o ser humano

– o Teste de Turing

- “Comportamento inteligente” – a capacidade de um artefacto obter desempenho comparável ao desempenho humano em todas as actividades cognitivas. [Turing, 1950]
- Teste de Turing – é uma definição operacional de comportamento inteligente de nível humano:
 - Consiste em submeter o artefacto a um interrogatório realizado por um ser humano através de um terminal de texto.
 - Se o humano não conseguir concluir se está a interrogar um artefacto ou outro ser humano, então, esse artefacto é inteligente.
- Os sistemas deste tipo serão o objectivo principal da “Inteligência Artificial”?

Ficção científica: HAL9000

- Agente artificial no filme “*Space Odyssey*” [1968], capaz de diálogo em linguagem falada, leitura labial, detecção caras, raciocínio e processamento de emoções



Ficção científica: C-3PO

- Droide de protocolo no filme “*Star Wars*” [1977], auxiliando na etiqueta e tradução, e fluente "em mais de seis milhões de formas de comunicação"!!...



Ficção científica: amor artificial ...

- No filme de Spielberg “A.I.” [2001], um robô humanoide semelhante a uma criança humana é programado para demonstrar amor pelos seus donos



Voltando à definição de “Inteligência Artificial”

- “Inteligência Artificial” é a disciplina que estuda as teorias e técnicas necessárias ao desenvolvimento de “artefactos” inteligentes. [Nilsson, 1998]
- Direcções seguidas [Russell & Norvig, 1995]:

Pensar como o ser humano	Pensar racionalmente
Agir como o ser humano	Agir racionalmente

IA “restrita”, “forte”, etc.

- A “IA restrita” centra-se no desenvolvimento de sistemas de IA especializados que apenas conseguem resolver uma tarefa ou uma gama limitada de tarefas.
- A “IA geral” centra-se no desenvolvimento de máquinas capazes de aplicar a inteligência a qualquer problema.
- A “IA fraca” na literatura filosófica refere-se a visões da IA em que as simulações (parciais) da mente humana não são consideradas mentes.
- A “IA forte” na literatura filosófica refere-se a visões da IA em que os sistemas que simulam corretamente a mente humana são, eles próprios, considerados mentes.

A “sala chinesa” de Searle

- Um humano, que apenas fala uma língua ocidental, documentado com um conjunto de regras escritas num livro nessa língua, e dispondo de folhas de papel, está fechado numa sala.
- Através de uma abertura na sala, o humano recebe folhas de papel com símbolos indecifráveis.
- De acordo com as regras, e em função do que recebe, o humano escreve outros símbolos (que igualmente desconhece) nas folhas brancas e envia-as para o exterior da sala.
- No exterior, no entanto, o que se observa é folhas de papel com mensagens escritas em caracteres chineses a serem introduzidas na sala e respostas inteligentes a essas mensagens a serem devolvidas do interior da sala.

O argumento de Searle

- O humano não percebe chinês
- A sala não percebe chinês
- O livro de regras e as folhas de papel também não percebem chinês
- Logo, não há qualquer compreensão de chinês naquela sala
- No entanto, podemos contra-argumentar: embora, individualmente, os componentes do sistema (a sala, o humano, o livro, as folhas de papel) não compreendam chinês, o sistema no seu conjunto compreende chinês

História até à “Inteligência Artificial”

- Século IV a.C. – Aristóteles estabelece os fundamentos da lógica e do pensamento puramente racional.
- Séculos XVI-XVII – Bacon e Locke estabelecem os fundamentos do “empirismo”: “Nada está na compreensão que não tenha estado primeiro nos sentidos”.
- Séculos XIX-XX – Duas correntes na psicologia: “comportamentalismo” e “cognitivismo”.
- 1940 – início da era dos computadores
- 1943 – McCulloch & Pitts propõem um modelo de computação vagamente inspirado no cérebro humano: redes de unidades chamadas “neurónios” podiam implementar qualquer função.
- 1951 – primeiro programa capaz de jogar xadrez
- 1956 – a expressão “Inteligência Artificial” foi usada pela primeira vez.

História da “Inteligência Artificial” - I

- 1958 – McCarthy usa lógica de primeira ordem para representar conhecimento numa espécie de “sistema pericial”.
- 1959 – *GPS: General Problem Solver* – aqui surge um assunto hoje clássico – pesquisa para resolução de problemas
- 1962 – Rosenblatt propõe o uso do “perceptrão” (rede de neurónios) para aprendizagem e reconhecimento de padrões
- 1966 – *CLS: Concept Learning System* – primeiro sistema de aprendizagem simbólica
- 1966 – Shakey, o primeiro robô autónomo inteligente
- 1970 – Prolog, uma linguagem de programação em lógica
- 1971 – DENDRAL: primeiro sistema pericial (reconstruía a estrutura de moléculas orgânicas)

História da “Inteligência Artificial” - II

- 1986 – Robôs comportamentalistas
- 1986 – Ressurgimentos das redes neuronais
- 1988 – Fundamentos do raciocínio probabilístico
- 1997 – O IBM Deep Blue vence em xadrez contra Kasparov
- 1997 – Primeiro campeonato mundial de futebol robótico
- 2000 – Google é o principal motor de pesquisa para a Web
- 2006 – Aprendizagem profunda (“Deep learning”)
- 2017 – Redes neuronais do tipo “Transformer”
- 2022 – ChatGPT

Variedade de sistemas

- Sistemas de apoio à decisão
 - Utilizam o conhecimento numa determinada área para interpretar dados de problemas específicos e propor decisões a tomar por humanos
- Sistemas autónomos
 - Executam tarefas delegadas por humanos, com um grau de autonomia que depende da aplicação
- Sistemas de Inteligência Artificial Geral

Inteligência Aumentada

- “Aumentar a capacidade de uma pessoa para abordar um problema complexo, para obter uma compreensão adequada às suas necessidades específicas e para derivar soluções para os problemas.” (Engelbar, 1962)

Tópicos de Inteligência Artificial

- Agentes
 - Noção de agente
 - Objectivo da Inteligência Artificial
 - Agentes reactivos e deliberativos
 - Propriedades do mundo de um agente
 - Arquitecturas de agentes
- Representação do conhecimento
- Técnicas de resolução de problemas

