



Universidade do Minho

Escola de Engenharia

Departamento de Informática

Agentes

Mestrado Integrado em Engenharia Informática

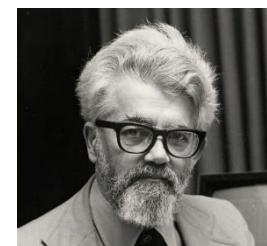
Mestrado em Engenharia Informática

Perfil SI :: Agentes e Sistemas Multiagente

The expression Artificial Intelligence appear for the first time in 1955 and at the Dartmouth Conference (1956) it was defined as:

“(Artificial Intelligence is) making a machine behave in ways that would be called intelligent if a human were so behaving.”

John McCarthy



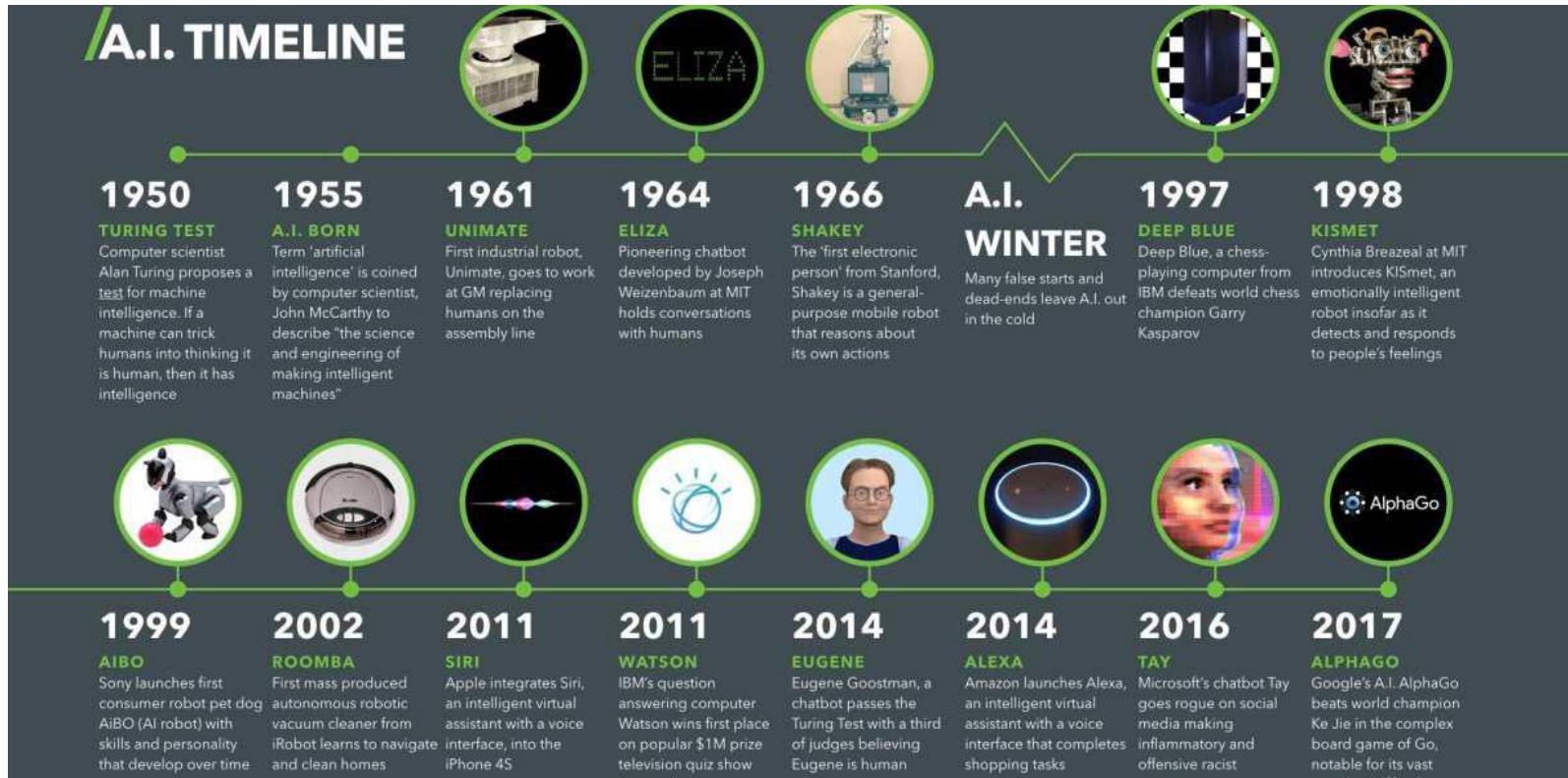
Source: Wired
<https://www.wired.com/2011/10/john-mccarthy-father-of-ai-and-lisp-dies-at-84/>

Artificial Intelligence

It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.

John McCarthy, 2007

AI Timeline



Source: Paul Marsden

<https://digitalwellbeing.org/artificial-intelligence-timeline-infographic-from-eliza-to-tay-and-beyond/>

**“Our Intelligence is what makes us
human, and AI is an extension of that
quality”.**

Yann LeCun (A.M. Turing Award 2018)

- Inteligência Artificial
- Inteligência Artificial Distribuída
- Agente Inteligente
 - Noção, Definição e Propriedades
- Fontes de Inspiração
- Tipologia de um Agente
- Arquiteturas
 - Reativas, Deliberativas e Híbridas;
 - *Beliefs, Desires and Intentions* (BDI)
- Tipos de Agentes
- Áreas de Aplicação
- Conclusões
- Referências



- Inteligência

- “faculdade de compreender, um talento, o raciocínio, a habilidade”;

- Inteligência Artificial:

- [...]
 - Ramo da informática que estuda o desenvolvimento de sistemas computacionais com base no conhecimento sobre a inteligência humana.

- Inteligência Artificial:

- “Ramo das Ciências da Computação que tem a ver com a automatização de comportamentos inteligentes, próprios dos seres humanos”.



[Luger&Stubblefield, 1998]

Inteligência Artificial Distribuída

- Inteligência Artificial Distribuída (IAD):

- “Ramo da Inteligência Artificial que estuda a resolução de problemas através de sistemas computacionais distribuídos”.

[Bond & Gasser, 1988]



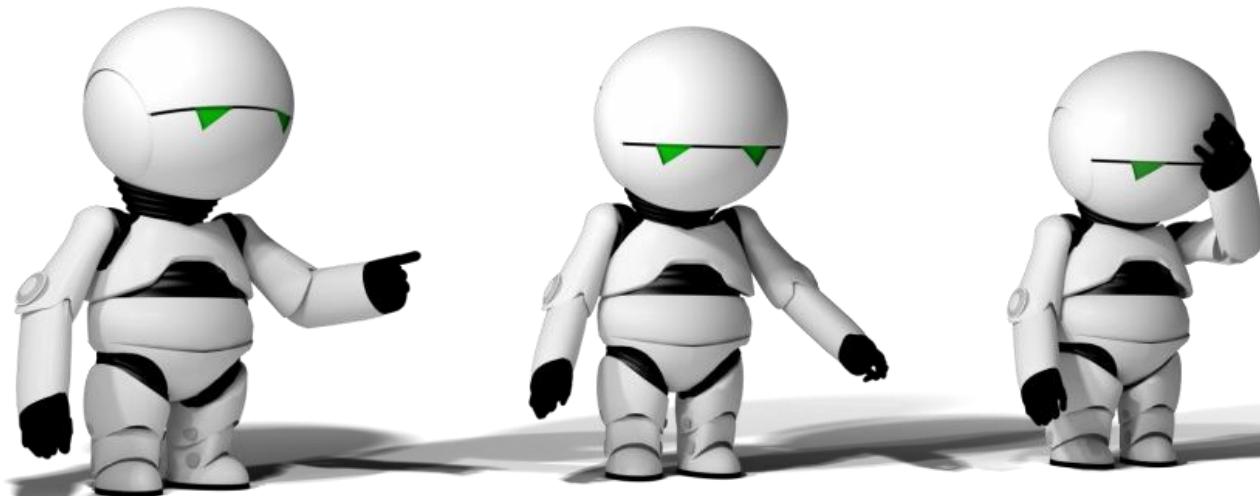
Inteligência Artificial Distribuída Agentes e Sistemas Multiagente

- “Estudo, construção e aplicação de sistemas em que diversas entidades computacionais (agentes ou Sistemas Multiagente - SMA) interagem e perseguem um conjunto de objetivos e/ou realizando um conjunto de tarefas”.

[Weiss, 1999]



- “As plataformas de computação e os sistemas de informação atuais são heterogéneos, abertos e distribuídos”.
- Os agentes e os SMA surgem como uma nova Metodologia Computacional da Engenharia de *Software*;
- A ideia chave passa a ser “O agente inteligente que interage”.



- “Algo que age; capaz de produzir um efeito”.



- Agente: Apercebe-se do ambiente através de sensores e age nesse ambiente através de atuadores
- Humano:
 - Sensores: Olhos, ouvidos, nariz, tacto, gosto, outros
 - Atuadores: Pernas, braços, mãos, outros
- Agente robótico:
 - Sensores: câmaras, sonares, sensores de infravermelhos, microfone...
 - Atuadores: motores, rodas, alto-falante, etc.

Noção de Agente (*de software*)

- “Um agente é algo que obtém informação e conhecimento do ambiente através de sensores e atua nesse ambiente através de atuadores”;

[Russell & Norvig, 1995]

- “Agentes como componentes persistentes e ativos que percebem, raciocinam, atuam e comunicam”;

[Huhns & Singh, 1997]

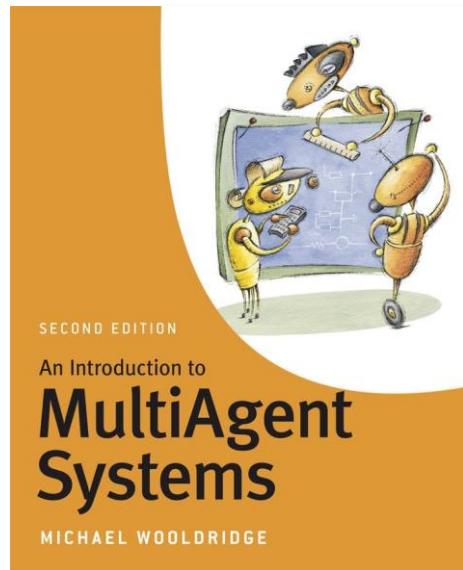
- “Agentes são entidades que habitam em ambientes complexos, sentem esse ambiente e atuam de modo autônomo, procurando executar um conjunto de tarefas para as quais receberam procuração”.

