



FUNDAÇÃO EDSON QUEIROZ
UNIVERSIDADE DE FORTALEZA
ENSINANDO E APRENDENDO

Inteligência Artificial



Unidade II – Agentes Inteligentes

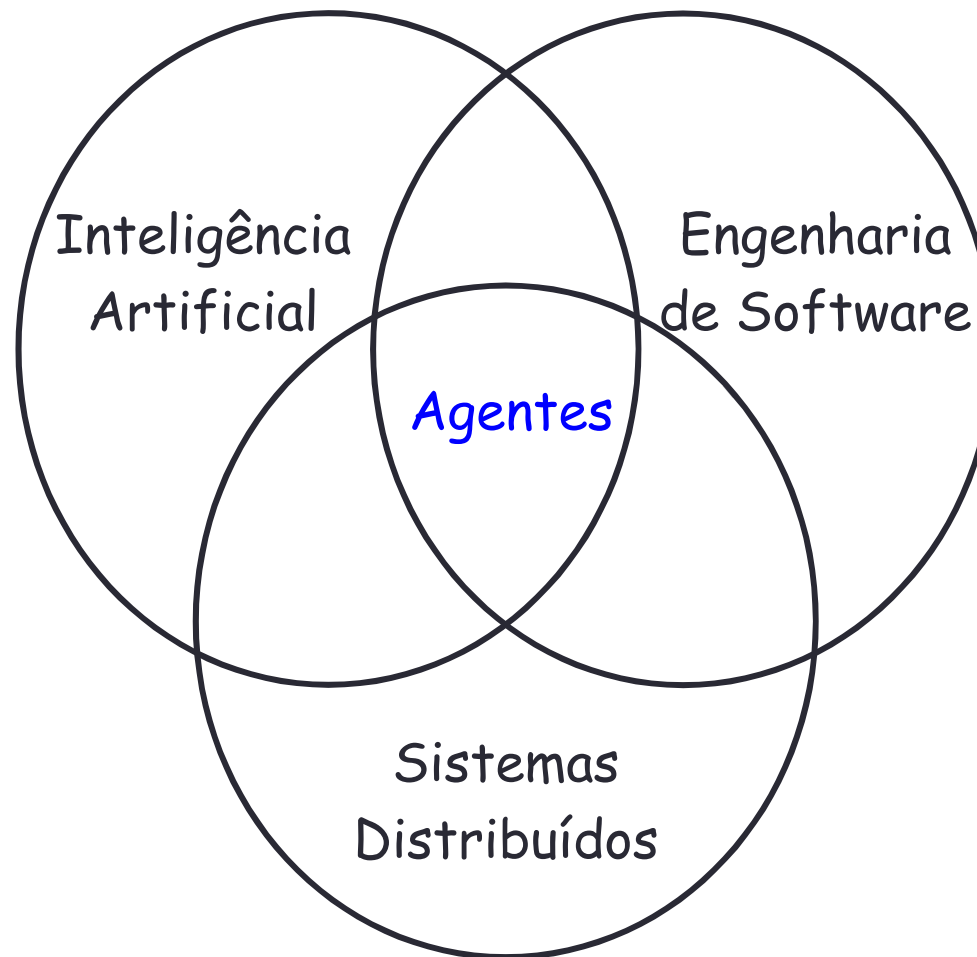
Profa. Vlândia Pinheiro

Adaptação de conteúdo dos Prof. André Coelho e Prof. Vasco Furtado

Quais as razões do interesse em agentes?

- ➡ • **Racionalidade vista como um referencial mais objetivo para definição de inteligência**
 - Pode-se lançar mão de todo um arcabouço teórico provido por estudos nas áreas de economia, tomada de decisão, psicologia social, política, etc.
 - Fornece metodologias de desenvolvimento de sistemas inteligentes estendendo as técnicas tradicionais de engenharia de software
 - Ajuda a embutir a IA em sistemas legados
 - Permite tratar melhor a interação da máquina com o seu ambiente
 - Possibilita extensão natural para a IA distribuída
- ➡ • **Fornece visão unificadora das várias subáreas da IA**

Agentes no contexto da computação



Terminologia usada para Agentes

- ↖ *agentes de software*
- ↖ *agentes inteligentes*
- ↖ *agentes artificiais*
- ↖ *interfaces inteligentes*
- ↖ *interfaces adaptativas*
- ↖ *shopbots, knowbots, softbots, userbots, taskbots,*
- ↖ *Agentes pessoais*
- ↖ *network agents.*

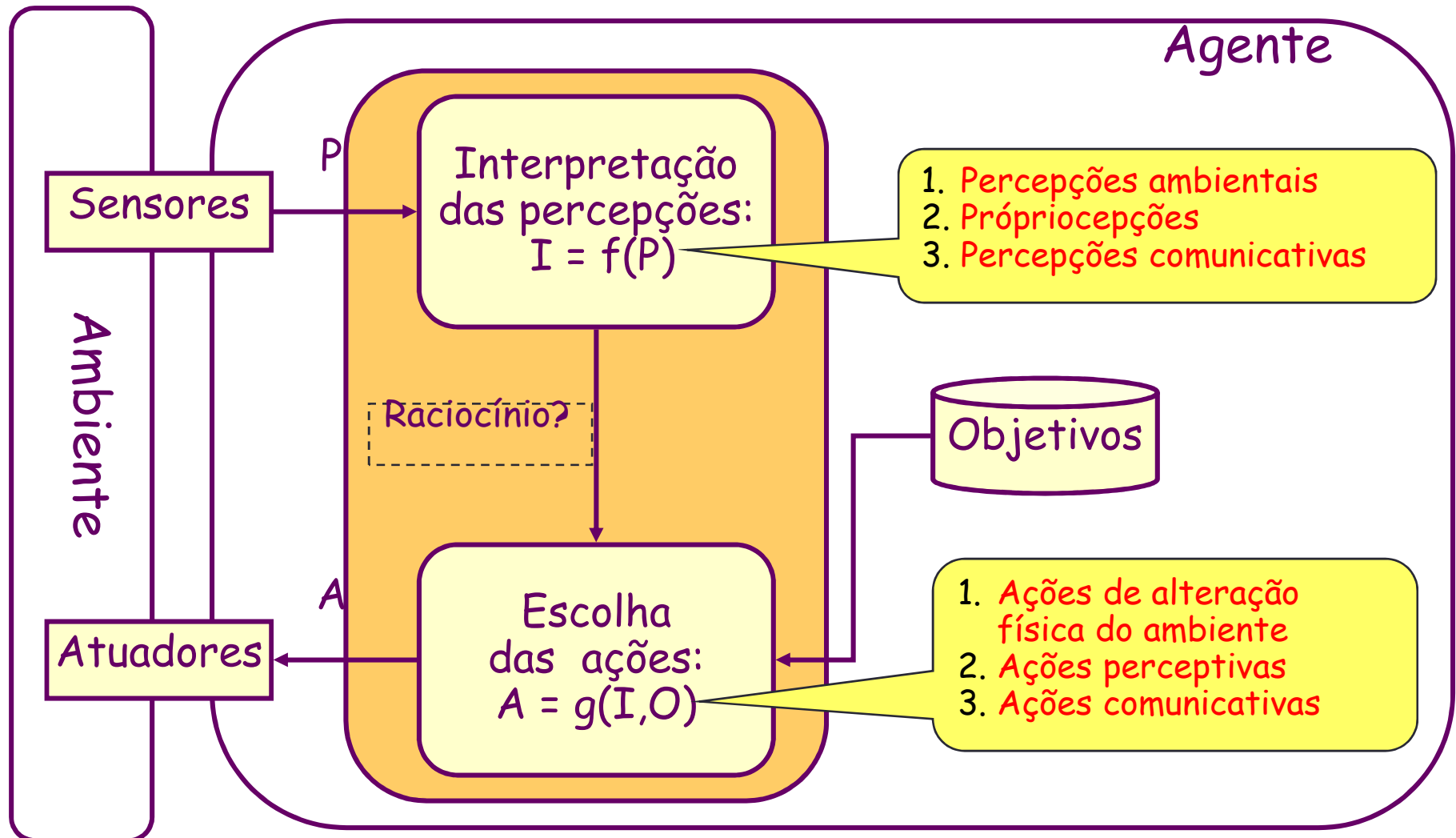
Definição

- ↖ “Um agente é qualquer coisa que possa perceber um ambiente por meio de sensores e atuar no mesmo por meio de atuadores”. Russell & Norvig (1995)
- ↖ “Agentes autônomos são sistemas computacionais que habitam em um ambiente complexo e dinâmico, sensoreiam e atuam autonomamente neste ambiente, realizando, desta maneira, uma série de metas e tarefas para as quais foram projetados” (Pattie Maes – MIT)