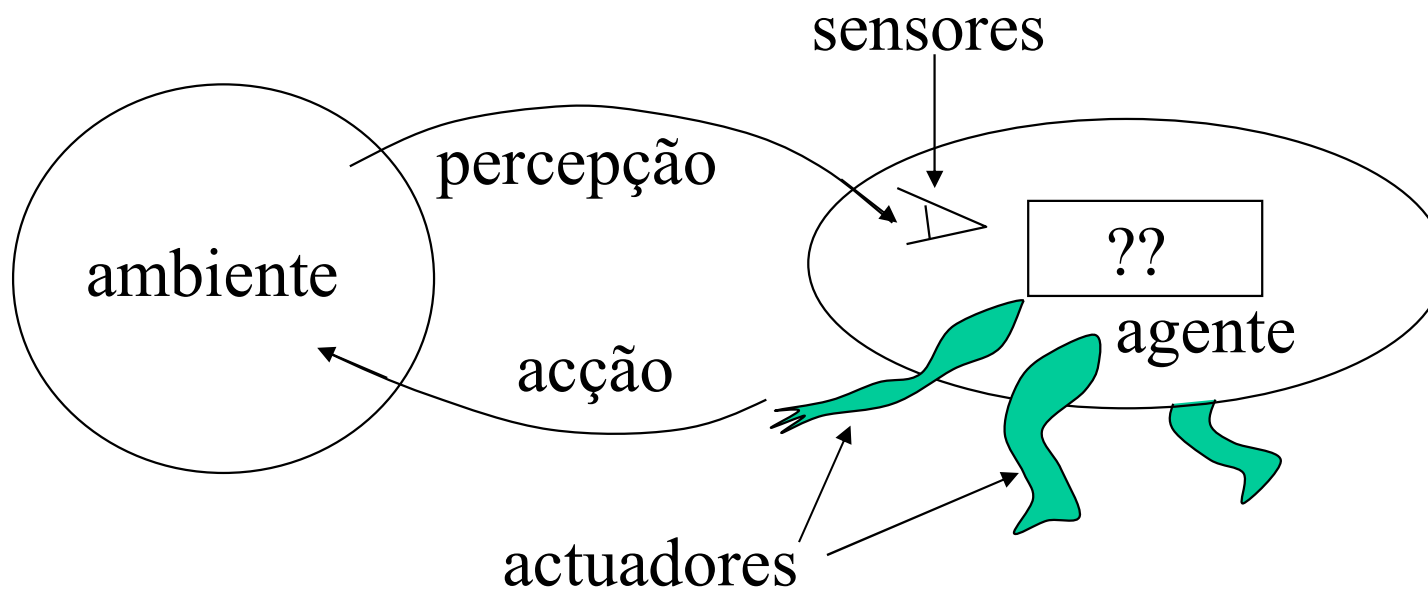
Agentes

- A Inteligência Artificial centra-se no desenvolvimento de “agentes inteligentes”
- O termo “agente” tem dois significados principais:
 - Entidade com poder ou autoridade para agir
 - Entidade que actua em nome de outra

Definição de “Agente”

- Agente – entidade com capacidade de obter informação sobre o seu ambiente (através de “sensores”) e de executar acções em função dessa informação (através de “actuadores”).
 - [Russell & Norvig, 1995]
- Exemplos:
 - Agente físico: robô anfitrião
 - Agente de software: agente móvel de pesquisa de informação na internet

Agente



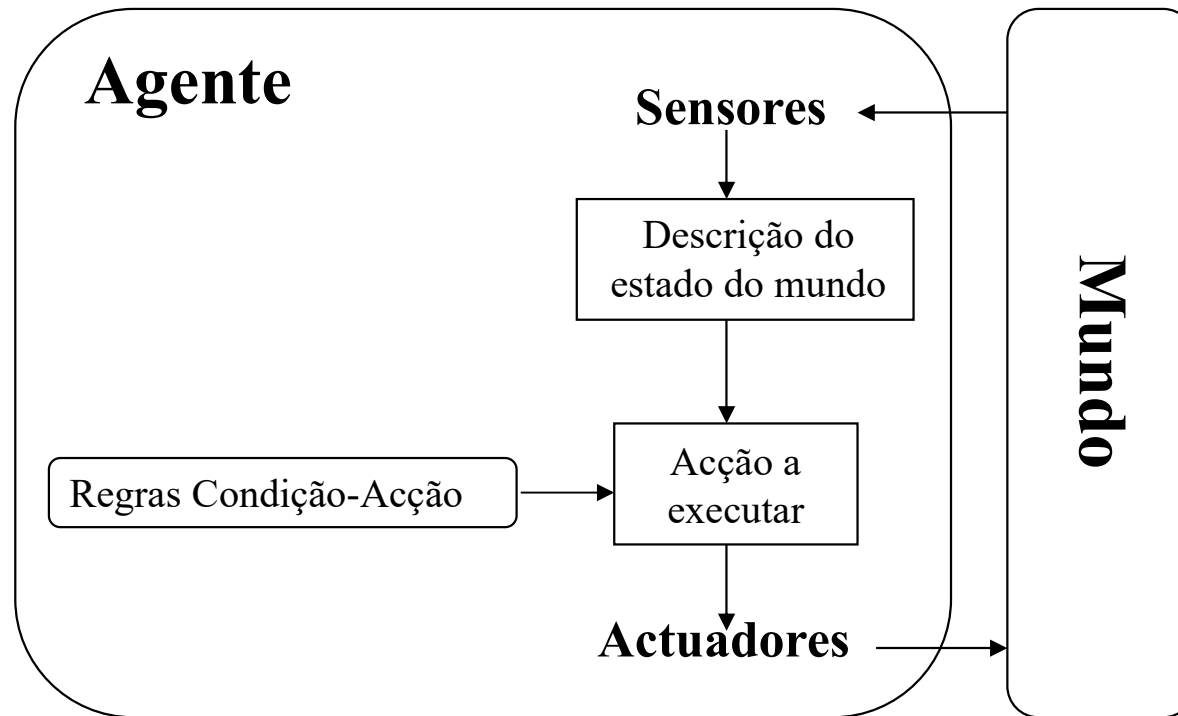
Exemplos de agentes

Tipo de agente	Percepção	Acção	Objectivos	Ambiente
Sistema de diagnóstico médico	Sintomas, respostas do paciente	Perguntas, testes, tratamentos	Saúde do paciente, custo mínimo	Paciente, hospital
Sistema de análise de imagens de satélite	Imagem	Devolver uma categorização da cena	Categorização correcta	Imagens de um satélite em órbita
Braço robótico para em embalagem	Imagem, sinal de força	Colocar peças em caixas	Colocar as peças na posição correcta	Alimentador de peças, caixas
Controlador de refinaria	Temperatura, pressão	Abrir e fechar válvulas; ajustar temperatura	Pureza, segurança	Refinaria
Tutor de inglês interactivo	Palavras introduzidas	Propôr exercícios, corrigi-los, dar sugestões	Maximizar o resultado dos alunos num teste	Conjunto de alunos