[Pattie Maes, 1990]

Definição de Agente Inteligente

- “Um agente corporiza um sistema computacional capaz de revelar uma ação **autónoma** e **flexível**, desenvolvido num determinado universo de discurso. A flexibilidade do agente está relacionada com as suas capacidades de **reação, iniciativa, aprendizagem e socialização**.”

[Wooldridge, 1999]



“An agent is a computer system that is situated in some environment, and that is capable of autonomous action in order to meet its delegated objectives.”

[Wooldridge, 2009]

Normalmente, o **controle** sobre o ambiente é apenas **parcial**. Os agentes devem estar preparados para o insucesso/falha. Propósito: “delegar objetivos de alto nível sem ter que especificar como alcançá-los!” Questão: que propriedades podemos utilizar para descrever agentes?

Fontes de inspiração

Inteligência Artificial

Resolução de Problemas
Raciocínio e Conhecimento
Planeamento
Aprendizagem

Sistemas Distribuídos e Redes de Computadores

Arquiteturas
Sistemas Multiagente
Comunicação e Coordenação

Sociologia

Sociedades Virtuais
Interação

Engenharia de *Software*

Agente como abstração
Programação orientada por Agentes

Teoria dos Jogos e Economia

Negociação
Resolução de Conflitos
Mecanismos de Mercado

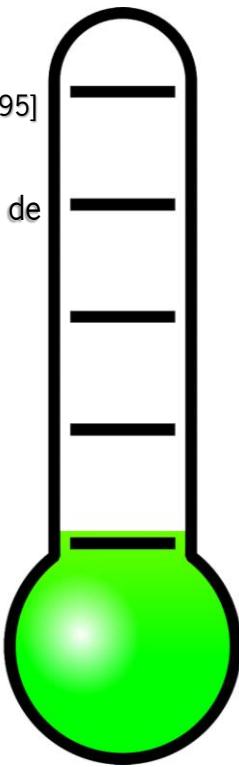
Noção Fraca de Agente Propriedades

- Noção **fraca** de agente:

- conjunto mínimo de propriedades/características que um agente inteligente deve exibir:

[Wooldridge & Jennings, 1995]

- Autonomia
 - os agentes operam sem intervenção de outros agentes e controlam as suas ações e o seu estado de conhecimento interno;
 - Reatividade
 - os agentes percecionam os eventos que ocorrem no seu universo de discurso e respondem adequadamente e atempadamente a mudanças ocorridas nesse ambiente;
 - Iniciativa:
 - os agentes tomam iniciativa, conduzindo as suas próprias ações mediante um comportamento dirigido por objectivos;
 - Sociabilidade:
 - os agentes relacionam-se com outros agentes, comunicando, competindo ou cooperando na resolução de problemas que lhes sejam colocados.



Noção Forte de Agente Propriedades

- Noção **forte** de agente:

- um agente é considerado forte, quando as entidades com que se depara são eminentemente cognitivas, passíveis de desenvolver a sua própria consciência, de se apresentar como tendo um conjunto de mais-valias como a percepção, a sentimentalidade e o emocionismo.

[Wooldridge & Jennings, 1995] [Ferber, 1999], [Nwana, 1996], [Russell & Norvig, 1995]

- Mobilidade:

- capacidade de se movimentar através da rede formada pelos seus pares, executando as tarefas de que foi incumbido;

- Intencionalidade:

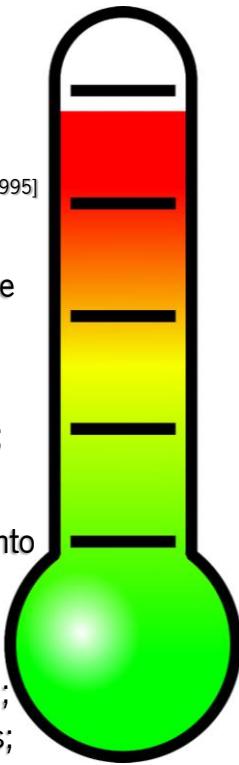
- capacidade que o agente apresenta para a definição de objetivos e das estratégias para os atingir;

- Aprendizagem:

- capacidade que o agente ostenta de adquirir conhecimento; a atualização da base de conhecimento é feita através da assimilação de padrões de comportamento ou de preferências;

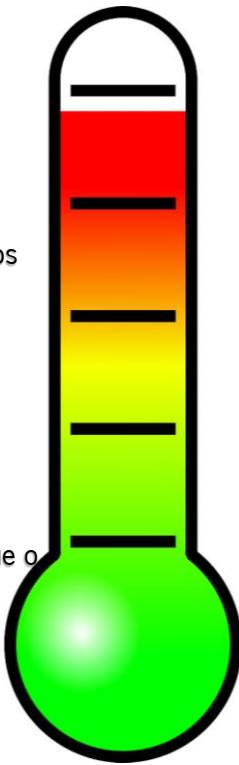
- Competência:

- um agente é competente quando conduz com sucesso e eficiência as tarefas de que é incumbido; a competência está normalmente relacionada com a confiança depositada no agente por terceiros;



Noção Forte de Agente Propriedades

- Veracidade:
 - Um agente exibe veracidade quando não fornece, de forma intencional, informação falsa;
- Racionalidade:
 - um agente racional não aceita realizar tarefas que avalie impossíveis de executar, contraditórias com os seus princípios ou quando não são compensados em termos do risco, custo ou esforço;
- Benevolência:
 - um agente benevolente adota como seus os objetivos de terceiros, desde que estes não entrem em conflito com os seus princípios de natureza ética e/ou deontológica, o que significa que não realizarão todas as tarefas que lhes sejam atribuídas;
- Emotividade:
 - certas características próprias do ser humano têm vindo a migrar e a estabelecer-se como parte constituinte dos agentes.
- Credibilidade:
 - capacidade do agente de criar uma suspensão de “descrença”, levando o utilizador a aceitar temporariamente que o agente está “vivo” ou que é um “personagem real”.
- Personalidade
 - capacidade do agente de se comportar de forma individual, o que torna o agente distingível de seus pares.



Em busca do “Santo Graal”

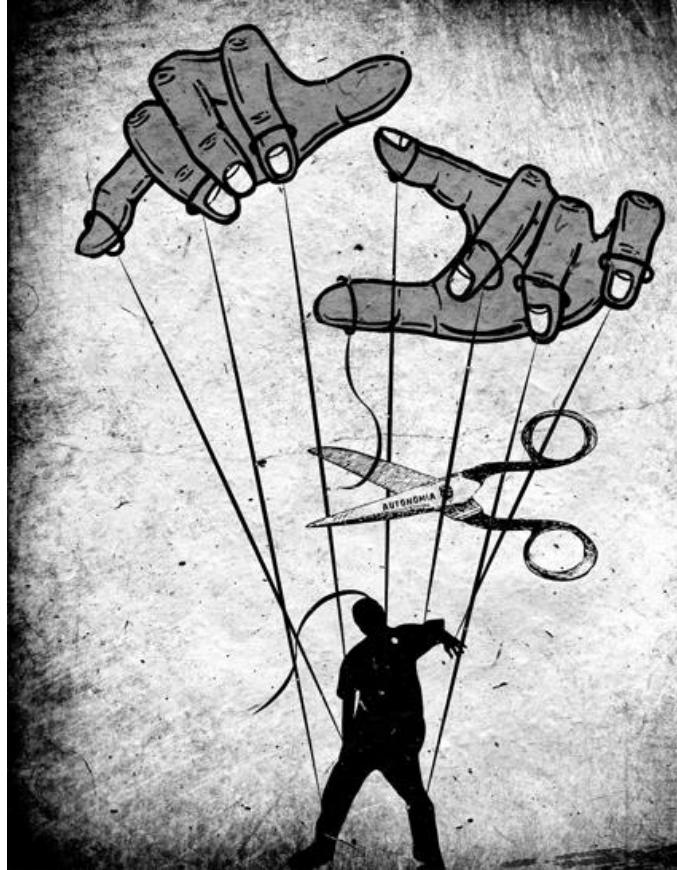
- A IA visa construir sistemas que possam (em última instância) comunicar em linguagem natural, reconhecer e compreender contextos, aprender, raciocinar, usar o “bom senso”, pensar criativamente, planear, agir, ...
 - Quando construímos um agente, queremos (simplesmente) um sistema que possa escolher a ação adequada a ser executada, normalmente num certo domínio.
 - Não temos que resolver todos os problemas de IA para construir um agente útil!

Problema da autonomia

- Podemos detetar diferentes graus de autonomia, dado que é um conceito relacional (social)
[Castelfranchi, 1995].
- Um agente só é autónomo em relação à influência dos outros agentes;
 - Qual a relação entre a autonomia de um agente e a sua capacidade de colaboração e aprendizagem?
- Tendo em conta que os agentes são situados, adaptando-se ao ambiente onde se inserem e a outros agentes, então a autonomia de um agente tem, necessariamente, de ser limitada!
- Como é que o comportamento humano responde e se adapta ao ambiente, e ao mesmo tempo, é independente dos estímulos externos?
[Descartes].

- Autonomia em relação ao contexto físico (ambiente):
 - Este tipo de autonomia entra em conflito com a necessidade de adaptação de um agente, ou seja, quanto mais autônomo o agente é do ambiente, menos necessidade tem de se adaptar a ele e de interagir com ele.
- Autonomia em relação ao contexto social (aos outros):
 - Este tipo de autonomia tem que ver com os objetivos dos agentes.
 - Pode-se dizer que o grau de autonomia é dado pelo número e valor dos objetivos que o agente tem que não consegue atingir sem a ajuda de outros;
 - Este tipo de autonomia pode ainda ser decomposto em dois tipos:
 - autonomia relativa aos meios de execução;
 - autonomia relativa aos objetivos.

Problema da autonomia



Problema da autonomia

- **Postulados** de um agente autónomo e social

[Castelfranchi, 1995].

- O agente:

- tem os seus próprios objetivos;
- é capaz de tomar decisões relativas a vários objetivos que podem estar em conflito;
- adota objetivos de outros agentes e do exterior (é influenciável):
 - adota como consequência de uma escolha;
 - adota se vê que a sua adoção o leva a atingir os seus próprios objetivos.



Agentes versus Objetos

- Agentes são **autónomos**, podendo recusar pedidos;
- Agentes controlam o seu **estado** e o seu **comportamento**;
- A comunicação nos agentes, o conceito de **mensagem** e de **linguagem** é definido ao nível da comunidade de agentes em que a ação se desenvolve.