Definição técnica de agente

- Qualquer entidade (humano, robô, software):
 - Imersa ou situada em um **ambiente** (físico, virtual/simulado);
 - Que percebe esse ambiente via **sensores** (olhos, câmera, *socket*);
 - Que age sobre esse ambiente via **atuadores** (mãos, roda, *socket*);
 - Que possui **objetivos próprios**, i.e., alcançar estados do ambiente mais preferidos (sendo estes explícitos ou implícitos); e
 - Que escolhe suas **ações** em função das suas percepções para atingir seus objetivos.
- Ciclo de processamento típico de um agente:
 - Percebe o ambiente P
 - Interpreta suas percepções $I = f(P)$ **função de interpretação**
 - Escolhe suas ações $A = g(I, O)$ para atingir objetivos O **função de tomada de decisão**
 - Executa ações A

Arquitetura genérica



Objeto

×

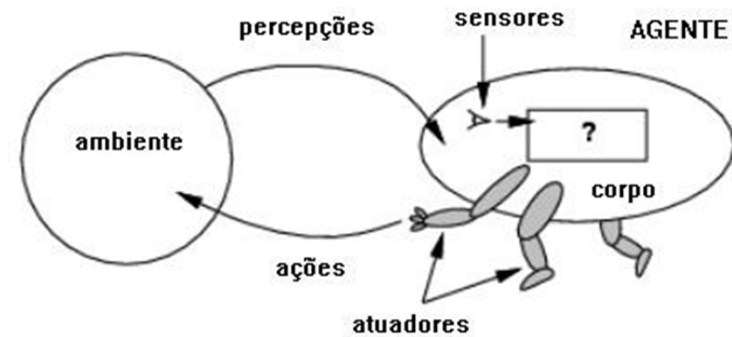
Agente

- Sem objetivo próprio
- Sem autonomia de decisão
 - Executa *apenas* quando invocado por outros objetos
 - Executa *sempre* que invocado por outros objetos
- Entrada e saída: parâmetros e resultado de métodos
- Descontinuidade temporal: ativo apenas durante invocação dos seus métodos
- Encapsulamento

- Intencionalidade
 - Apresenta objetivos próprios (mesmo que implicitamente), além de dados e métodos
- Com autonomia de decisão/ação
 - Pode iniciar ação por sua própria iniciativa para satisfazer seus objetivos (**pró-atividade**)
 - Pode negar a atender um pedido de ação de outro agente (negociação)
- Entrada e saída mais complexas: percepções e ações
- Continuidade temporal: sempre monitorando o ambiente
- Granularidade maior:
 - Encapsula código do tamanho de um pacote ou componente de software
 - Composto de vários objetos quando implementado no paradigma de OO

Terminologia básica

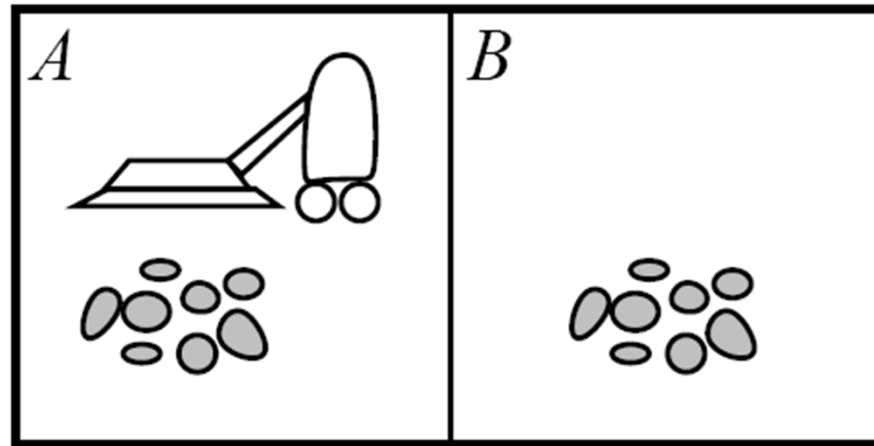
- **Agência** (*agency* ou *agenthood*)
 - Propriedade de ser agente
- **Agenda**
 - Lista de tarefas a cumprir
- **Sensores**
 - Elementos/dispositivos de **sensação** e **percepção**:
 - câmeras, microfones, termômetros, teclados, transdutores, *sockets*...
- **Atuadores**
 - Elementos/dispositivos de **atuação**:
 - garras, braços articulados, acionadores, propulsores, vídeo, impressora...
- **Ambiente** e **estado**



Terminologia básica

- **Percepção**
 - Entradas sensório-perceptivas em qualquer instante
- **Seqüência perceptiva**
 - História completa de tudo que o agente já percebeu
 - A escolha da próxima ação do agente pode depender da seqüência inteira de percepções observadas até o momento
- **Função de agente**
 - Descrição externa abstrata (matemática) acerca do que o agente deve fazer em todas as circunstâncias
 - Mapeamento: seqüência perceptiva → ação
- **Programa de agente**
 - Implementação concreta de uma função de agente
 - Relacionada à arquitetura física do agente

Exemplo: Mundo do aspirador de pó



- **Ambiente:** quadrados *A* e *B*
- **Objetivo:** deixar o chão limpo em cada turno de trabalho
- **Percepções:** de localização e de existência de sujeira
- **Ações:** mover-se p/ esquerda e direita, não fazer nada, despejar ou aspirar
- **Função de agente:** se o quadrado atual estiver sujo, então aspirar; senão, mover-se para o outro quadrado

Exemplo: Mundo do aspirador de pó

Seqüência perceptiva	Ação
[A, limpo]	Direita
[A, sujo]	Aspirar
[B, limpo]	Esquerda
[B, sujo]	Aspirar
[A, limpo], [A, limpo]	Direita
[A, limpo], [A, sujo]	Aspirar
...	...
[A, limpo], [A, limpo], [A, limpo]	Direita
[A, limpo], [A, limpo], [A, sujo]	Aspirar
...	...

A função de agente especificada por esta tabela é uma dentre várias.

➡ **Pergunta-chave:** Será que ela especifica um agente inteligente???

Racionalidade

- A definição do que é racional (inteligente) depende de 4 fatores:
 - A **medida de desempenho**, que define o critério de sucesso do agente;
 - O **conhecimento anterior** (crenças) que o agente tem sobre o ambiente (domínio do problema);
 - O **repertório de ações** que o agente pode executar; e
 - A **sequência de percepções** do agente até o momento.