Tipos de agentes

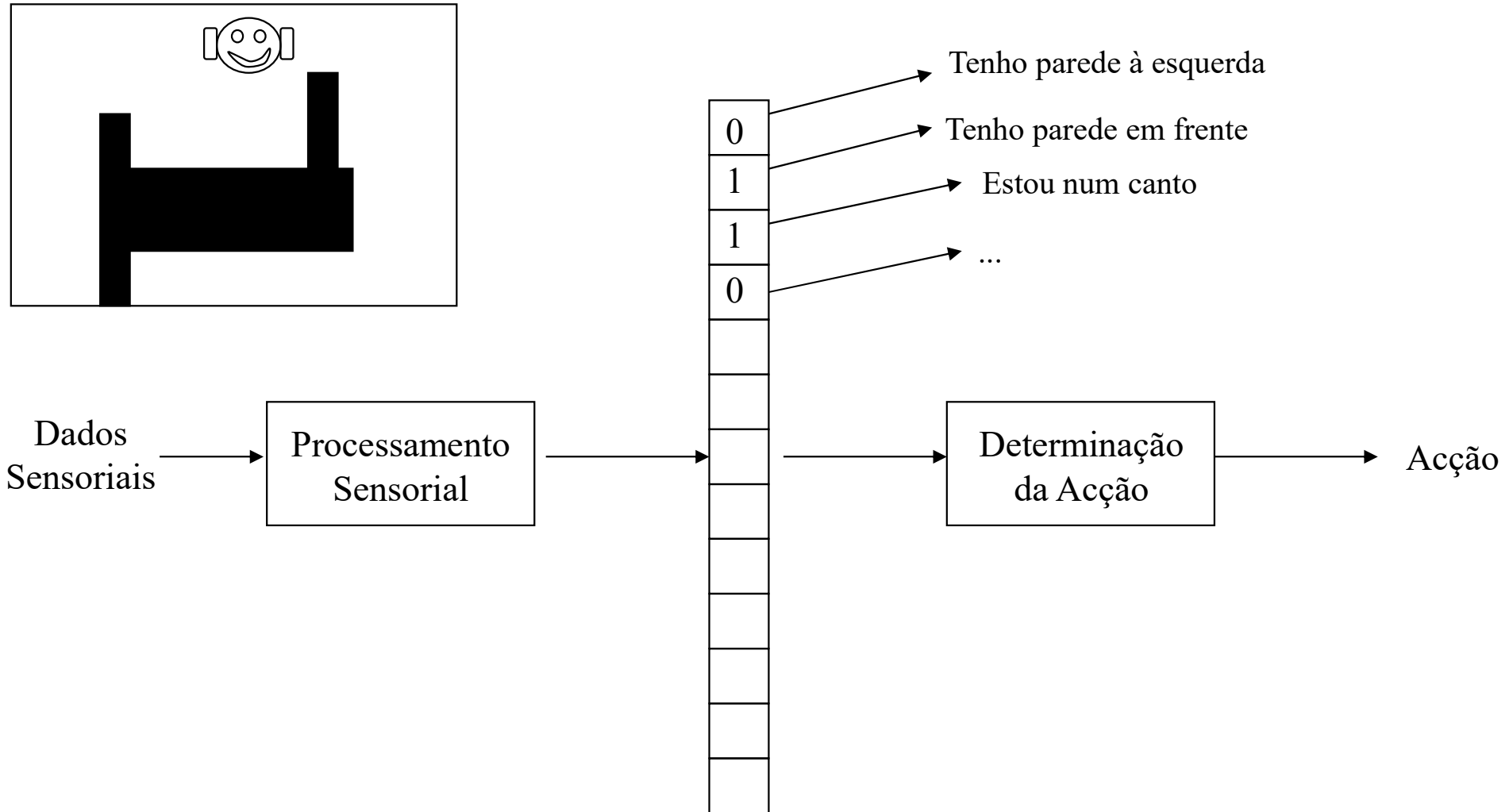
- Reactivos
 - Simples
 - Com estado
- Deliberativos
 - Orientados por objetivos
 - Orientados por medidas de utilidade

Agente reactivo: simples

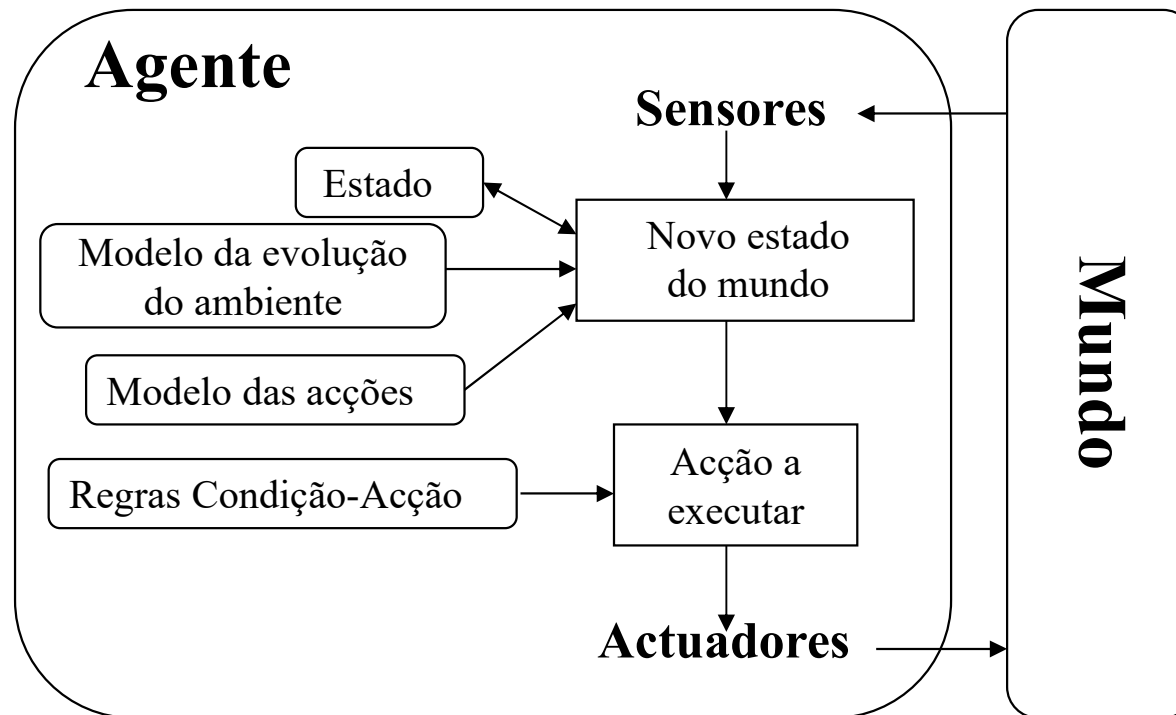


- O conceito de “regra de condição-acção” é também conhecido como “regra de situação-acção” ou “regra de produção”.
- Os agentes ou sistemas reactivos simples são também conhecidos como “sistemas de estímulo-resposta” ou “sistemas de produção”