- Objetos não são autónomos, não podendo recusar pedidos;
- Objetos controlam o seu estado mas não o seu comportamento;
- A comunicação nos objetos faz-se pela definição de métodos e da sua invocação;

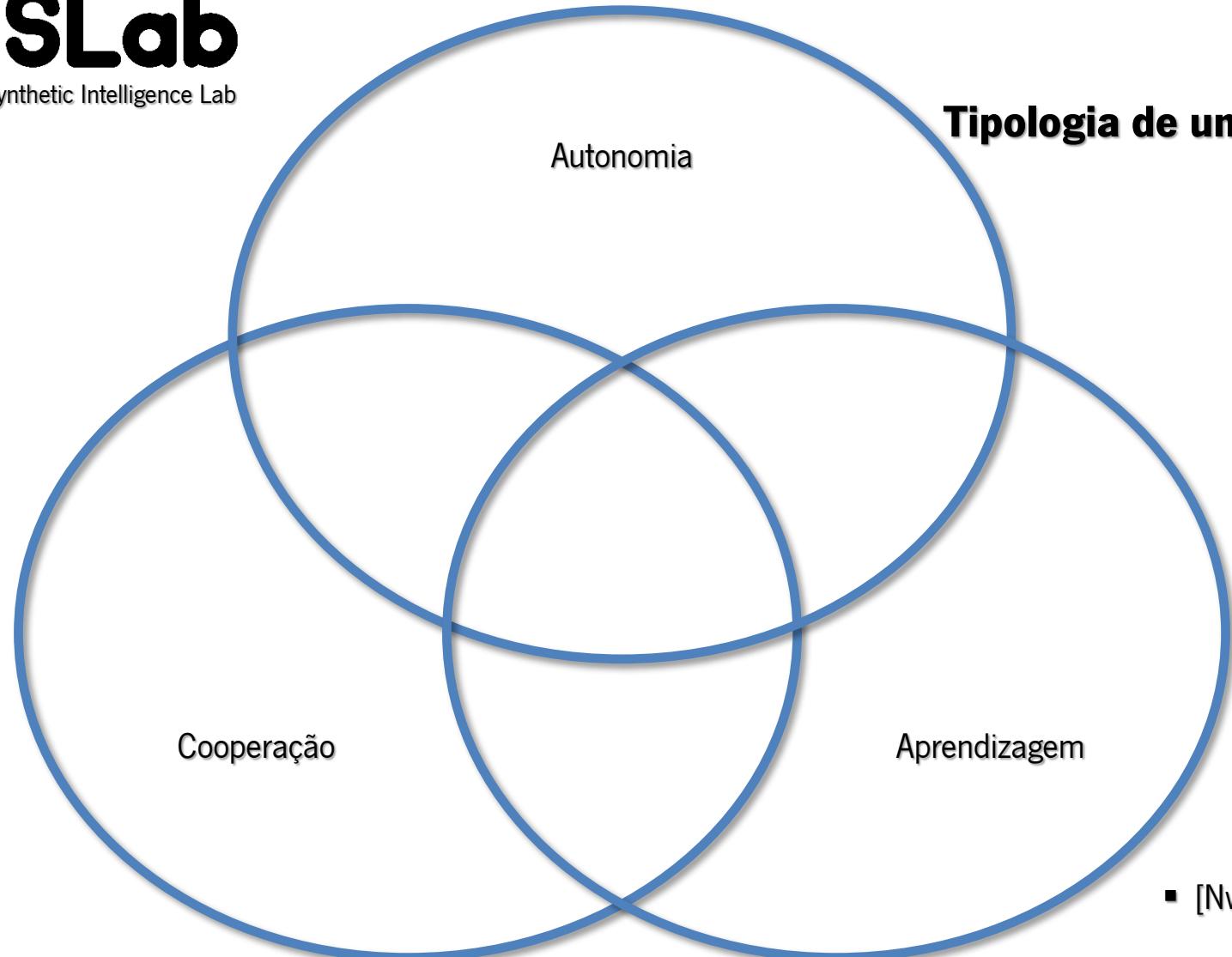


Agentes *versus* Sistemas Periciais

- Um Sistema Pericial tem como principal objetivo o apoio à resolução de problemas e à tomada de decisão num determinado domínio de conhecimento, funcionando como um consultor.
 - lidam com uma representação do universo de discurso, não o manipulando diretamente nem percecionando, no imediato, o resultado das suas ações;
 - destinam-se, essencialmente, a assistir peritos numa determinada área do conhecimento, enquanto que os agentes se envolvem na resolução de problemas;
 - aplicam-se, geralmente, a tarefas de alto nível, enquanto que os agentes de dedicam a tarefas comuns;
 - não têm capacidade de tomar iniciativa nem têm autonomia, respondendo de modo passivo e realizando sempre a mesma ação para os mesmos dados sensoriais.

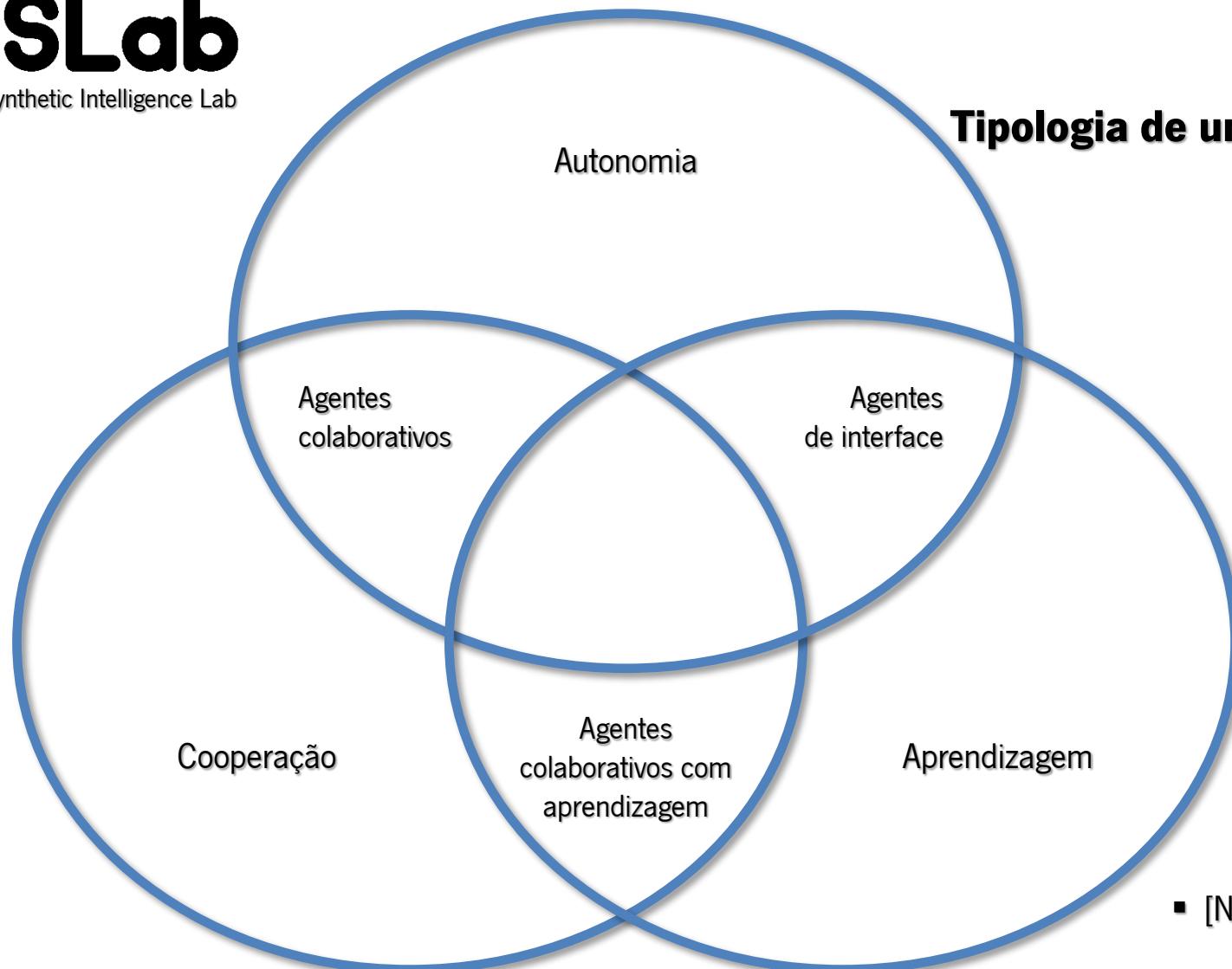


Tipologia de um Agente



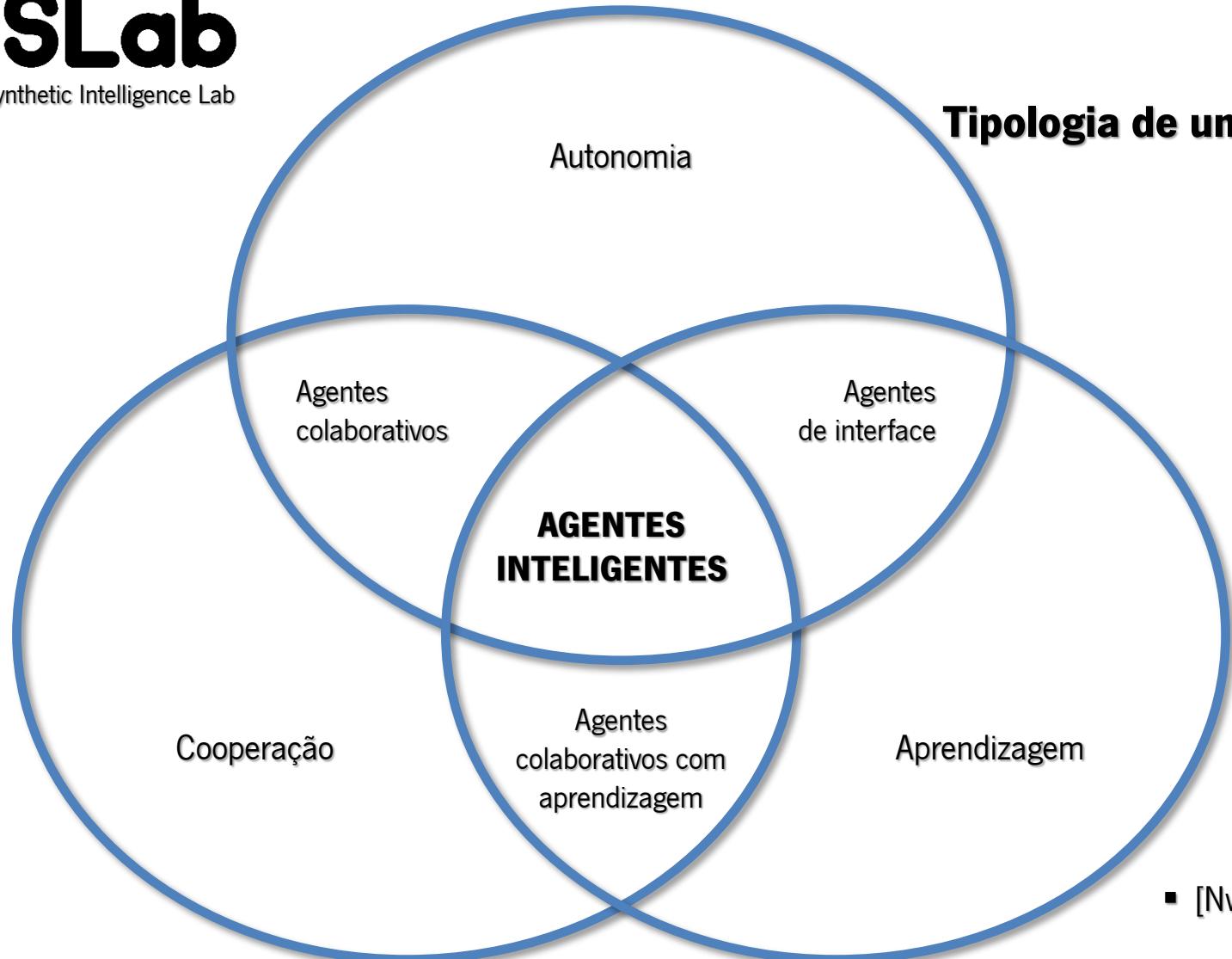
■ [Nwana, 1995]