Racionalidade

- Exemplo para o mundo do aspirador de pó
 - ➔ **Se** as circunstâncias forem estas:
 - **Medida de desempenho**: um ponto por quadrado limpo ao final de cada turno de trabalho
 - **Conhecimento anterior**: a geografia do ambiente é conhecida *a priori*, mas a distribuição de sujeira e posição inicial do agente não são conhecidas
 - **Repertório de ações**: as únicas ações disponíveis são *Esquerda*, *Direita*, *Despejar*, *Aspirar* e *NoOp*
 - **Seqüência de percepções**: o agente percebe corretamente sua posição e se ela contém sujeira
 - ➔ **Então** a função de agente anterior leva à especificação de um agente realmente **racional**

Agente racional

(McCarthy & Hayes 69, Newell 81)

- Definição informal: É aquele que faz a melhor coisa ***possível***
- Definição mais criteriosa:
 - “Para cada sequência de percepções possível, um agente racional deve selecionar uma ação que ele **espera** venha a maximizar sua medida de desempenho, dada a evidência fornecida pela sequência de percepções e por qualquer conhecimento interno do agente” (Russell & Norvig)
 - Ação certa fará o agente obter maior sucesso

Medidas de desempenho

- Critério(s) que define(m) o grau de sucesso do comportamento de um agente na realização de uma dada tarefa
 - Medida objetiva a ser imposta de fora para dentro
- No entanto, a escolha de uma medida de desempenho nem sempre é trivial
 - Má escolha de tais critérios pode acarretar em comportamento indesejado
 - Compromissos entre múltiplos critérios conflitantes
 - Nem sempre é factível se saber *a priori* qual o melhor momento para se avaliar o desempenho

Medidas de desempenho

- Exemplo de medida de desempenho inadequada:
 - “Medir o desempenho do aspirador de acordo com a quantidade de sujeira limpa em um certo tempo”
 - O agente poderia maximizar esse critério limpando a sujeira e, em seguida, despejando tudo no chão, depois limpando novamente e assim por diante → seria racional!
- Exemplo de medida de desempenho + adequada:
 - “Recompensar o agente por deixar o chão limpo”



• Regra geral:

- Deve-se projetar medidas de desempenho de acordo com o resultado realmente desejado no ambiente, em vez de criá-las de acordo com o comportamento esperado do agente

Medidas de desempenho

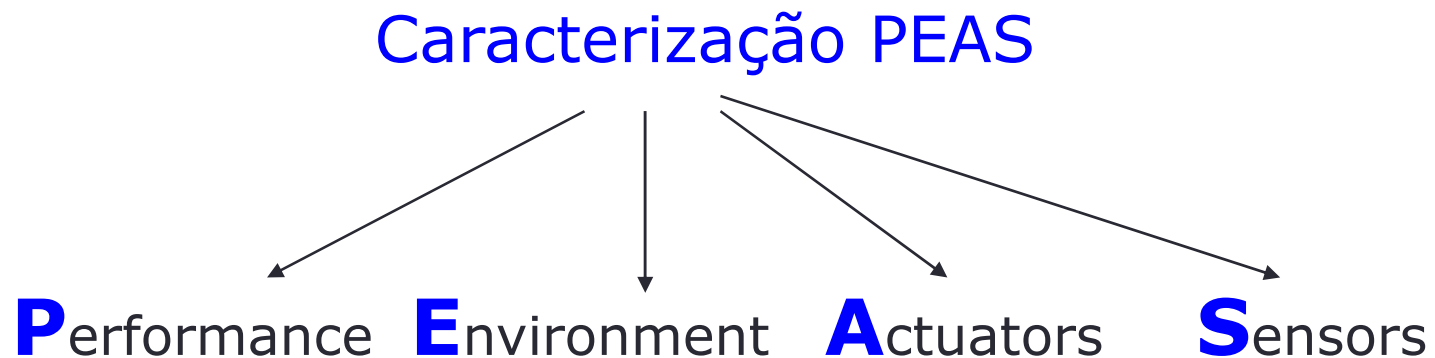
- E se a medida de desempenho do aspirador de pó anterior incluísse uma penalidade de um ponto para cada movimento à esquerda ou direita, o agente anterior continuaria a ser racional?
 - Não! A melhor atitude seria não fazer nada, caso percebesse que os quadrados estivessem limpos
- Quais os critérios a serem adotados para avaliar os seguintes agentes?
 - Provedor de teoremas, filtragem de *e-mails*, policial de trânsito, jogador de futebol de robôs, piloto automático de um boeing 747...

Racionalidade × Onisciência

- Racionalidade não é o mesmo que perfeição!
 - Um agente onisciente sabe o resultado real de suas ações a todo momento e pode agir de acordo com ele → algo impossível na prática!
 - A racionalidade maximiza o desempenho **esperado**, enquanto a perfeição maximiza o desempenho real (só descoberto posteriormente)
 - É irracional atravessar uma rua se, durante a travessia, cai um míssil na sua cabeça?
 - Contudo, agir com a finalidade de obter mais dados perceptivos faz parte da racionalidade.
 - Um agente racional deve realizar a ação de “olhar para os lados” antes de iniciar a travessia de uma rua, pois maximiza o desempenho esperado

Ambientes de tarefa

- Ao se projetar um agente, a primeira etapa deve ser sempre especificar o **ambiente de tarefa** de forma tão completa quanto possível
 - Ambiente de tarefa \leftrightarrow “Aspectos do Problema”



Ambientes de tarefa - PEAS

Tipo de agente	Medida de desempenho	Ambiente	Atuadores	Sensores
Motorista de táxi	Viagem segura, rápida, econômica, confortável	Estradas, pedestres, clientes	Direção, freio, acelerador, sinal, buzina	Câmeras, velocímetro, GPS, hodômetro
Sistema de diagnóstico médico	Paciente saudável, minimizar custos	Paciente, hospital, equipe	Perguntas, diagnose, exames, indicações	Sintomas, ressonância magnética, respostas
Sistema análise imagens satélite	Classificação correta imagem	Imagens de satélite	Exibir a classificação da cena	Pixels
Controlador de refinaria	Maximizar pureza, segurança	Refinaria, operadores	Válvulas, aquecedores, bombas	Leituras de temperatura, pressão
Tutor de português	Maximizar nota aluno em teste	Conjunto de alunos	Exibir dicas, exercícios	Palavras digitadas
Filtro de emails	No. de emails filtrados corret.	Mailbox	Aceitar, organizar, rejeitar	Mensagens

Ambiente de tarefa: características

↩ **Completa x Parcialmente observável:**

- Completamente observável quando os sensores do agente são capazes de coletar todas as informações relevantes sobre o ambiente.
- Parcialmente observável quando apenas um conjunto de informações é coletável.

Ex: Ambiente com diferentes dimensões...

Ambiente de tarefa: características

↩ **Determinístico x Estocástico:**

- Determinístico se o próximo estado do ambiente é definido a partir da atuação do agente sobre o estado atual.
 - Próximo estado = Estado atual + ação do agente
- Caso contrário, ele é dito estocástico (incerto). Um ambiente com outros agentes se torna estocástico, pois seus estados futuros são afetados pelos outros agentes.
Ex: Trânsito urbano
- Mundo do aspirador de pó simples: determinístico
- Mundo do aspirador de pó c/ sujeira surgindo ao acaso: estocástico

Ambiente de tarefa: características

↩ Episódico x Sequencial:

- Em um ambiente episódico, a atuação do agente é dividida em episódios atômicos onde o agente percebe e efetua apenas uma ação por episódio, independente dos anteriores. Não há relação ou dependência entre episódios.
- Nos ambientes sequenciais, por sua vez, existe a dependência. Por exemplo, um jogo de xadrez é sequencial, pois decisões de curto prazo podem ter consequências a longo prazo.

Ambiente de tarefa: características

↩ Estático x Dinâmico:

- Quando as alterações que podem ocorrer no ambiente independem da atuação do agente, este ambiente é dito dinâmico. Em geral, a dinamicidade ocorre por conta das ações realizadas por outros agentes, enquanto o agente estiver deliberando sobre qual será a sua próxima ação. Exemplos:
- Jogo de palavras cruzadas é estático.
- Trânsito urbano é dinâmico

Ambiente de tarefa: características

↩ Discreto x Contínuo:

- Um ambiente discreto tem um número definido de variáveis e valores possíveis.
- Um ambiente contínuo, por sua vez, podem ter variáveis com um número infinito de valores (contínuos).
- Ex. O xadrez por exemplo tem um número definido de variáveis e de movimentos, logo é discreto.

Ambiente de tarefa: características

↩ Agente único x Multi-Agente:

- O ambiente pode ser composto por apenas um agente ou muitos agentes.
- No ambiente multiagente, é necessário contar com agentes responsáveis pela coordenação dos agentes envolvidos (colaboração, competição ou ambos). Tais sociedades de agentes são denominadas de Sistemas Multi-Agentes (SMA) ou Multi Agent Systems (MAS).