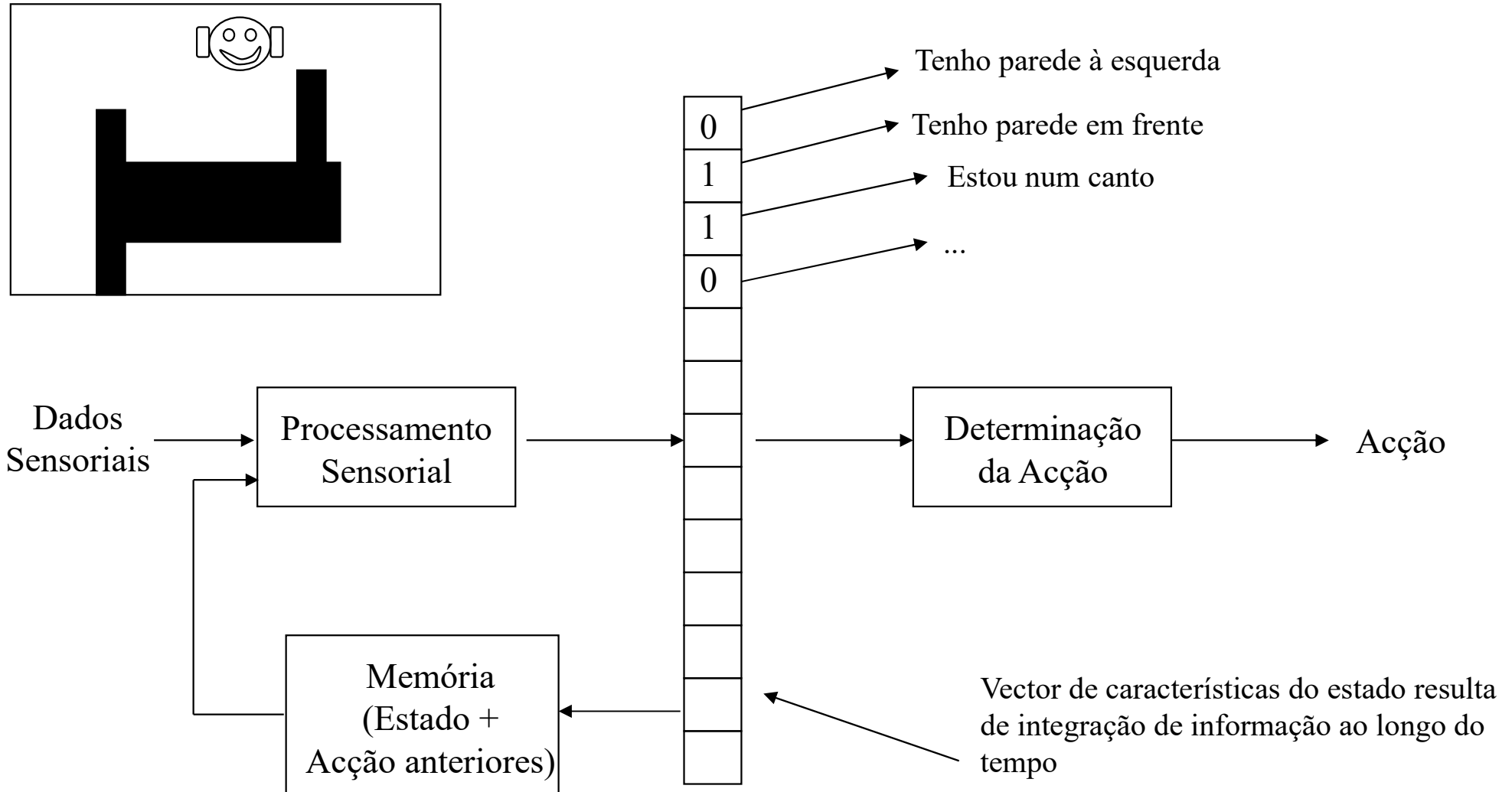
Representando a percepção através de um vector de características



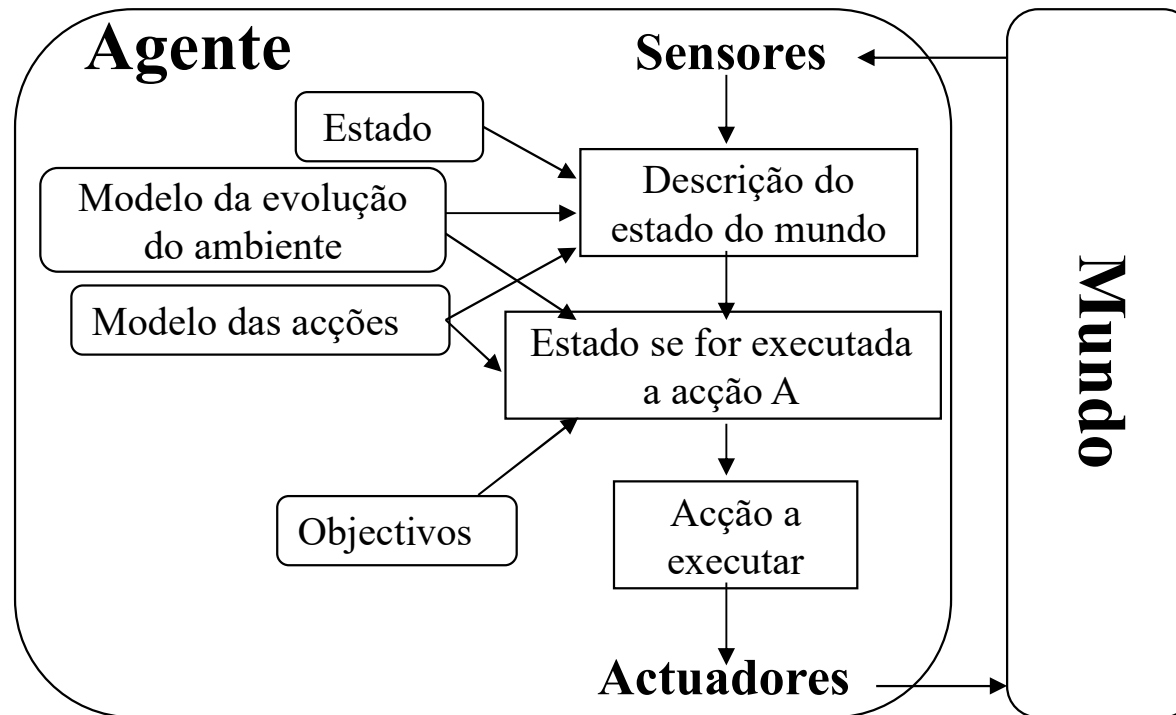
Agente reactivo: com estado interno



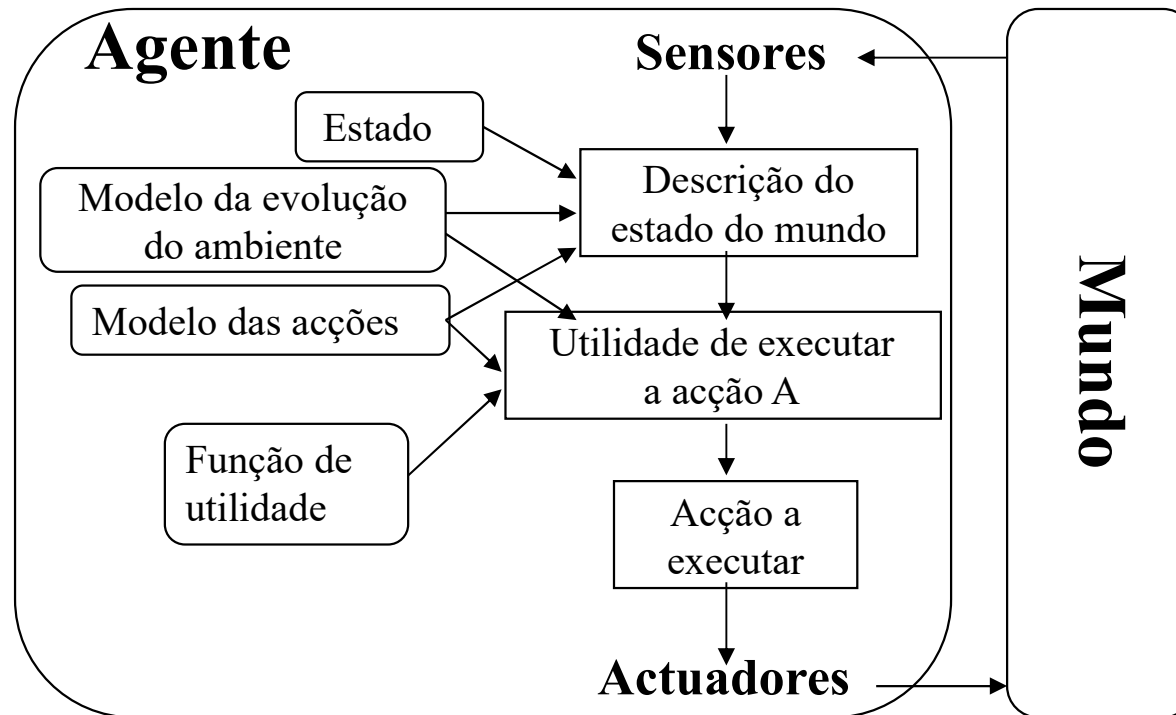
Representando a percepção através de um vector de características: agente com estado