Tipologia de um Agente



■ [Nwana, 1995]

Tipologia de um Agente



■ [Nwana, 1995]

- **Ambiente (o meio):**

como um conjunto finito de estados discretos $E = \{e, e', e'', \dots\}$

- **Ações**

- Um conjunto finito de ações disponibilizadas pelo agente $Ac = \{\alpha, \alpha', \alpha'', \dots\}$

- **Execução**

- sequência de estados e ações intercalados $r: e_0 \xrightarrow{\alpha_0} e_1 \xrightarrow{\alpha_1} e_2 \xrightarrow{\alpha_2} \dots \xrightarrow{\alpha_{n-1}} e_n$

Em que:

- R seja o conjunto de todas as possíveis sequências finitas (sobre E e Ac)
 - R^{Ac} é o subconjunto daqueles que terminam com uma ação
 - R^E seja o subconjunto daqueles que terminam com um estado de ambiente
 - Uma função de transformação de estado representa o comportamento do ambiente: $\tau: R^{Ac} \rightarrow \varphi(E)$
 - Se $\tau(r) = \emptyset$, então não há estados sucessores possíveis para r . Neste caso, dizemos que o sistema encerrou a sua execução

- Ambiente

Amb é um triplo Amb= $\langle E, e_0, \tau \rangle$ onde:

- E é um conjunto de estados do ambiente
- $e_0 \in E$ é o estado inicial
- τ é uma função de transformador de estado

- Agentes

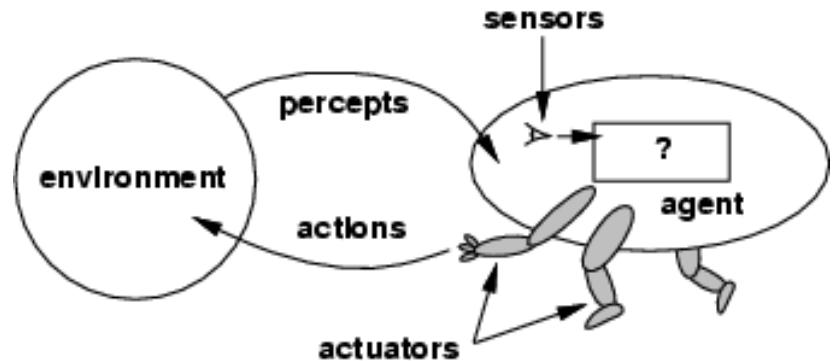
Agente é uma função que mapeia a execução de ações:

- $Ag: R^E \rightarrow Ac$
- Um agente toma uma decisão sobre qual ação executar com base no histórico do sistema.
- Seja AG o conjunto de todos os agentes
- R^E é o subconjunto de execuções que terminam com um estado de ambiente

- Um **sistema** é um **par** que contém um **agente (ou mais)** e um meio **Ambiente**
- Qualquer sistema terá associado a ele um conjunto de execuções possíveis;
 - denota o conjunto de execuções do agente Ag no ambiente A por R (Ag, A)
- Assuma-se que R (Ag, Amb) contém apenas execuções encerradas
- Formalmente: $(e_0, \alpha_0, e_1, \alpha_1, e_2, \dots)$ representa uma execução de um agente Ag no ambiente Amb= $\langle E, e_0, \tau \rangle$ se:
 - e_0 é o estado inicial de A;
 - $\alpha_0 = Ag(e_0)$
 - Para n>0
 - $e_0 \in \tau(e_0, \alpha_0, \dots, \alpha_{n-1})$ onde $\alpha_0 = Ag(e_0, \alpha_0, \dots, e_n)$

Tipos de Agentes

- Estruturas de Dados internas são atualizadas usando percepções e usadas para tomar a decisão das ações a executar (melhor ação)
- Tipos de Agentes (Russel e Norvig, 1995):
 - Agentes reflexos simples
 - Agentes com representação do mundo
 - Agentes baseados em objetivos
 - Agentes baseados em utilidade
 - Agentes com Aprendizagem

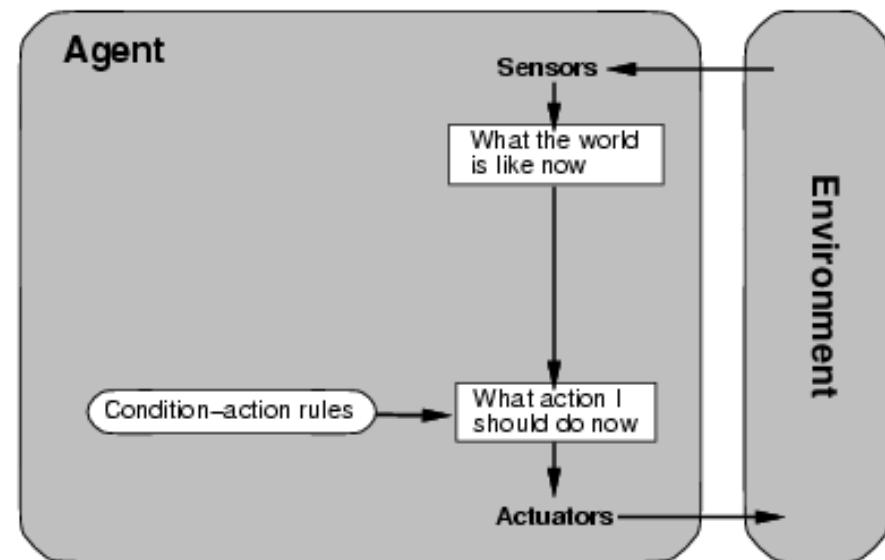


- Agentes que decidem o que fazer sem referência à histórico, baseiam a sua tomada de decisão nas informações do momento.
- Agentes Reativos
 - $\alpha: A \rightarrow \text{Ac}$
- Um termostato pode ser visto como um agente reativo

$$\sigma_e \begin{cases} off & \text{Se } e = \text{temperatura } 0k \\ On & \text{em qualquer outro caso} \end{cases}$$

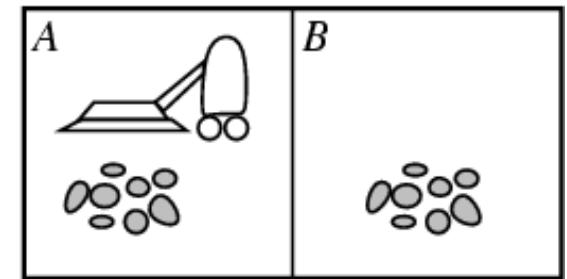
Agentes Reflexos Simples

- Baseados em tabelas de regras condição-acção (regras if-then)



Exemplo: O Mundo do Aspirador

- Percepções: local e conteúdo
 - Exemplo: [A, sujo]
- Ações: Esquerda, Direita, Aspirar, NoOp



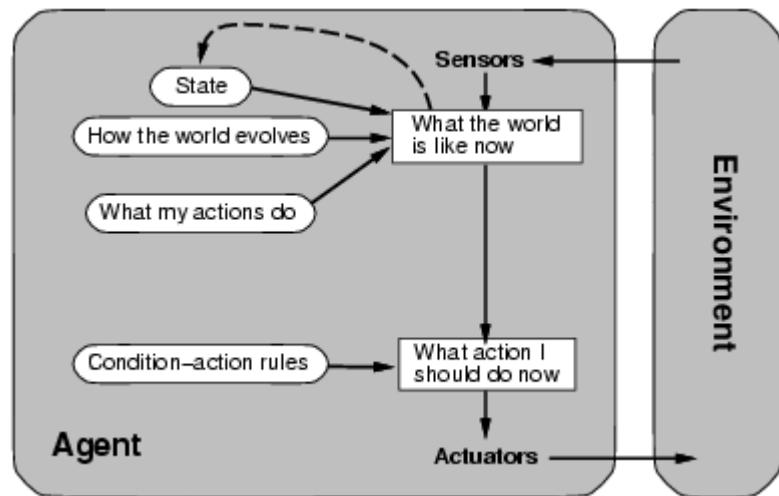
Exemplo: Agentes Reflexos Simples

- Regras condição-ação (regras se-então) fazem uma ligação direta entre a percepção atual e a ação;
- O agente funciona apenas se o ambiente for completamente observável e a decisão correta puder ser tomada com base apenas na percepção atual.

```
Função AGENTE-ASPIRADOR-REATIVO (posição, estado)
retorna uma ação
    se estado = Sujo então retorna Aspirar
    senão se posição = A então retorna Direita
    senão se posição = B então retorna Esquerda
```

Agentes com Representação do Mundo (com Estado)

- Mantém um estado interno (representação do mundo)



Agentes com Representação do Mundo

Função AGENTE-REATIVO-COM-Rep_Mundo(*perceção*) retorna uma ação
Variáveis estáticas:

estado, uma descrição do estado atual do mundo

regras, um conjunto de regras condição-ação

ação, a ação mais recente, inicialmente vazio

estado \leftarrow ATUALIZA-ESTADO(*estado*, *ação*, *percepção*)