Ambiente de tarefa: características

- Diferentes ambientes requerem diferentes agentes.
- O pior ambiente possível:
 - Inacessível
 - Não-determinístico
 - Não episódico
 - Dinâmico
 - Contínuo
 - Ex: dirigir um carro.

Ambientes de tarefa: exemplos

Ambiente	Observável	Determinístico	Episódico	Estático	Discreto	Multiagente
Jogo palavras cruzadas	Sim	Sim	Não	Sim	Sim	Não
Xadrez com relógio	Sim	Sim (Estratégico)	Não	Semi	Sim	Sim
Pôquer	Não	Sim (Estratégico)	Não	Sim	Sim	Sim
Gamão	Sim	Não	Não	Sim	Sim	Sim
Direção de táxi	Não	Não	Não	Não	Não	Sim
Diagnóstico médico	Não	Não	Não	Não	Não	Não
Tutor	Não	Não	Não	Não	Sim	Sim
Análise de imagens	Sim	Sim	Sim	Semi	Não	Não
Robô de seleção de peças	Não	Não	Sim	Não	Não	Não
Controlador de refinaria	Não	Não	Não	Não	Não	Não

Agentes de Software X Software Convencional

- Agentes de Software são diferentes, segundo o MIT, porque são:
 - Semi-autônomos
 - Proativos
 - Adaptativos
 - Duram mais
 - Comportam-se mais como assistentes do que como ferramentas

Composição de um agente básico

- Agente = Arquitetura (1) + Programa (2)

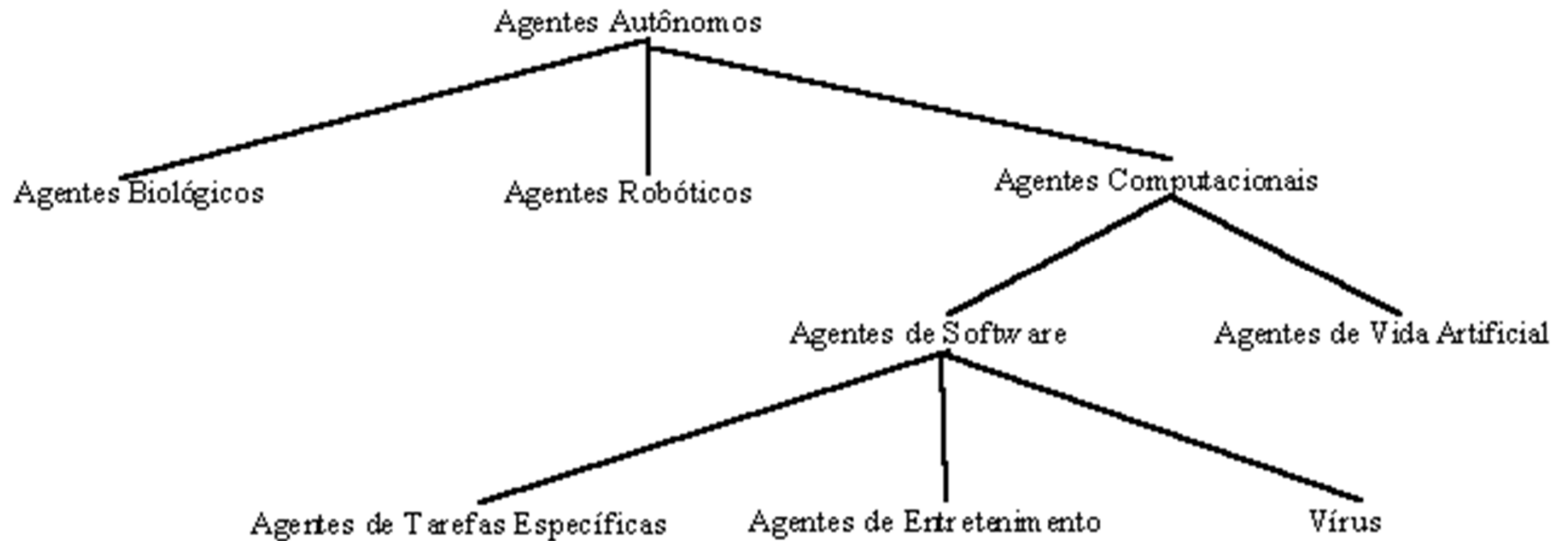
(1) Composição física do agente, representando os seus sensores e atuadores → **corpo**

- Interage com o ambiente, tornando as percepções dos sensores disponíveis para o programa de agente e retornando as opções de ação geradas pelo programa para os atuadores

(2) Implementação computacional da função de agente

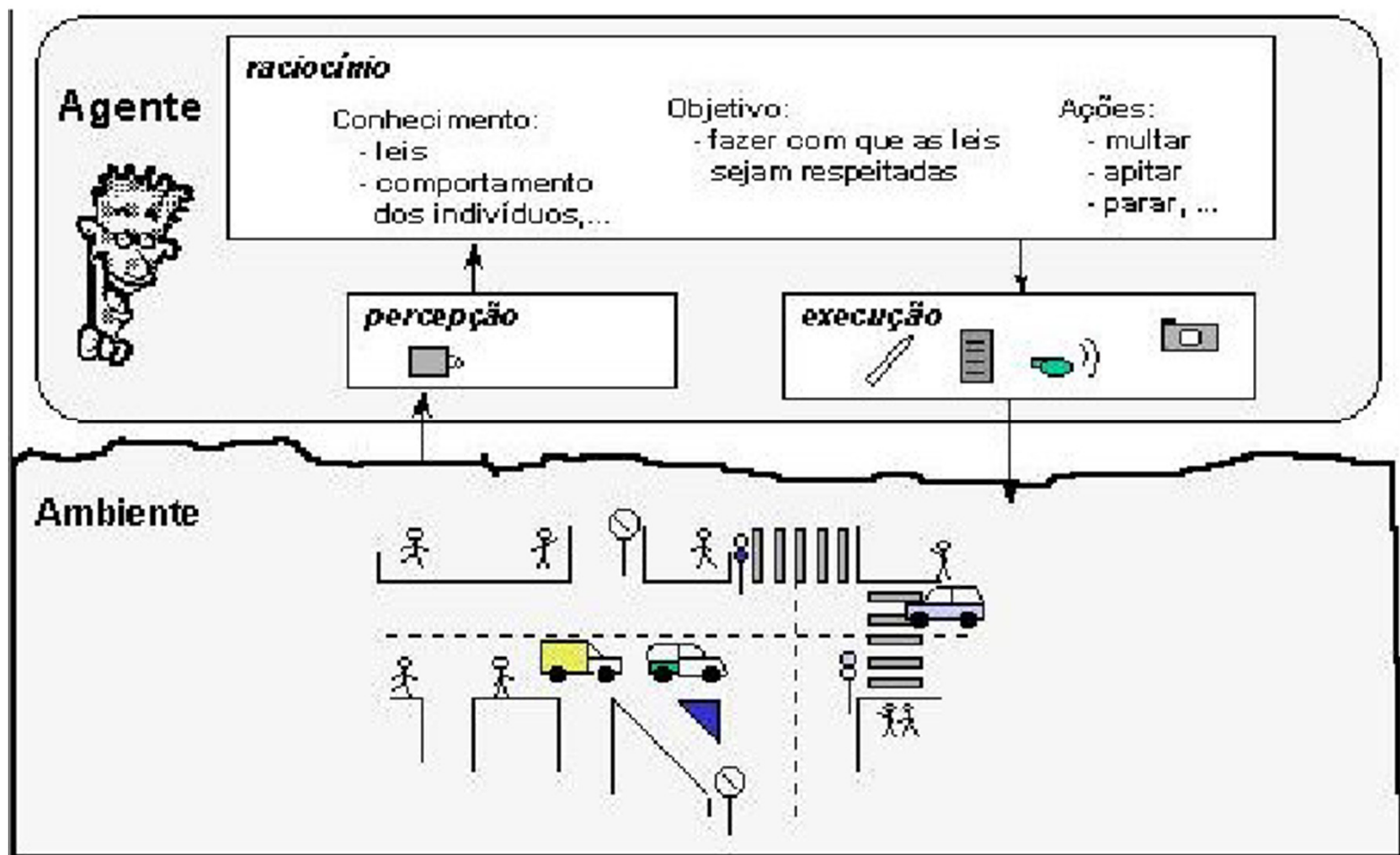
- **Estrutura básica**: Receber somente a percepção **atual** como entrada dos sensores e retornar a próxima ação
 - Se as ações do agente dependerem de sequências de percepções, então terá que memorizá-las!