Agente deliberativo: orientado por objectivos



Agente deliberativo: orientado por função de utilidade



Propriedades do mundo de um agente

- Acessibilidade – o mundo é “acessível” se os sensores do agente permitem obter uma descrição completa do estado do mundo; o mundo será “efectivamente acessível” se é possível obter toda a informação relevante ao processo de escolha das acções.
- Determinismo – o mundo é “determinístico” se o estado resultante da execução de uma acção é totalmente determinado pelo estado actual e pelos efeitos esperados da acção.
- Mundo episódico – no caso em que cada episódio de percepção-acção é totalmente independente dos outros.
- Dinamismo – o mundo é “dinâmico” se o seu estado pode mudar enquanto o agente delibera; caso contrário, o mundo diz-se “estático”.
- Continuidade – o mundo é “contínuo” quando a evolução do estado do mundo é um processo contínuo ou sem saltos; caso contrário o mundo diz-se “discreto”.

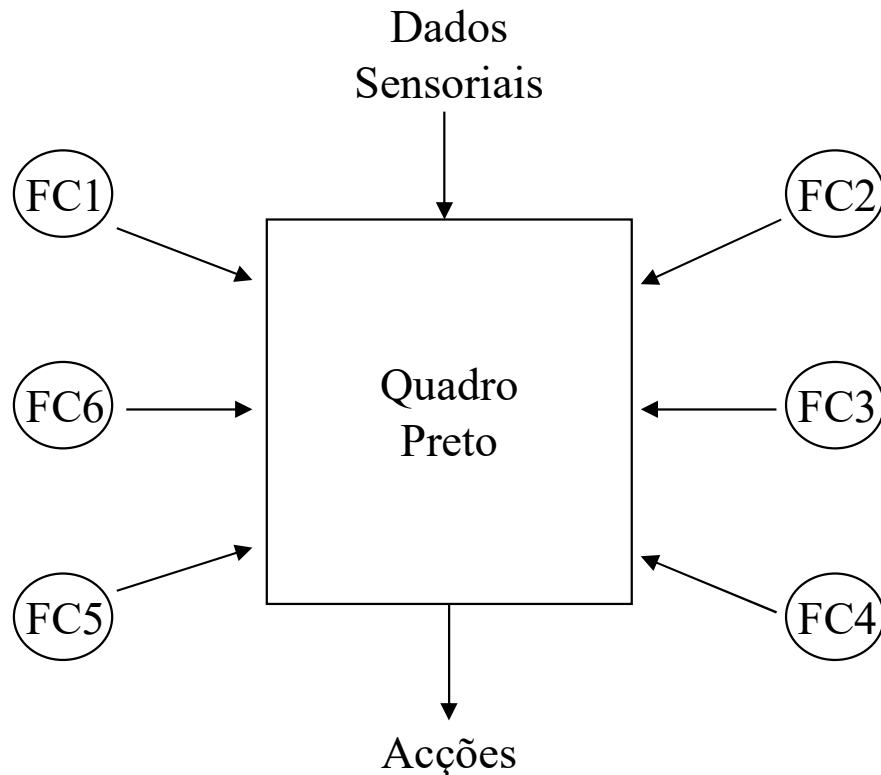
Mundo de um agente: Exemplos

Mundo	Acessível	Determinístico	Episódico	Dinâmico	Contínuo
Xadrês s/ relógio	Sim	Sim	Não	Não	Não
Xadrês c/ relógio	Sim	Sim	Não	Semi	Não
Poker	Não	Não	Não	Não	Não
Condução de carro	Não	Não	Não	Sim	Sim
Diagnóstico médico	Não	Não	Não	Sim	Sim
Sistema de análise de imagem	Sim	Sim	Sim	Semi	Sim
Manipulação robótica	Não	Não	Sim	Sim	Sim
Controlo de refinaria	Não	Não	Não	Sim	Sim
Tutor de Inglês interactivo	Não	Não	Não	Sim	Não