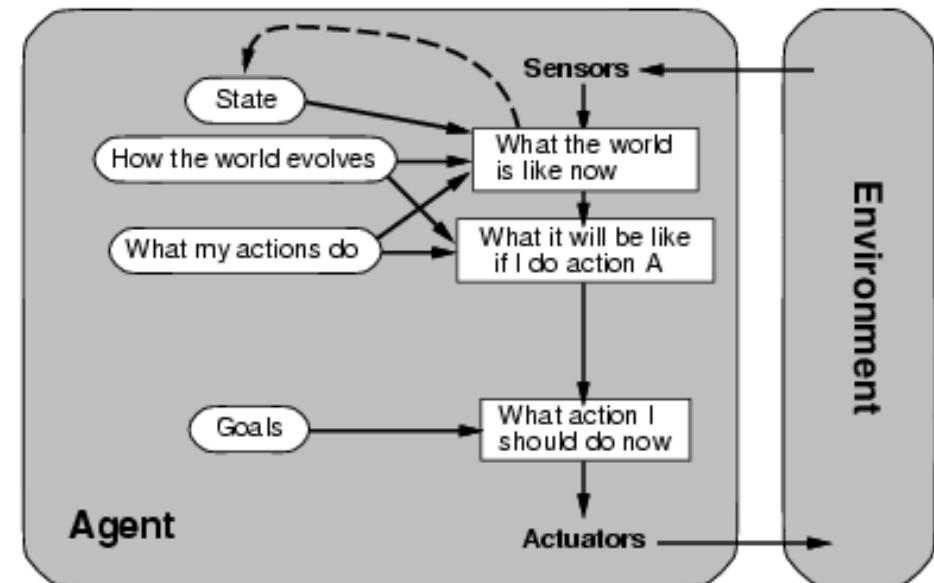
regra \leftarrow DETERMINA_REGRA(*estado*, *regras*)

ação \leftarrow AÇÃO-DA-REGRA(*Regra*)

retornar *ação*

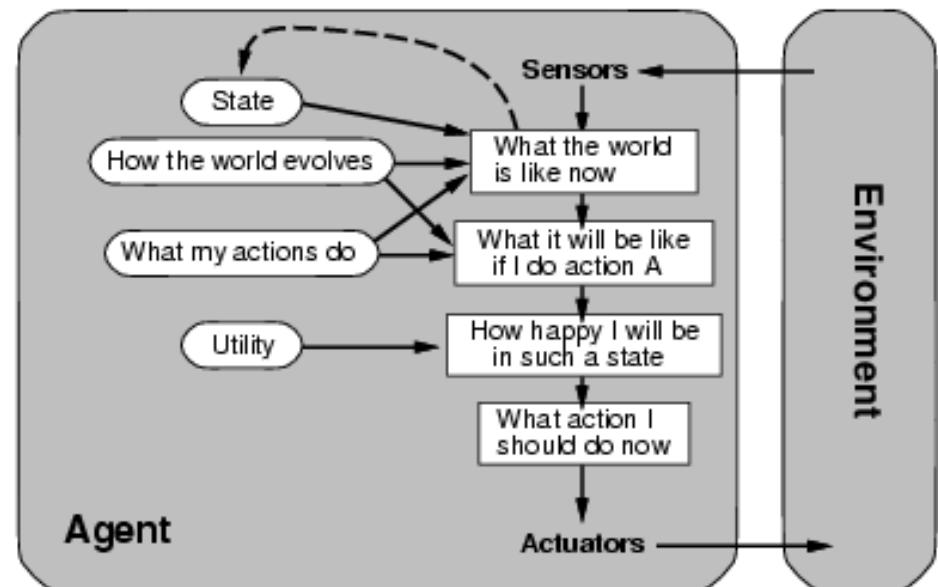
Agentes Baseado em Objetivos

- Descrição do estado do mundo e do objectivo a atingir
- Resolução de problemas por Pesquisa, Planeamento



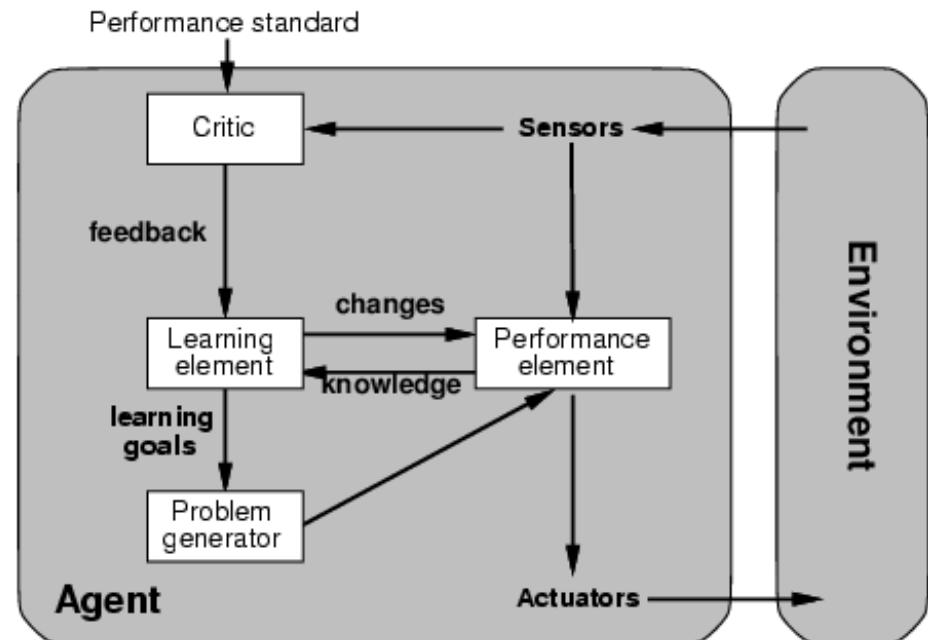
Agentes Baseados em Utilidade

- Utilidade: Espécie de grau de felicidade do agente!
- Mapeia o estado atual num valor!



Agentes com Aprendizagem

- Elemento de Aprendizagem e elemento de avaliação de desempenho



CONTROLEX: Agente para Controlar a Temperatura de uma Sala

Apresente um diagrama e o pseudo-código para um agente simples reflexo – CONTROLEX - para controlar a temperatura de uma sala. Suponha que dispõe das percepções T1 e T2 correspondentes à temperatura da sala e à temperatura exterior e as ações AQ ligar o aquecedor, NAQ – Desligar o aquecedor, AC – ligar o ar frio, NAC – Desligar o ar frio, AJ – abrir as janelas, NAJ – fechar as janelas.

Pretende-se que a temperatura da sala esteja entre os 22 e os 24 graus. Sempre que seja possível usar as janelas para controlar a temperatura (não desperdiçando energia), tal deve ser efetuado. Sempre que a temperatura esteja mais de 2 graus afastada da banda desejada (ou seja se a temperatura for inferior a 20 ou superior a 26 graus), deve-se fechar as janelas e em vez disso, usar o aquecedor ou ar frio para repor a temperatura dentro da banda desejada.

Como poderia construir um agente um pouco mais inteligente para este problema (que tipo de agente, percepções, estado do mundo, etc., usar)?

- Perceções:
 - T1 – Temperatura Interior
 - T2 – Temperatura Exterior
- Ações:
 - AQ ligar o aquecedor
 - NAQ – Desligar o aquecedor
 - AC – ligar o ar frio
 - NAC – Desligar o ar frio
 - AJ – abrir as janelas
 - NAJ – fechar as janelas
- Objetivo:
 - Manter a temperatura da sala entre os 22 e os 24 graus
- Agente mais inteligente:
 - Câmaras para analisar quantas e quais as pessoas no interior da sala
 - Ajuste da temperatura em função dos gostos das pessoas
 - Utilização de previsões meteorológicas da Internet

Exercício: Tópicos de Resolução

Interpretação da Perceção:

M_QUENTE = $T1 > 26$

QUENTE = $T1 > 24$ e $T1 \leq 26$

NORMAL = $T1 \geq 22$ e $T1 \leq 24$

FRIO = $T1 \geq 20$ e $T1 < 22$

M_FRIO = $T1 < 20$

FORA_UTIL = $T2 < 24$ e QUENTE ou
 $T2 > 22$ e FRIO

Regras Condição-Ação:

SE NORMAL

Então NAQ; NAC; NAJ

Se (QUENTE ou FRIO) e FORA_UTIL

Então NAQ; AJ; NAC

Se QUENTE e não(FORA_UTIL) ou M_QUENTE

Então NAQ; NAJ; AC

Se FRIO e não(FORA_UTIL) ou M_FRIO

Então AQ; NAJ; NAC

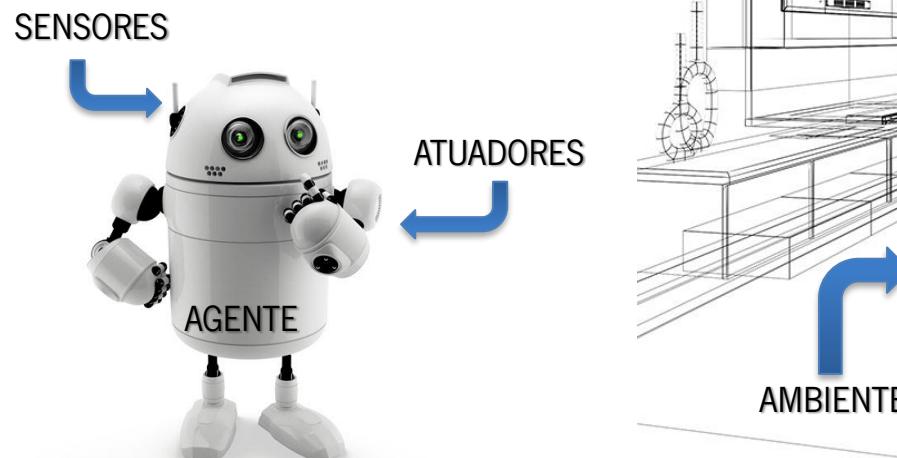
Abordagens baseadas na utilidade Dificuldades

- De onde vêm os números?
- Não pensamos em termos de números!
- Dificuldade na formulação de tarefas nestes termos ...