Categorias de agentes



Taxonomia de Agentes proposta por Franklin e Graesser (1996).

Agentes Biológicos: humano (policia de transito)



Agentes Robóticos

- ↖ Dispositivos mecânicos controlados por computador, capazes de realizar tarefas de processamento, montagem e transporte
- ↖ Usado para fins industriais, domésticos, entretenimento ou de pesquisa

Agentes de Software

- ↖ Agentes de Interface, de auxílio ao usuário
- ↖ Agentes de Aconselhamento – palpites/dicas, financeiro
- ↖ Agentes de Internet – busca, viagens, compras, etc

Outras classificações:

- ↖ Classificação simples em Reativos e Cognitivos
- ↖ Classificação de Russel e Norvig (1995)
- ↖ Classificação de Woldridge (1999)

Agentes Reativos x Agentes Cognitivos

↩ Agentes Reativos

- ↩ Reagem aos estímulos do ambiente.
- ↩ Agentes simples.
- ↩ Tem pouca autonomia.
- ↩ Não possuem uma representação interna do ambiente.
- ↩ Não há memória das ações realizadas.
- ↩ Limitado em relação as possibilidades de atuação.
- ↩ Baseado em regras simples do tipo estímulo reação.

Agentes Reativos x Agentes Cognitivos

↩ Agentes Cognitivos/Deliberativos

- ↩ Agentes mais complexos
- ↩ Tem suas ações voltadas para seus objetivos.
- ↩ Autônomos
- ↩ Tem uma representação interna do ambiente.
- ↩ Baseado em regras complexas.
- ↩ Pode aprender.

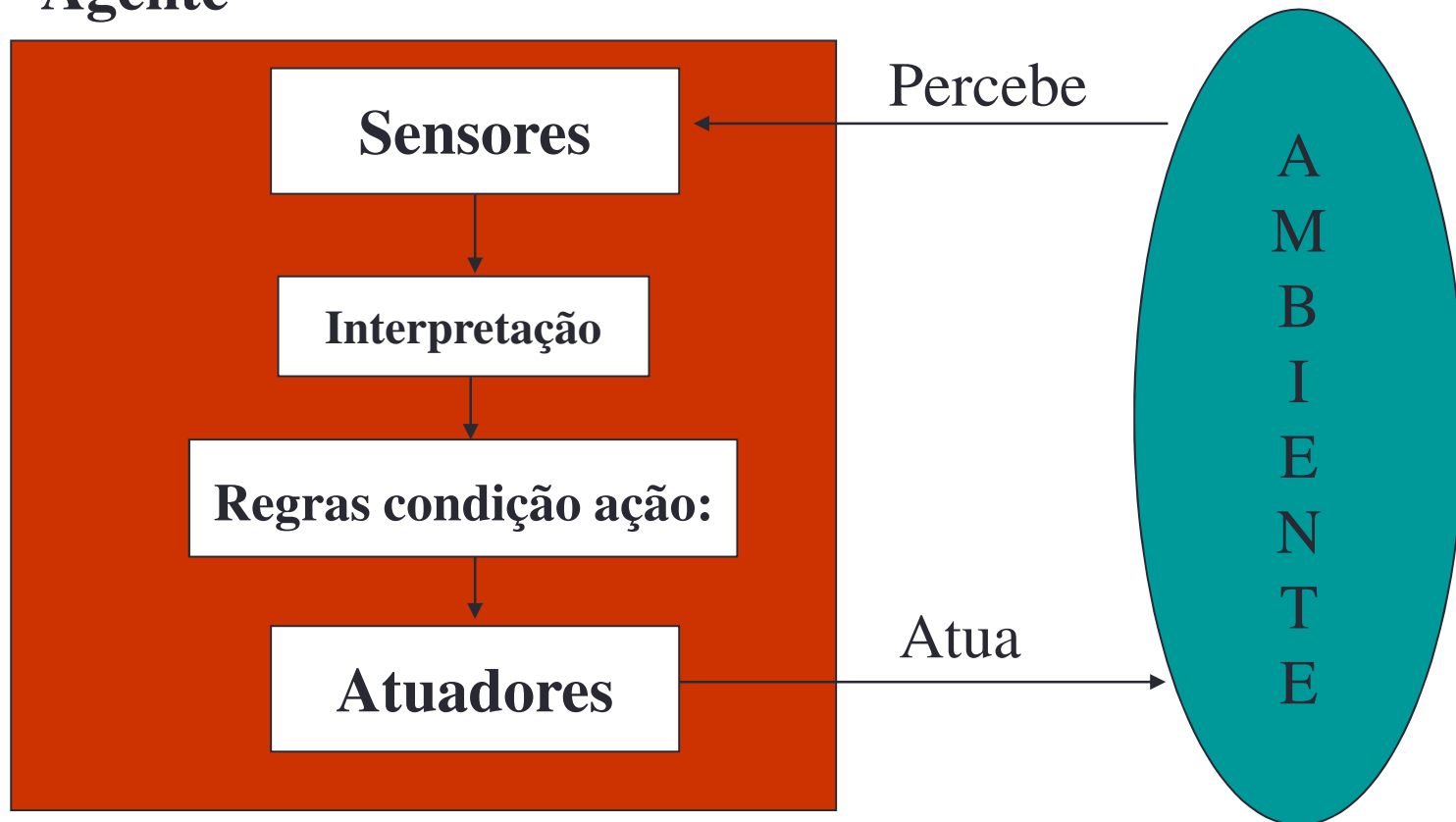
Classificação de Russel e Norvig

- Colocam classes intermediárias entre reativos e cognitivo:
 - Reativos simples
 - Reativos baseados em modelo
 - Baseados em objetivos
 - Baseados na utilidade
 - com aprendizagem

Reativos Simples

- É a estrutura de agente mais simples.
- Seleciona suas ações baseado apenas na percepção atual do ambiente.