Arquitecturas de agentes

- Quadro Preto
- BDI
- Três torres
- Três camadas
- Subsunção
- CARL

Arquitecturas de agentes: quadro preto

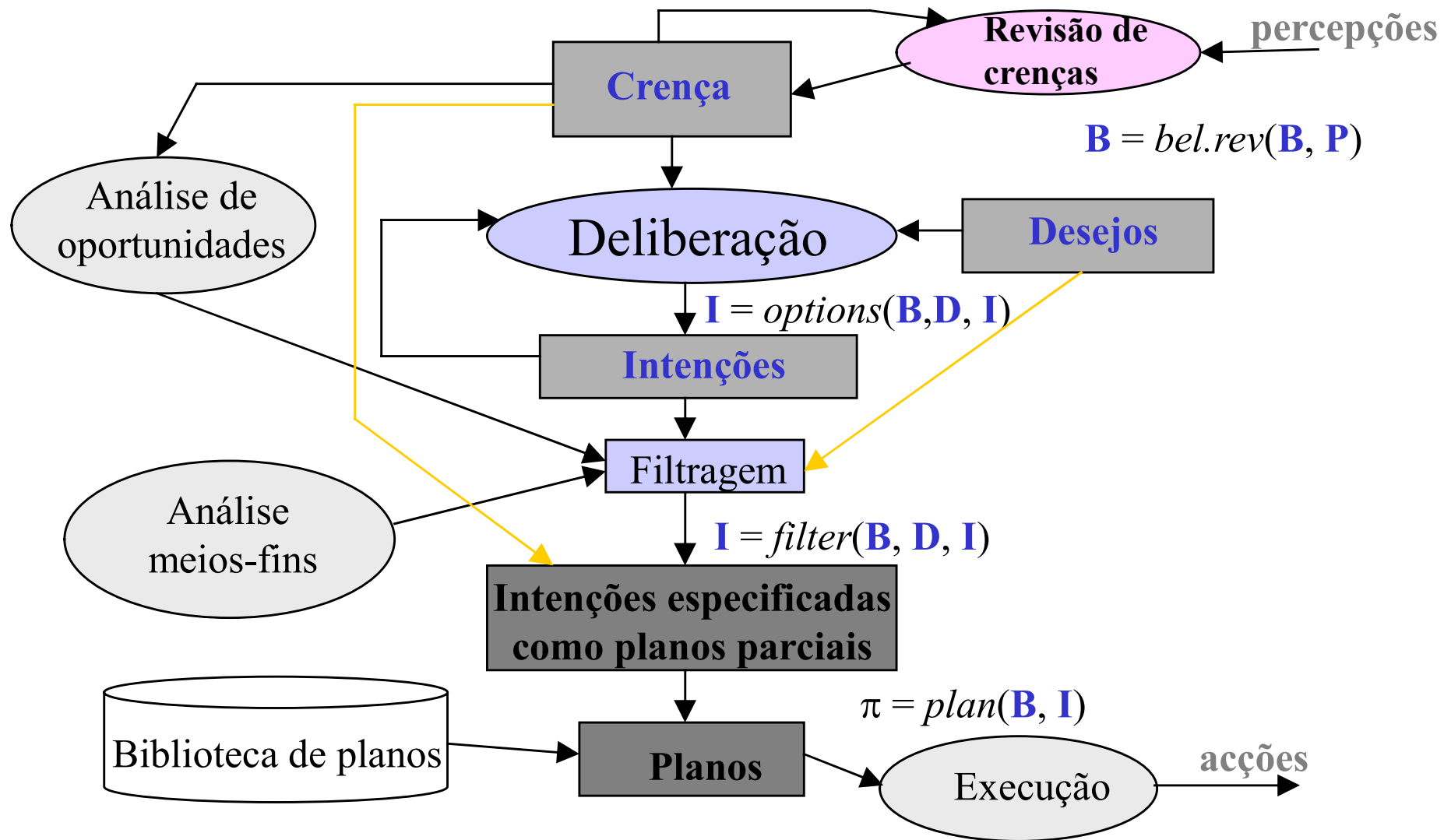


- Podem ser vistos como uma elaboração dos sistemas reactivos com estado interno.
- Uma “fonte de conhecimento” (FC) é um programa que vai fazendo alterações no Quadro Preto.
- Uma FC pode ser vista como um especialista num dado domínio.
- Tipicamente, cada FC rege-se por um conjunto de regras de situação-acção.

Arquitecturas de agentes: BDI

- BDI = Beliefs-Desires-Intentions
 - Ou seja: crenças-desejos-intenções
 - Crenças – informação que o agente possui sobre o mundo
 - Desejos – estados de coisas que o agente gostaria de ver
 - Intenções – estados de coisas que o agente se propõe atingir
- Esta é uma arquitetura para agentes com recursos limitados
 - Combina a análise meios-fins com uma avaliação comparativa de alternativas

Arquitecturas de agentes: BDI



Arquitetura BDI: ciclo de controlo

B $\leftarrow B_0$

D $\leftarrow D_0$

I $\leftarrow I_0$

while true **of**

$p \leftarrow$ Get Next percept

B \leftarrow **bel.rev**(**B**,**P**)

I \leftarrow **options**(**B**,**D**,**I**)

I \leftarrow **filter**(**B**,**D**,**I**)

$\pi \leftarrow$ **plan**(**B**,**I**)

execute (π)

end while

□ Grades:

- The selection and filtering of intent implies a commitment attitude
- In turn, commitment implies a temporal persistence of intentions: after being selected, an intention should not be immediately discarded

Arquitetura BDI: ciclo de controle

B $\leftarrow B_0$

I $\leftarrow I_0$

D $\leftarrow D_0$

while true of

Get Next percept p

B $\leftarrow \text{bel.rev}(\text{B}, \text{P})$

I $\leftarrow \text{options}(\text{B}, \text{D}, \text{I})$

I $\leftarrow \text{filter}(\text{B}, \text{D}, \text{I})$

$\pi \leftarrow \text{plan}(\text{B}, \text{I})$

while not (Empty (π) or succeeded (**I**, **B**) Or impossible (**I**, **B**)) **of**

$\alpha \leftarrow \text{head}(\pi)$

execute (α)

$\pi \leftarrow \text{tail}(\pi)$

$p \leftarrow \text{Get Next percept}$

B $\leftarrow \text{bel.rev}(\text{B}, \text{P})$

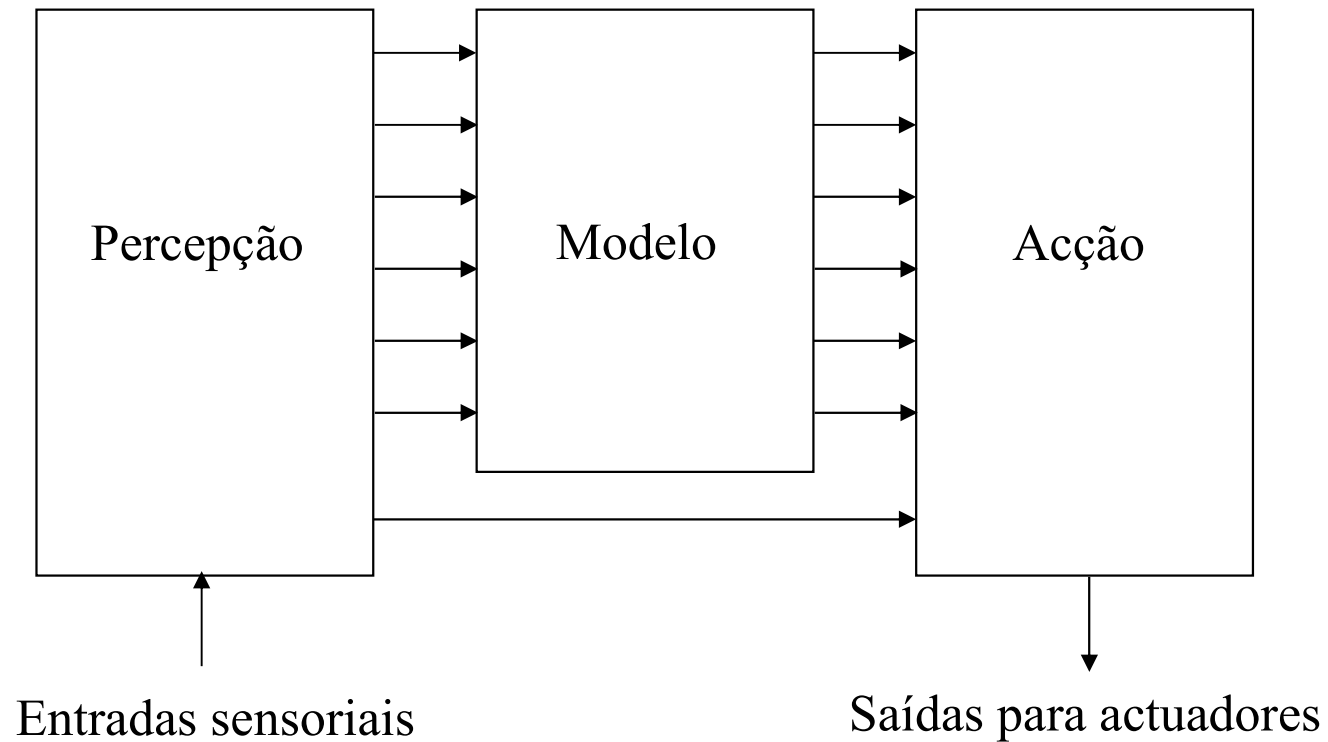
if not sound (π , **I**, **B**) **then**