- Outra possível abordagem baseada na especificação de predicados (eg., Lógica)

Arquiteturas de Agentes Inteligentes

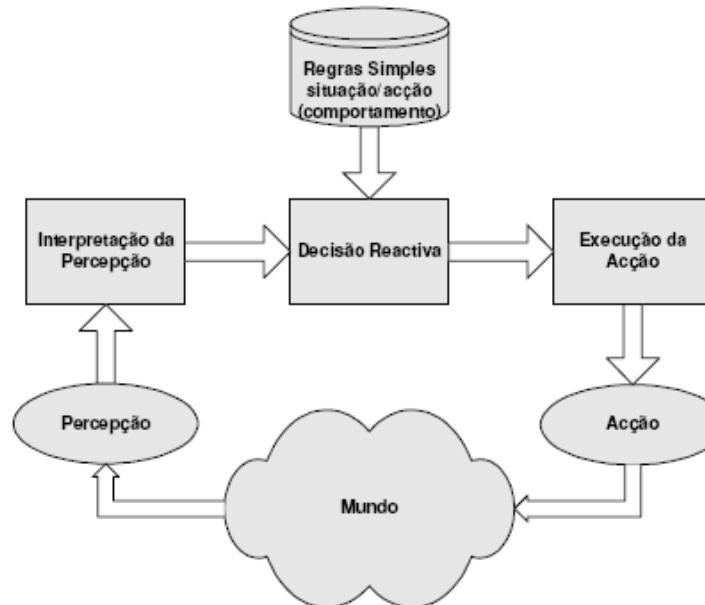
- As arquiteturas nas quais se baseiam o desenvolvimento de agentes inteligentes distinguem-se em:
 - Arquiteturas reativas
 - Arquiteturas deliberativas
 - Arquiteturas BDI – Beliefs, Desires and Intentions
 - Arquiteturas híbridas



- Um agente reativo não possui qualquer representação simbólica do universo de discurso, nem requer formas de raciocínio complexas.
- Um agente reativo comporta-se como um autómato envolvido pelo meio ambiente que o rodeia, agindo por reação a estímulos.
- Procura lidar com a percepção que tem do mundo através da receção de itens de informação do tipo atómico, que lhe são passados através de sensores.



- Um agente reativo não possui qualquer representação simbólica do universo de discurso, nem requer formas de raciocínio complexas.

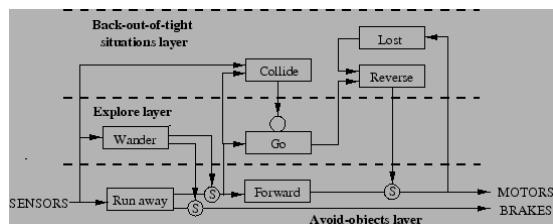


Luis Paulo Reis, 2003

Subsumption Architecture Rodney Brooks

“Elephants don't play chess” Rodney Brooks

- Recusa da utilização da “representação simbólica”
- Esta arquitetura de controle foi proposta em oposição as arquiteturas mais tradicionais da IA .
- Em vez de o comportamento de se guiar por representações simbólicas do mundo, esta arquitetura agrupa a informação do meio, via sensores, à escolha da ação.
 - Um Comportamento inteligente pode ser gerado sem representações explícitas do tipo que a IA simbólica propõe;
 - O comportamento inteligente pode ser gerado sem raciocínio abstrato explícito do tipo que a IA simbólica propõe;
 - A inteligência como uma propriedade emergente de sistemas complexos.
- O comportamento inteligente emerge como resultado da interação de um agente com seu ambiente.



Exemplo de um sistema de controle



Fonte: Wikipedia
Mars Exploration Rover

Arquiteturas Deliberativas

- Seguem a abordagem clássica da IA, onde os agentes atuam com pouca autonomia e possuem modelos simbólicos explícitos dos seus ambientes (universo de discurso), cujo estado de conhecimento pode ser modificado por alguma forma de raciocínio lógico-matemático.

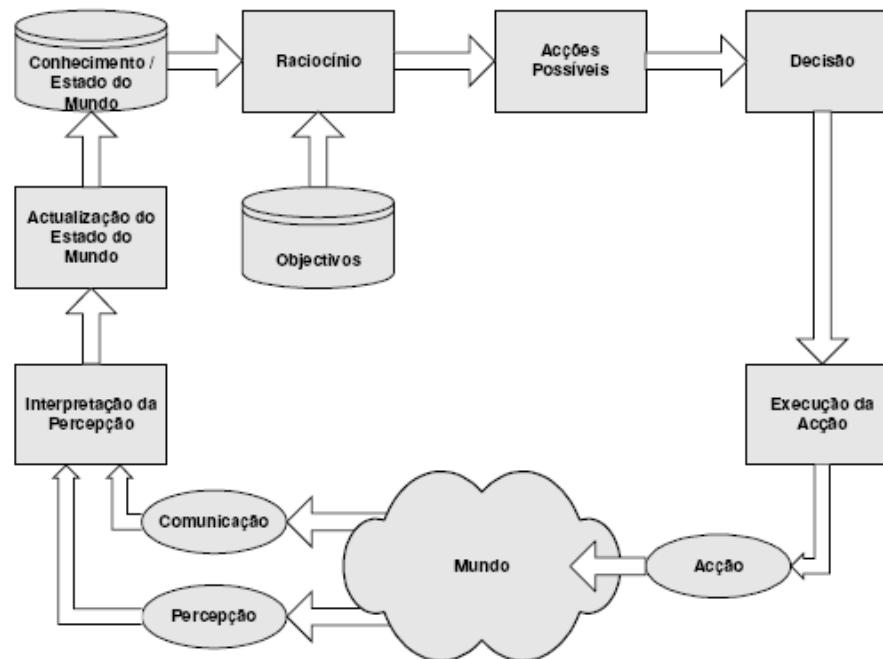
- Um problema de transposição e representação:
 - como traduzir o mundo real em termos de um programa em lógica?

- Um problema de raciocínio:
 - como levar os agentes a raciocinar?



Arquiteturas Deliberativas

- Seguem a abordagem clássica da IA, onde os agentes atuam com pouca autonomia e possuem modelos simbólicos explícitos dos seus ambientes (universo de discurso), cujo estado de conhecimento pode ser modificado por alguma forma de raciocínio lógico-matemático.



Luis Paulo Reis, 2003

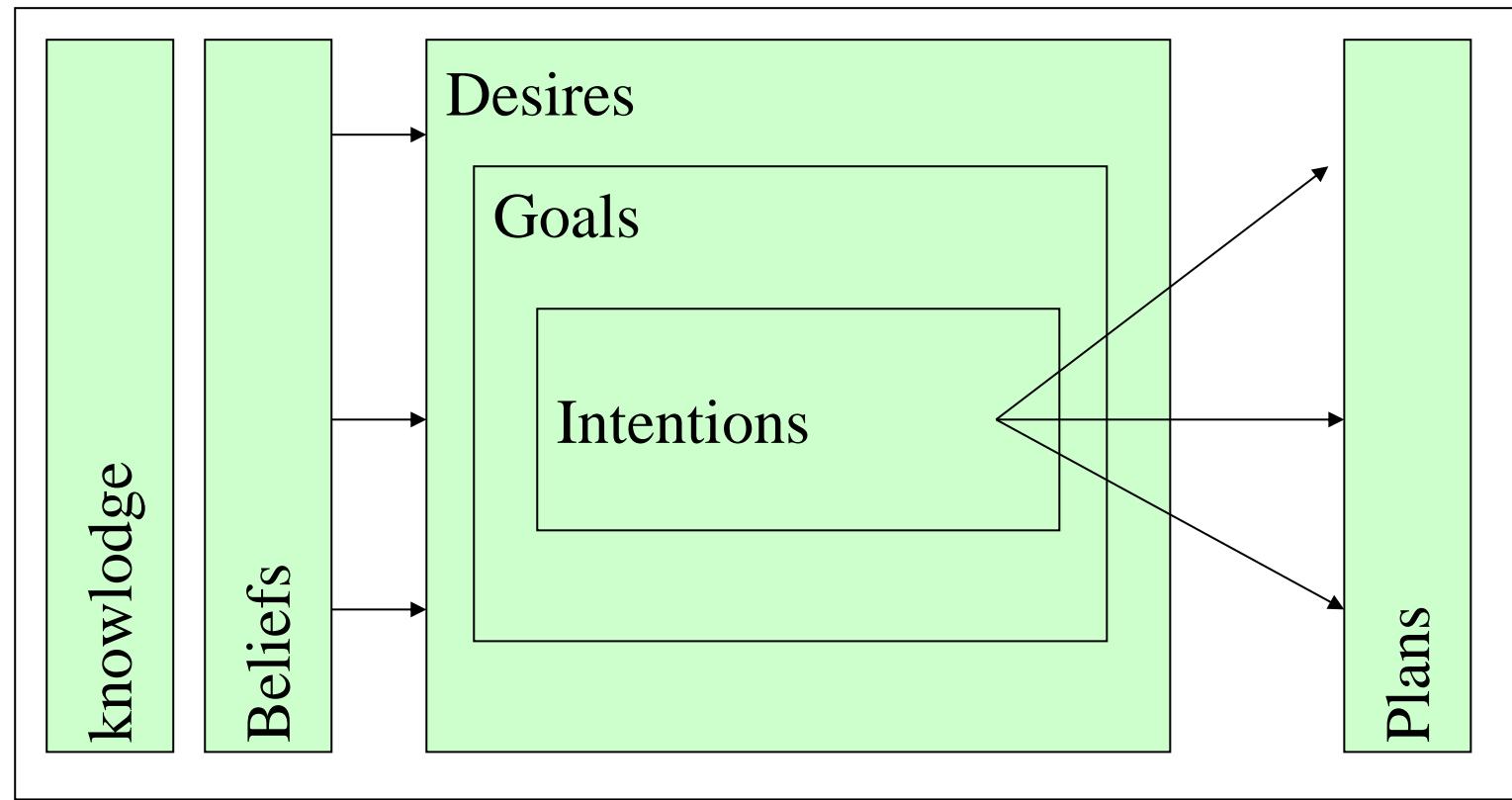
- Agentes como sistemas intencionais
 - Fornecemos ao agente uma especificação abstrata do sistema e deixamos o mecanismo de controle descobrir o que fazer, sabendo que ele atuará de acordo com alguma teoria embutida de atuação
- Arquitetura do Agente Deliberativo:
 - Contém um explícito modelo simbólico que representa o mundo;
 - Toma decisões por meio de raciocínio simbólico eg., sobre quais ações executar.
 - Atitudes informadas: Relacionadas com as informações que os agentes possuem sobre o meio ambiente (**crenças** e conhecimento)
 - Pró-atitudes: orienta as ações dos agentes (**desejos, intenções**, expectativas, comprometimento).