Agente

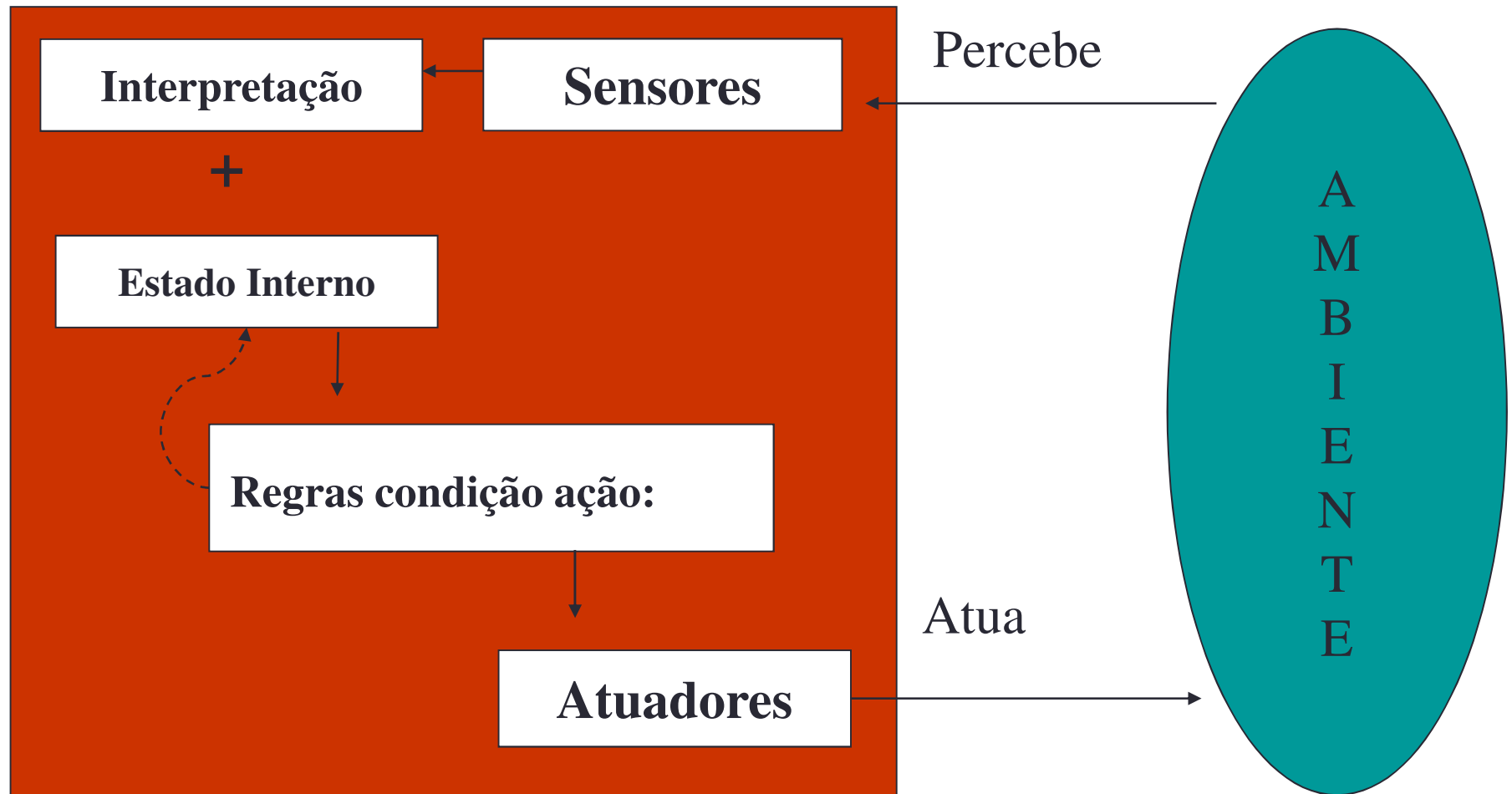


Reativos Baseados em Modelo

- Agente se baseia não só no estado atual do ambiente (o que ele está percebendo), mas em uma representação interna.
- A representação interna funciona como um histórico de percepções, uma base de experiências.

Reativos Baseados em Modelo: Arquitetura

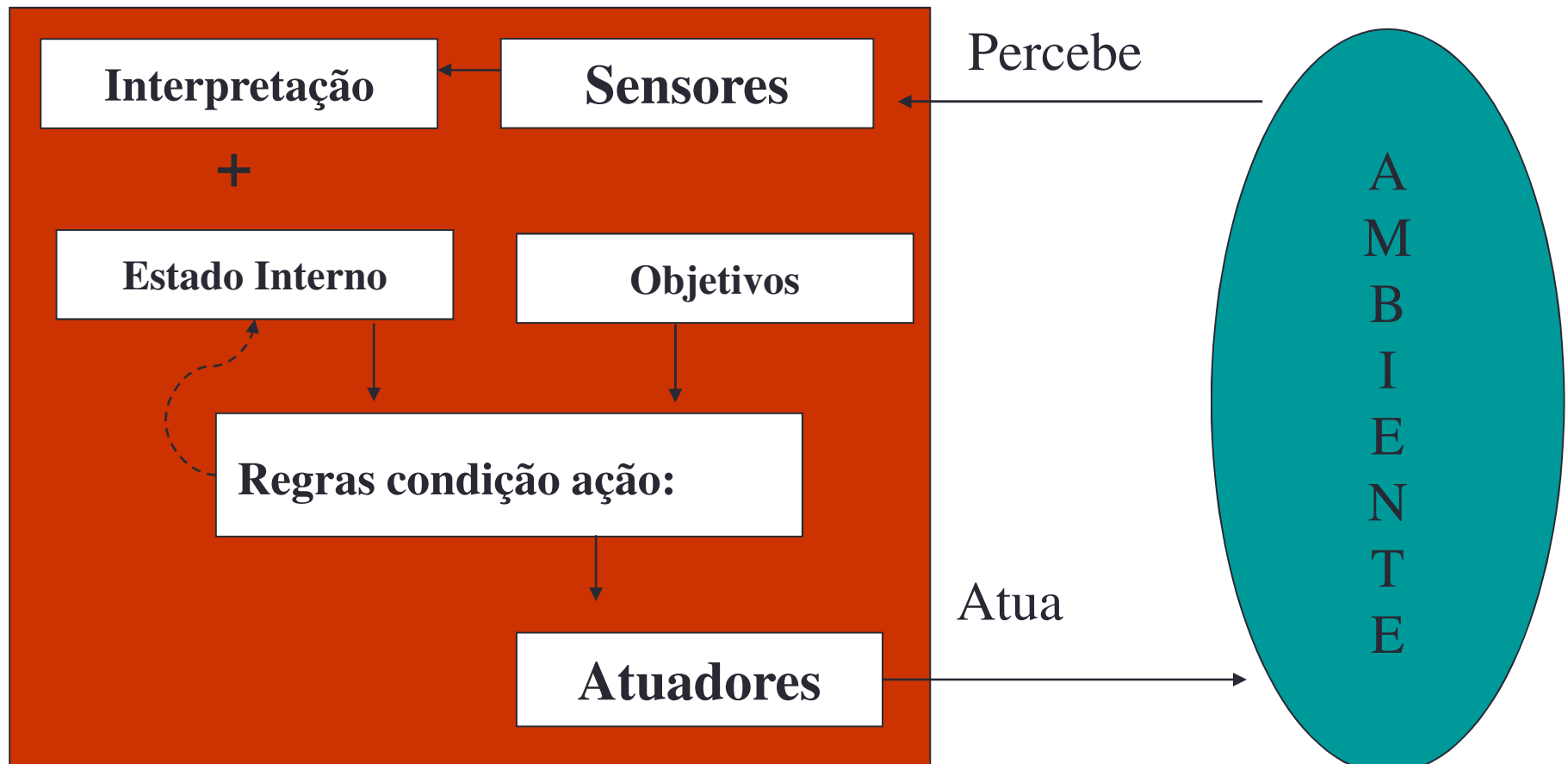
Agente



Baseados em Objetivos

- Agente se baseia não só no estado atual do ambiente e em uma representação interna, mas em um conjunto de objetivos que ele deve alcançar.