$\pi \leftarrow \text{plan}(\text{B}, \text{I})$

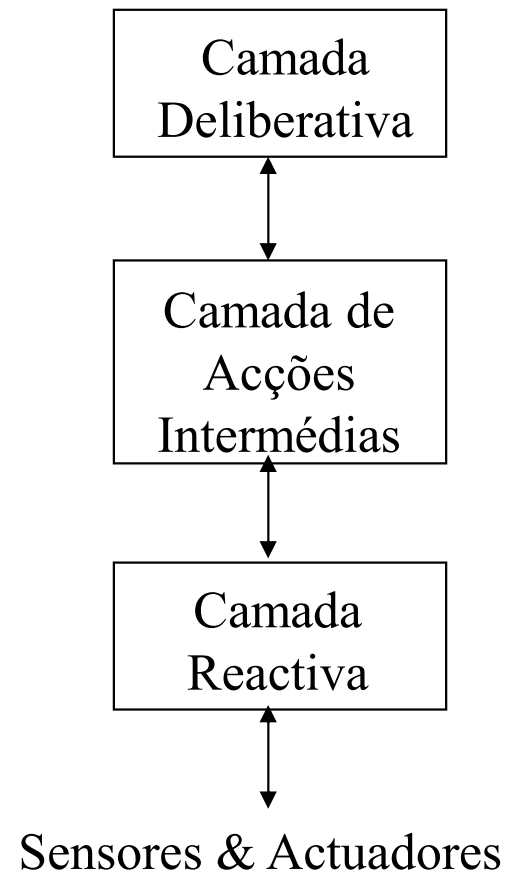
end while

end while

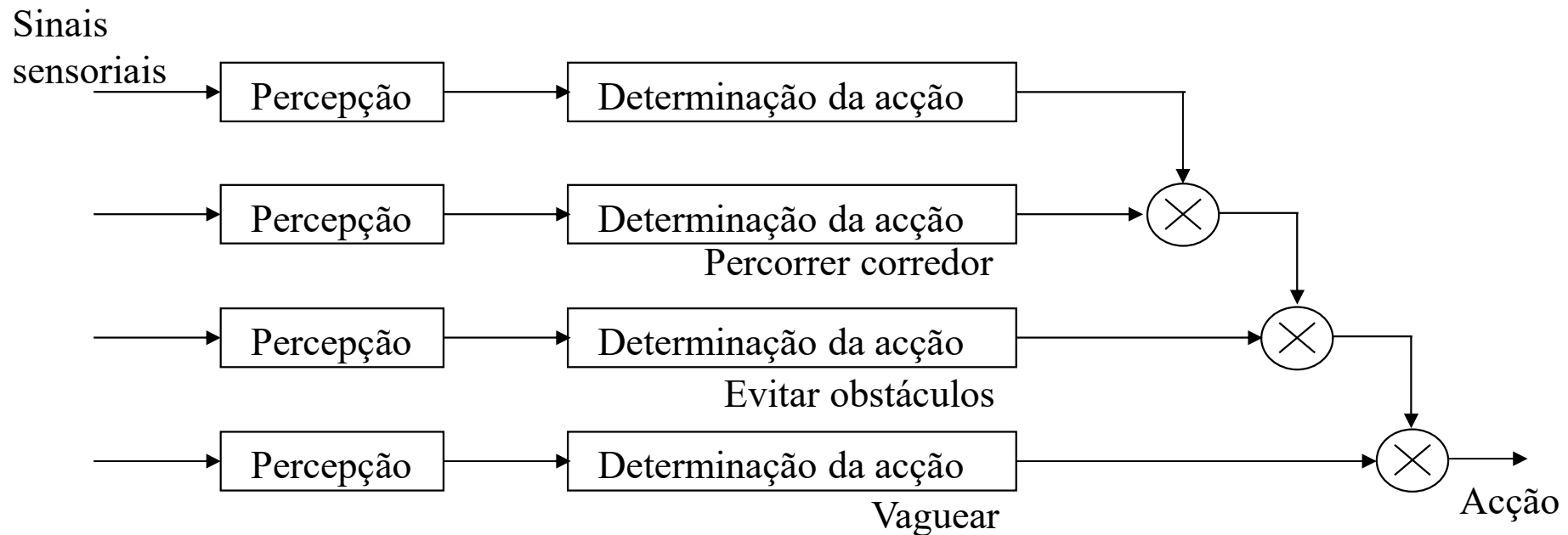
Arquitecturas de agentes: Três Torres



Arquitecturas de agentes: Três Camadas

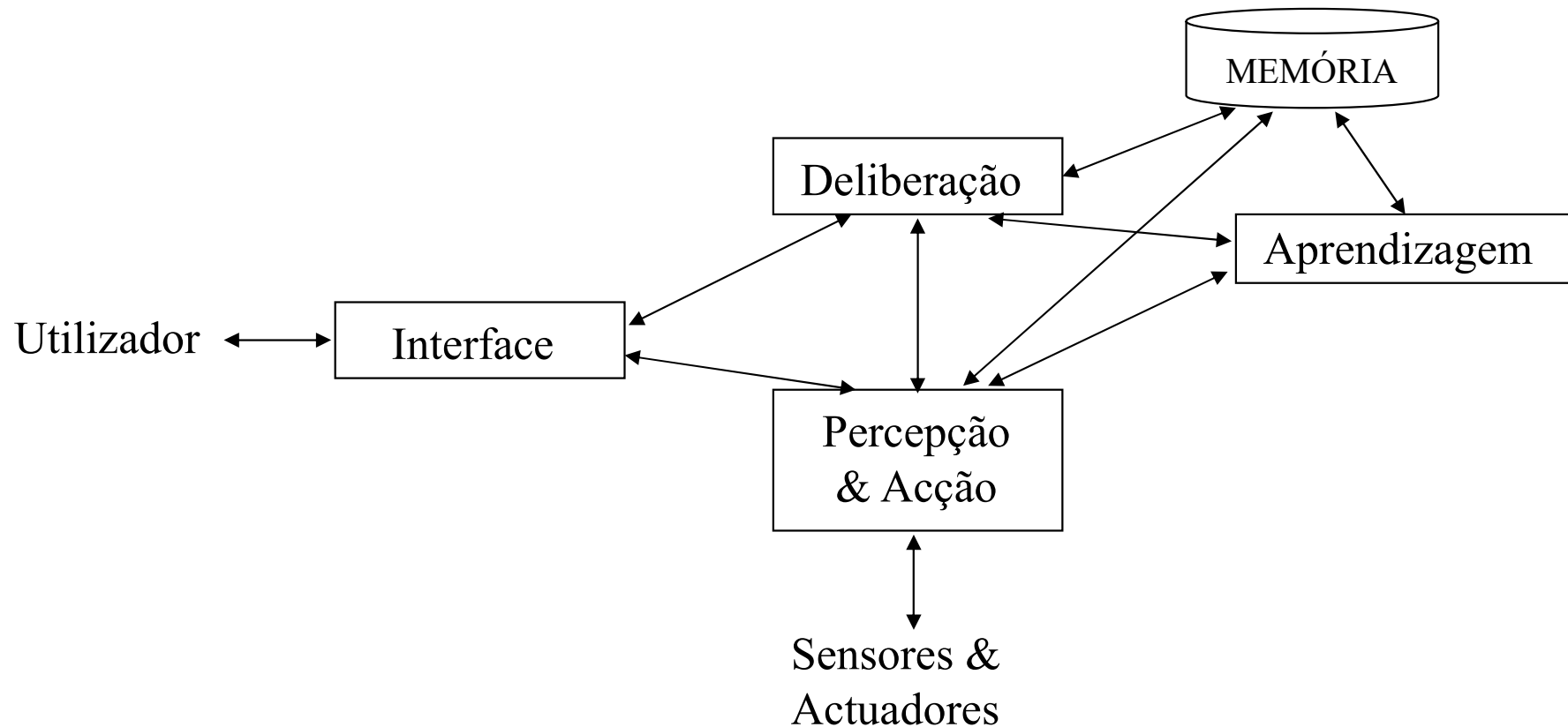


Arquitecturas de agentes: Subsunção



- A arquitectura de subsunção [Brooks,1986; Connell,1990] procura estabelecer a ligação entre percepção e acção a vários níveis – daqui resulta uma organização em camadas.
- A camada mais baixa é a mais reactiva.
- O peso da componente deliberativa aumenta à medida que se sobe na estrutura de camadas.

Arquitecturas de Agentes: CARL



Empresas a investir (muito dinheiro) em IA

- **Microsoft** – reconhecimento de voz a nível humano; Copilot; parceria com OpenAI;
- **IBM** – sistema de perguntas e respostas em linguagem natural (Watson) evoluindo para IA Geral
- **Google** – aprendizagem automática, processamento de linguagem, processamento visual, classificação de páginas, previsão
- **Apple** – agente virtual inteligente (Siri)
- **Meta** – interpreta os dados partilhados nas suas redes sociais; integração de IA generativa
- **Amazon / AWS** – agente virtual inteligente (Alexa), Amazon Machine Learning

Empresas a investir em IA (cont.)

- **CloudMinds** – sistemas baseados em inteligência na nuvem (cloud) para robôs
- **iCarbonX** – fornece análises de saúde personalizadas e previsões de índices de saúde
- **Tesla** – condução semiautónoma
- **ViSenze** – aprendizagem automática e visão computacional para recomendação de artigos visualmente semelhantes em compras online
- **Zebra Medical Systems** – aprendizagem profunda para detectar múltiplas doenças a partir de radiologia com uma precisão superior à humana
- **OpenAI** – foco em IA geral; IA generativa
- **Alphabet** – IA generativa; Gemini

Empresas a investir em IA (cont.)

- **NVidia** – GPUs; soluções diversas para empresas, cobrindo ciência de dados, agentes conversacionais, etc.
- **AMD** – Visão integrada da IA, cobrindo hardware e software

Obstáculos à IA nas empresas

Obstáculo	%cases
Não existe um caso de negócio definido	42%
Não está claro para que pode ser utilizada a IA	39%
Não possui as competências necessárias	33%
Antes, tem que investir na modernização da plataforma de gestão de dados	29%
Não tem orçamento	23%
Não tem a certeza do que é necessário para implementar um sistema de IA	19%
Os sistemas de IA não estão comprovados	14%
Não têm os processos ou a gestão correta	13%
IA é uma moda com pouca substância	11%
Não possui nem tem acesso aos dados necessários	8%
Não tem a certeza do que significa IA	3%

Áreas de aplicação

- Operações autónomas de serviços de campo
 - Optimização de rotas