- "Practical reasoning is a matter of weighing conflicting considerations for and against competing options, where the relevant considerations are provided by what the agent desires/values/cares about and what the agent believes." Bratman,M.E., (1990), "What is Intention?" - MIT Media Lab
- Os Sistemas Intencionais são a base dos agentes deliberativos, que seguem a postura intencional por meio do raciocínio prático.
- Raciocínio Prático = Deliberação + Raciocínio Means-Ends
 - Deliberação: decidir qual o “estado” que queremos alcançar
 - Raciocínio means-ends: decidir como queremos alcançar esse “estado”.

Arquiteturas **BDI** ***Beliefs, Desires and Intention***

- Nesta abordagem à problemática que está subjacente à construção de arquiteturas para agentes, tem-se como objetivo obter aquela que melhor se adequa à descrição do estado interno de conhecimento de um agente, com base nos seus estados mentais, estando estes de que o agente fará uso para determinar o seu curso de ação:
 - As intenções dos agentes guiam as Ações;
 - As intenções baseiam-se em Crenças.



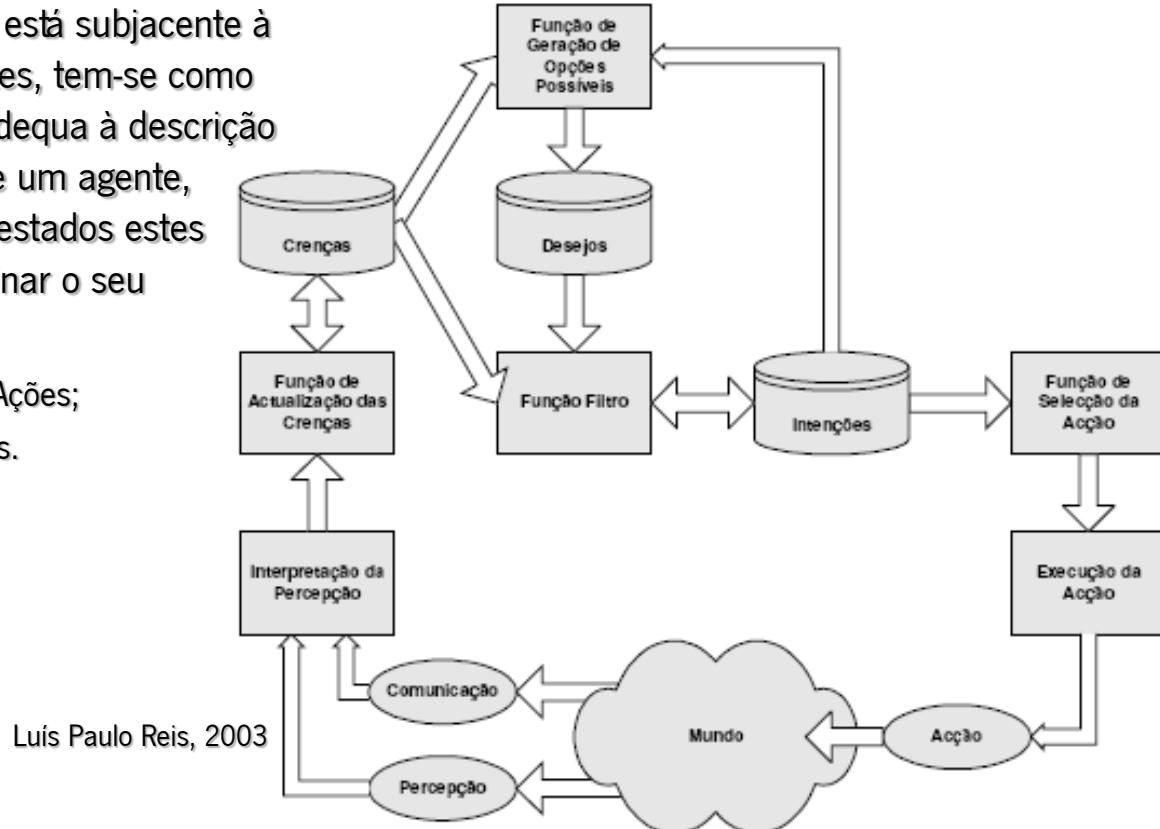


Arquiteturas **BDI** ***Beliefs, Desires and Intention***

- **Beliefs** (crenças) : Contém as visões fundamentais de um agente com relação ao seu ambiente.
- **Desires** (desejos): São derivados diretamente das crenças, contém os julgamentos de situações futuras do agente.
- **Goals** (metas): Subconjunto dos desejos contendo somente os estados que o agente realmente pode assumir.
- **Intentions** (intenções): Subconjunto das metas; se um agente decide seguir uma meta específica, esta meta torna-se uma Intenção.
- **Plans** (planos): Combina as intenções dentro de unidades consistentes.

Arquiteturas BDI *Beliefs, Desires and Intention*

- Nesta abordagem à problemática que está subjacente à construção de arquiteturas para agentes, tem-se como objetivo obter aquela que melhor se adequa à descrição do estado interno de conhecimento de um agente, com base nos seus estados mentais, estados estes de que o agente fará uso para determinar o seu curso de ação:
 - As intenções dos agentes guiam as Ações;
 - As intenções baseiam-se em Crenças.



Luis Paulo Reis, 2003

Subcomponentes do fluxo de controle geral da arquitetura BDI

- Revisão de crenças
 - Atualiza as crenças com informações “sensoriais” e crenças anteriores
- Geração opções
 - Use as crenças e intenções existentes para gerar um conjunto de alternativas/opções (= desejos)
- Filtragem
 - Escolhe entre alternativas concorrentes e compromete-se com sua realização
- Função de planeamento
 - Dadas as crenças e intenções atuais, geram um plano de ação
- Geração de ação: executa ações iterativamente segundo a sequência descrita no plano
- Eg., OASIS - System of air-traffic control (NASA)

Algumas limitações das arquiteturas reativas

- Os agentes sem modelos de ambiente deveriam ter informações suficientes disponíveis no “ambiente local”;
- Se as decisões são baseadas no “ambiente local”, como leva em consideração as informações não locais (tem de algum modo uma visão de "curto prazo");
- É muito difícil construir agentes reativos com capacidade de aprendizagem;
- É difícil projetar agentes com grande número de comportamentos (a dinâmica das interações torna-se muito complexa para ser entendida).

Algumas limitações das arquiteturas deliberativas

- Críticas aos Agentes Deliberativos:
 - A abordagem simbólica em si mesmo;
 - A ligação ao mundo real
 - A velocidade do processo de tomada de decisão num mundo real.

Arquiteturas Híbridas

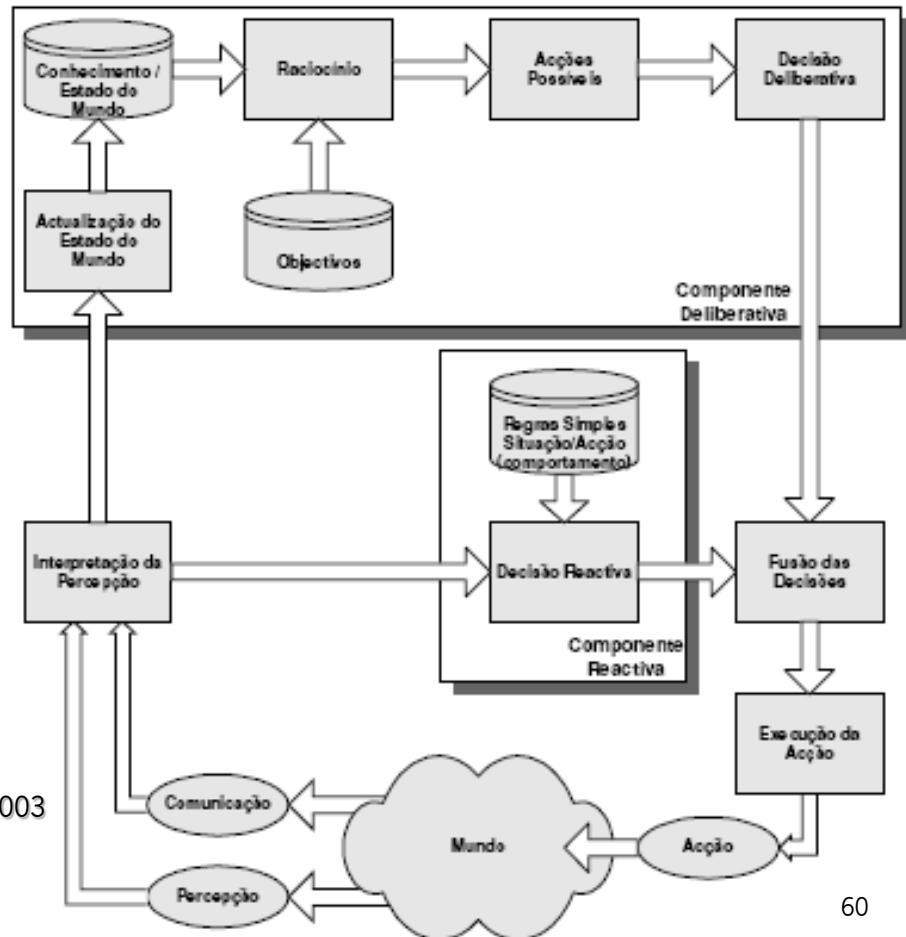
- Uma abordagem óbvia é construir um agente de dois (ou mais) subsistemas:
 - um deliberativo, contendo um modelo de mundo simbólico, que desenvolve planos e toma decisões;
 - um reativo, que é capaz de reagir a eventos sem raciocínio complexo.
- Agentes híbridos combinam as características deliberativa e reativa.
- A ideia principal passa por categorizar as funcionalidades do agente em camadas dispostas hierarquicamente.
- À camada reativa é atribuída alguma forma de prioridade sobre a deliberativa, de tal modo que se aproveite uma das suas características mais importantes que é a de dar resposta rápida a eventos detetados no ambiente.



- Agentes híbridos combinam as características deliberativa e reativa.
- A ideia principal passa por categorizar as funcionalidades do agente em camadas dispostas hierarquicamente.
- À camada reativa é atribuída alguma forma de prioridade sobre a deliberativa, de tal modo que se aproveite uma das suas características mais importantes que é a de dar resposta rápida a eventos detetados no ambiente.

Luis Paulo Reis, 2003

Arquiteturas Híbridas



The DARPA Grand Challenge (2003)

Veículos autónomos ao longo de 212 km do deserto de Nevada:
Stanley vencedor contra CMU.