Agente

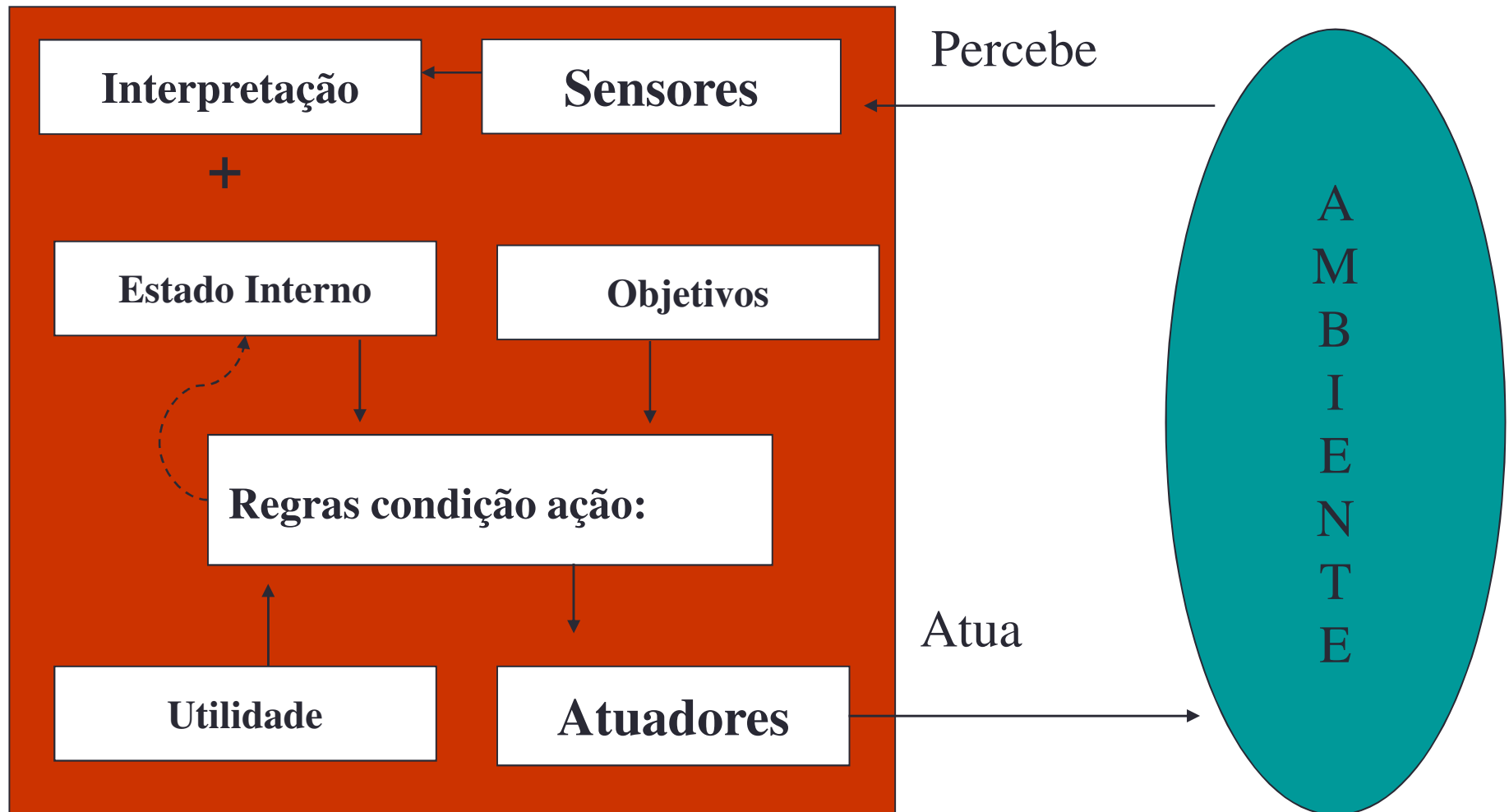


Baseados em Utilidade

- Não só o agente busca um conjunto de objetivos, mas os busca da melhor forma possível.
- Objetivos simplesmente dizem onde chegar, não se importando com a qualidade do processo ou caminho tomado.
- Utilidade:
 - Se a regra A associada a B realiza o objetivo com 30% de satisfação e a regra A associada a C realiza o objetivo com 70% de satisfação então utilizar A e C!

Baseados em Utilidade: Arquitetura

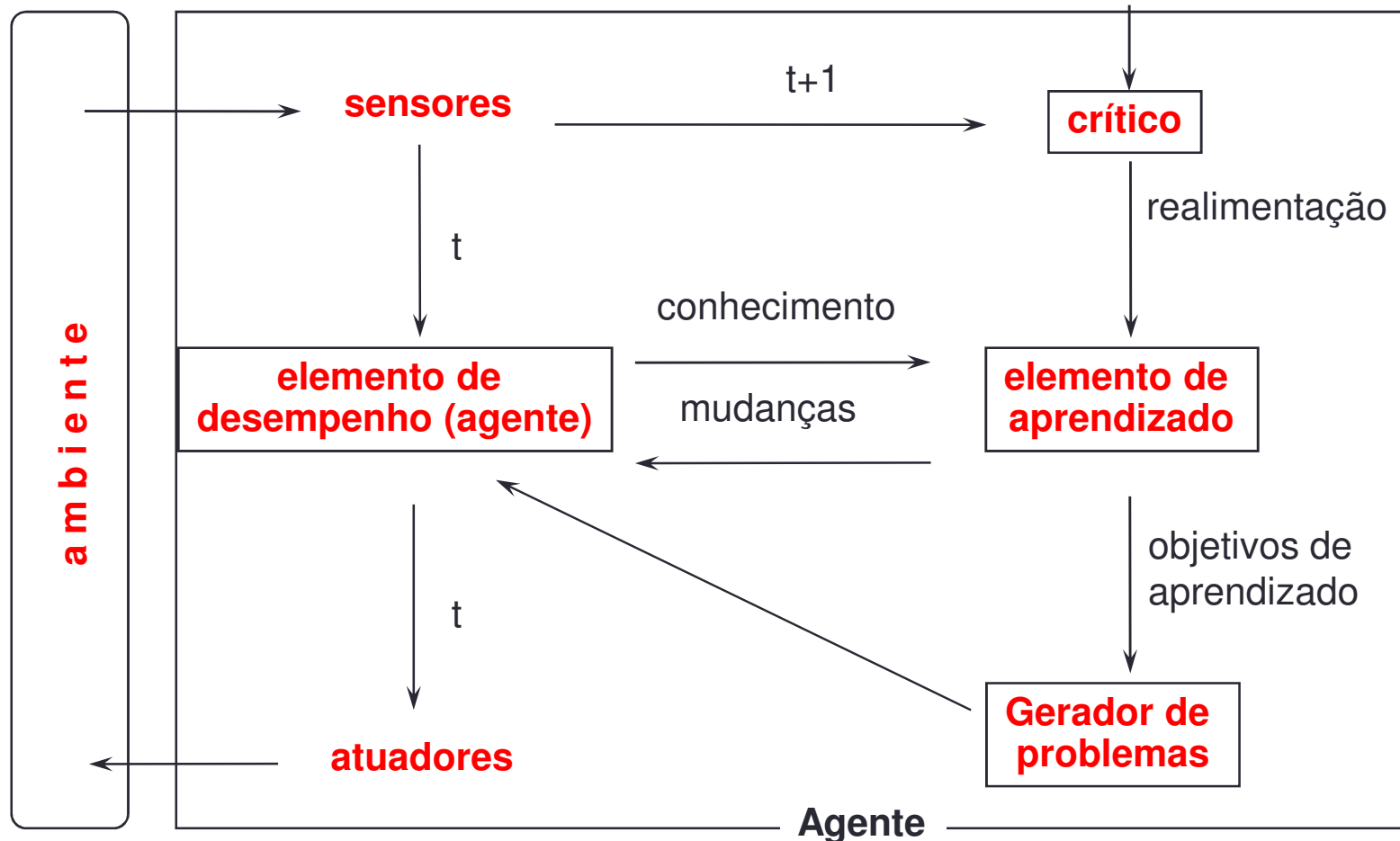
Agente



Agente com aprendizagem

- Composto de 4 componentes conceituais:
 - **Elemento de aprendizado**: executa aperfeiçoamentos
 - **Elemento de desempenho**: seleciona ações externas
 - É quem realmente realiza o ciclo de percepções-ações!
 - **Crítico**: avalia como o agente está se comportando em relação a um padrão fixo de desempenho (externo)
 - É necessário pois as percepções não oferecem nenhuma indicação (reforço ou penalidade) do sucesso do agente
 - **Gerador de problemas**: sugere ações exploratórias que levarão a experiências novas e informativas, melhorando o desempenho do agente a longo prazo
- Vantagem:
 - **Capaz de operar em ambientes inicialmente desconhecidos e de se tornar mais competente do que seu conhecimento inicial sozinho poderia permitir**