 - Outras otimizações logísticas
- Antecipação das necessidades dos clientes
 - Processar grandes volumes de dados dos clientes
- Personalização em serviços financeiros
 - Recomendações oportunas de produtos
 - Serviços de consultoria
- Saúde
 - Diagnóstico
 - Assistentes virtuais para tirar dúvidas dos pacientes

Principais tecnologias de IA

- Geração de linguagem natural
 - Interação com o cliente
 - Relatórios
- Reconhecimento de fala
 - Agentes virtuais, robótica, aplicações móveis
- Agentes virtuais
 - De agentes conversacionais a agentes mais avançados
- Plataformas de aprendizagem automática
 - Algoritmos, APIs, módulos de ferramentas (toolkits), conjuntos de dados, poder computacional

Principais tecnologias de IA (cont.)

- Hardware otimizado para IA
 - Abordado principalmente no contexto da aprendizagem profunda
- Gestão de decisões
 - Regras, lógica
 - Tecnologia madura
- Plataformas de aprendizagem profunda
 - Redes neurais utilizadas principalmente no reconhecimento e classificação de padrões com base em grandes conjuntos de dados
- Biometria
 - Reconhecimento facial, linguagem corporal, fala, ...

Principais tecnologias de IA (cont.)

- Robótica e automação
 - Substituir os humanos em tarefas perigosas ou fastidiosas
- Análise de texto e processamento de linguagem natural
 - Análise semântica, análise de sentimentos, reconhecimento de intenções
 - Mineração de texto, assistentes virtuais

Problemas em aberto na IA

- O significado das palavras
 - Os sistemas atuais processam os textos em grande parte sem compreender o significado das palavras
 - O significado deve ser, em última análise, baseado na percepção e na ação
- Programação de robôs
 - Comunicação intuitiva
 - Treino comportamental através de testes, possivelmente em simulação
- Para além dos jogos de tabuleiro
 - O Xadrez e o Go são jogos estruturados que não têm a complexidade dos problemas do mundo real

Problemas em aberto na IA

- Segurança e ética
 - Proteção contra hackers
 - Distinguir o certo do errado

Implicações éticas e sociais

- Em fevereiro de 2017, o Parlamento Europeu (PE) aprovou por maioria uma recomendação para a adoção de regras relativas à robótica e à inteligência artificial, tendo em conta:
 - Responsabilidade civil
 - Implicações éticas e sociais
- O PE recomendou ainda à Comissão Europeia que considerasse a criação de um estatuto jurídico para os robôs que atuam de forma autónoma em interação com humanos
- 2024/08/01 – Entrada em vigor da Lei da Inteligência Artificial na União Europeia

Evolução futura da IA

- Crescimento exponencial da tecnologia de IA?
- Superinteligência?
- Humanos dominados por máquinas?
- Redução do número de empregos?
- Mais tempo livre com rendimento mínimo garantido?