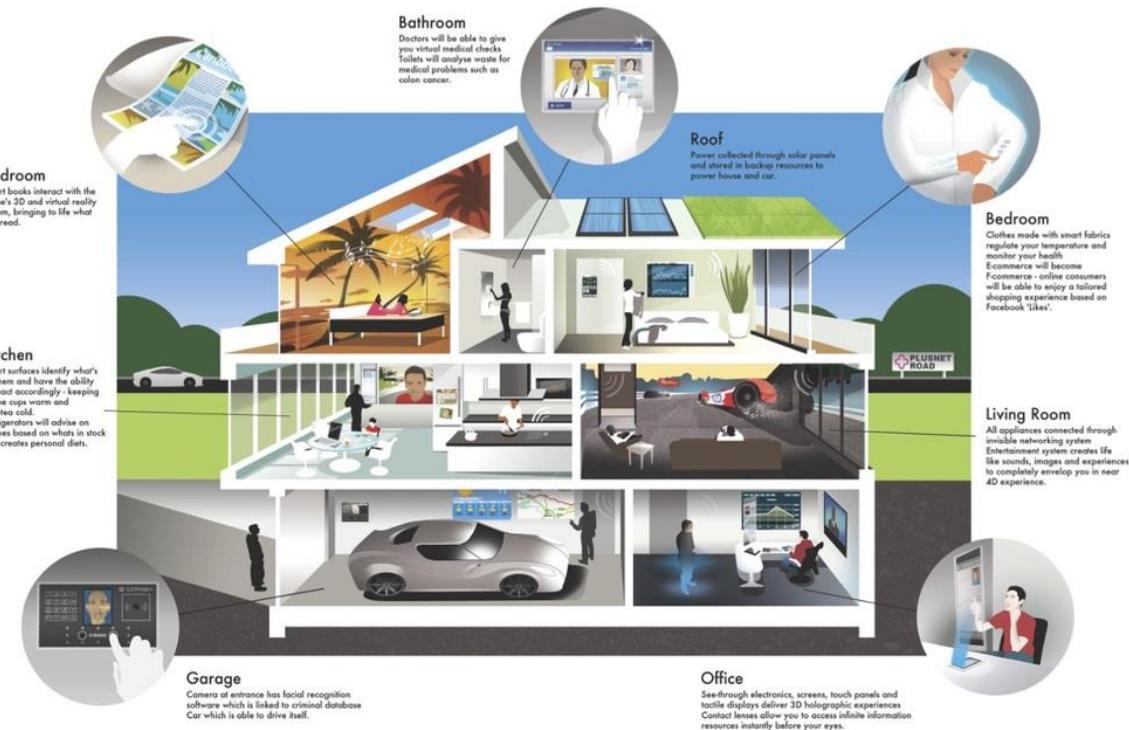


Sebastian Thrun, da Universidade de Stanford, bateu as 3 equipas da CMU, que tinha por sua vez batido a de Rodney Brooks do MIT no Concurso da NASA para a missão em Marte, (STANLEY, 2005).



- Internet
- Comércio Eletrónico
- Redes Colaborativas
- Simulação Social
- Aplicações Industriais
- Ambientes Inteligentes
- Tomada de Decisão em Grupo
- Direito

Áreas de Aplicação

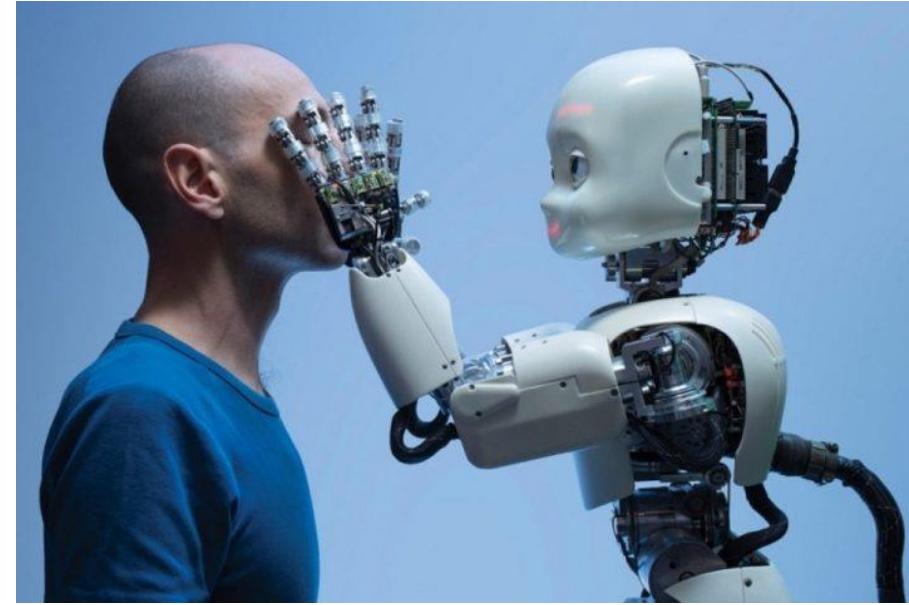


Conclusões

- Um agente é uma entidade computacional com um comportamento **autônomo** que lhe permite decidir sobre as suas próprias ações;
- Os agentes têm uma **existência própria**, independente da existência de outros agentes;
- Cada agente possui um conjunto de características comportamentais que definem a sua **competência**, um conjunto de objetivos, e a **autonomia** necessária para utilizar as suas capacidades comportamentais a fim de alcançar os seus objetivos;
- A **decisão** de qual a ação a levar a cabo é determinada pelo agente, tendo em consideração as mudanças que ocorrem no ambiente em que atua e o desejo de alcançar os seus objetivos.



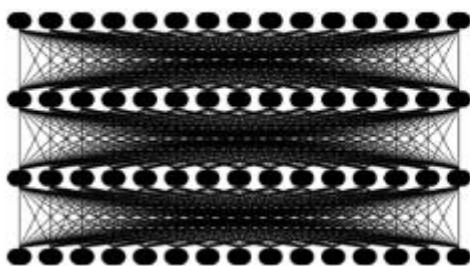
- É raro os agentes existirem isolados;
- É necessário trabalhar em conjunto - Sistemas Multiagentes;
- É necessário coordenar - trabalhar em conjunto de forma harmoniosa;
- Ubiquidade - Integração de microprocessadores nos objetos do dia a dia (computação embebida);
- Conectividade - isolado *versus* redes distribuídas;
- Inteligência - tarefas cada vez mais complexas;
- Delegação - delegação de tarefas críticas
(p.ex., piloto automático);
- Servir o humano - uso de metáforas “humanas”
em vez de interação ao nível da máquina.



Próximos passos

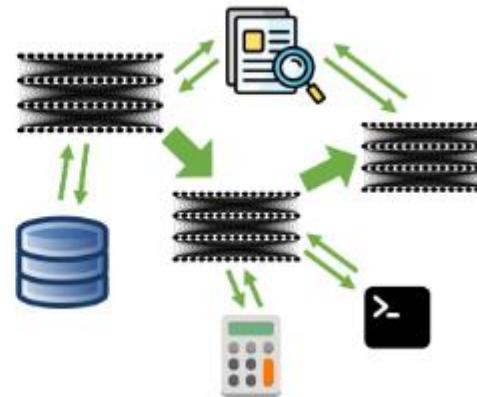
IA Generativa

Geração conteúdos como texto e imagens



Agentic

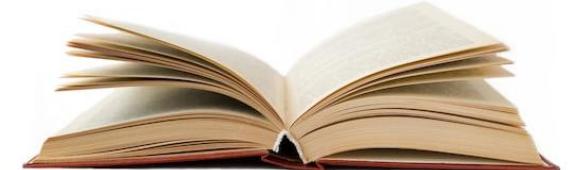
Executar tarefas complexas



Source: Zaharia et al. 2024. The Shift from Models to Compound AI Systems,
<https://bair.berkeley.edu/blog/2024/02/18/compound-ai-systems/>

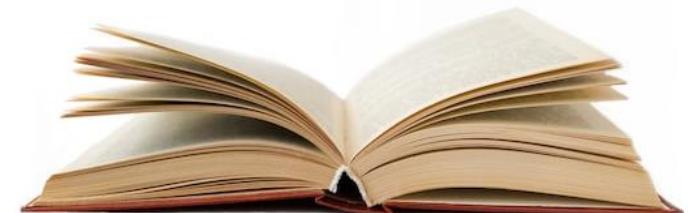
Referências

- [Russell & Norvig, 1995] Stuart J. Russell, Peter Norvig, "Artificial Intelligence – A Modern Approach", Prentice Hall International Inc., EUA, 1995.
- [Huhns & Singh, 1998] Michael N. Huhns, Munindar P. Singh, "Agents and Multiagent Systems: Themes, Approaches and Challenges", Huhns, Singh (editors), Readings in Agents, pp. 1 23, Morgan Kaufmann Publishers, San Francisco, USA,
- [Maes, 1990] Pattie Maes, "Situated Agents Can Have Goals", Designing Autonomous Agents, Maes (editor), MIT Press.
- [Wooldridge, 1999] Michael J. Wooldridge, "Intelligent Agents", in Multiagent Systems – A Modern Approach to Distributed Artificial Intelligence, Weiß (editor), chapter 1, pp. 27 77, MIT Press, Cambridge, USA.
- [Wooldridge & Jennings, 1995] Michael J. Wooldridge, Nicholas R. Jennings, "Intelligent Agents: Theory and Practice", Knowledge Engineering Review, 10 (2), pp. 115 152, 1995.
- [Nwana, 1996] Hyacinth S. Nwana, "Software Agents: An Overview", Knowledge Engineering Review, 11 (3), pp. 1-40.



Referências

- [Rao & Georgeff, 1995] Anand S. Rao, Michael P. Georgeff, “BDI Agents: from Theory to Practice”, Proceedings of the First International Conference on Multi Agent Systems – ICMAS’95, São Francisco, EUA.
- [Reis, 2003] Luís Paulo Reis, “Coordenação em Sistemas Multi-Agente: Aplicações na Gestão Universitária e Futebol Robótico”, Tese de Doutoramento, Faculdade de Engenharia da Universidade do Porto.
- [d'Inverno, Luck, 2003] d'Inverno M., Luck M., “Understanding Agent Systems”, Springer, ISBN: 978-3540407003.
- [Maes P., 1990] Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back, MIT Press.
- [Muller J., 1996] The Design of Intelligent Agents: a Layered Approach, Springer.
- [Zaharia et al. 2024] The Shift from Models to Compound AI Systems, Berkeley Artificial Intelligence Research, <https://bair.berkeley.edu/blog/2024/02/18/compound-ai-systems/>





Universidade do Minho

Escola de Engenharia

Departamento de Informática

Agentes

Mestrado Integrado em Engenharia Informática

Mestrado em Engenharia Informática

Perfil SI :: Agentes e Sistemas Multiagente