Agente adaptativo



Classificação de Woldridge

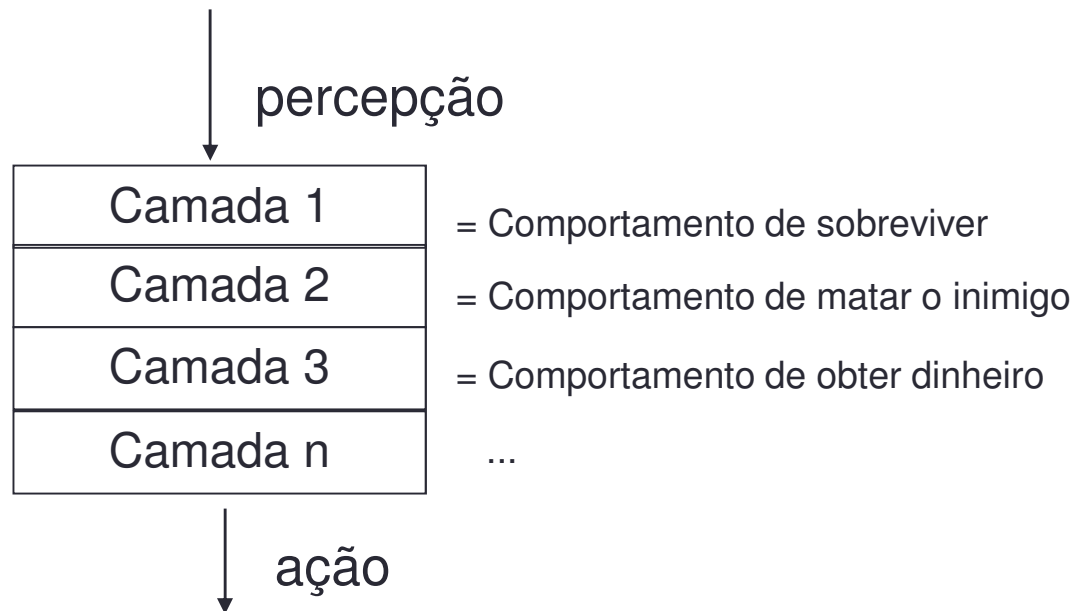
- Woldridge classifica as arquiteturas de agentes em quatro tipos:
 - Arquiteturas baseadas em lógica
 - Arquiteturas reativas
 - Arquiteturas em camadas
 - Arquiteturas BDI

Classificação de Woldridge

- **Arquiteturas baseadas em lógica:** Similar as arquiteturas baseadas em objetivos de Russel e Norvig.
- **Arquiteturas reativas :** Similar as arquiteturas reativas simples de Russel e Norvig.

Classificação de Woldridge

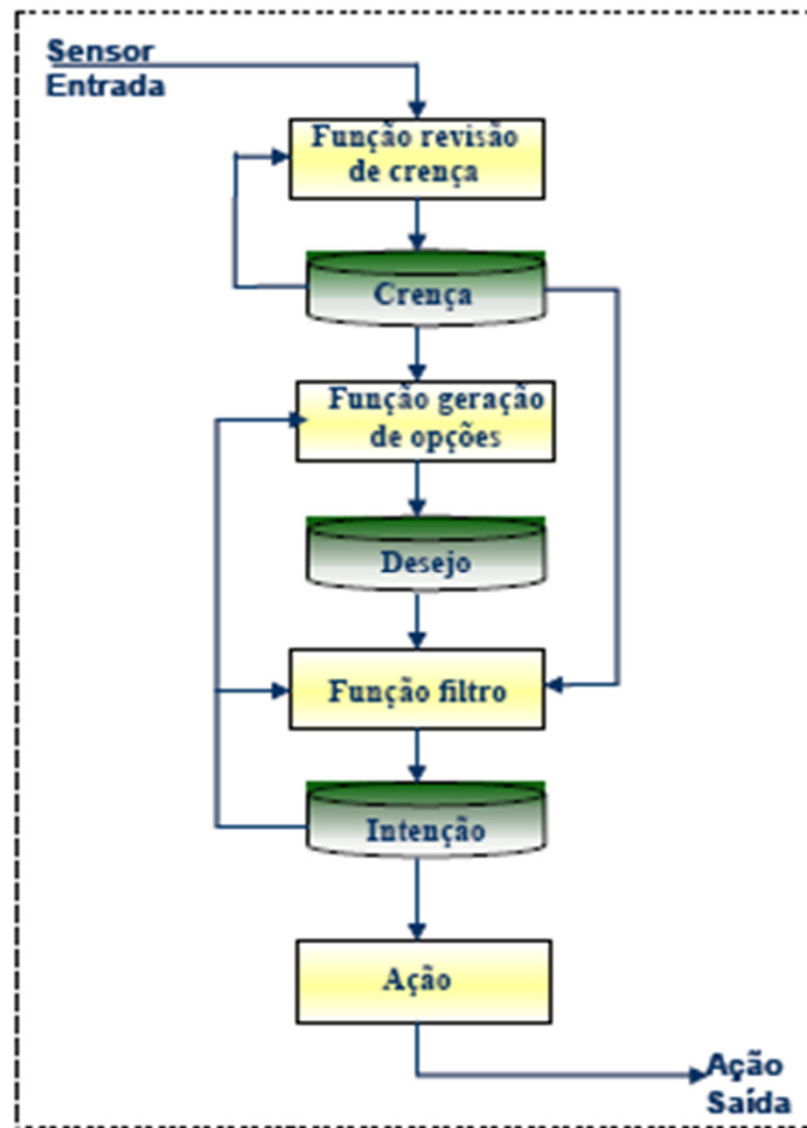
- Arquiteturas em camadas ou híbrida:
 - A decisão da ação a executar é feita passando por várias camadas.
 - Cada camada raciocina em diferentes níveis de abstração



Classificação de Woldridge

- **Arquiteturas BDI (Rao & George, 1995):**
 - Baseiam-se em teorias da sociologia e psicologia de que os seres humanos são regidos por três estados mentais fundamentais:
 - **Crenças:** o que o agente acredita (sabe) sobre o ambiente em que se encontra.
 - **Desejos:** o que motiva o agente.
 - **Intenções:** as ações do agente.

Arquitetura BDI



Aplicações de agentes:

- ↖ Assistentes pessoais
- ↖ Sistema para atendimento a público
- ↖ *Exploração de informação WEB*
- ↖ *Gerenciamento de redes*
- ↖ *Controle de tráfego aéreo / terrestre*
- ↖ *Sistemas de simulação de vida real*
- ↖ *Jogos educativos*
- ↖ *Sistemas tutores*
- ↖ *Automação Industrial*

Características de Agentes

↩ Reatividade

- ↩ *Um agente deve ser capaz de reagir apropriadamente a influências ou informação de seu ambiente.*
- ↩ *Um agente reativo comporta-se num modo estímulo-resposta, isto é, ele não tem memória da ação realizada no passado nem qualquer previsão da ação a ser tomada no futuro.*

Características de Agentes

↩ Pro-atividade

↩ *É um nível acima da reatividade.*

↩ *Um agente inteligente não apenas reage a mudanças de seu ambiente, porém ele próprio tem iniciativa sob circunstâncias específicas.*

↩ *Iniciativa de agir em busca de seus objetivos.*

Características de Agentes

↩ Capacidade de Aprendizado

↩ *Capacidade de incorporar dados e informações a suas bases de conhecimento e inferir sobre as mesmas de forma a adaptar-se às mudanças do ambiente modificando suas ações (adaptação).*

Características de Agentes

↩ Autonomia

↩ *Capacidade de agir de forma autônoma, sem intervenção externa, para o cumprimento de seus objetivos.*

↩ Autonomia é a habilidade de exercer controle sobre suas próprias ações.

Características de Agentes

↩ Mobilidade

- ↩ *Habilidade para mover-se de uma localização física ou lógica para outra, preservando seu estado interno.*
- ↩ *Capacidade de mover-se entre pontos distintos de uma rede.*

Características de Agentes

↩ Comunicação

↩ *Capacidade de trocar informações com outras entidades (agentes, humanos, objetos, ambiente)*

↩ Representatividade

↩ *Capacidade de representar, substituir, alguém ou alguma coisa em alguma atividade específica.*

↩ Benevolência

↩ *Suposição de que um agente não terá objetivos conflitantes, e que agente sempre tentarão fazer o que lhes foi pedido.*

Atividade Extra-sala

↩ Livro Russell & Norvig, Exercício 2.5

↩ Livro Russell & Norvig, Exercício 2.6

↩ Livro Russell & Norvig, Exercício 2.9