Conhecimento e Raciocínio
Agente Baseado em Conhecimento

Agentes que Resolvem Problemas

- representados através de algoritmos de busca
- agentes não possuem acesso ao conhecimento do domínio explicitamente representado
  - somente implicitamente, através das funções de geração de sucessores e teste de gol
- Agentes baseados em conhecimento podem superar essas limitações

3

Knowledge-based Agent

Agente Baseado em Conhecimento (Agente-KB)

- Possue conhecimento sobre o mundo e sobre as suas próprias ações
- Raciocina para inferir uma maneira de atingir um determinado objetivo: raciocina sobre as suas possíveis ações

4

Agente-KB simples

Deve ser capaz de :

- Conhecer e representar: estados, ações, etc.
- Incorporar novas percepções e atualizar representações internas do mundo
- Inferir propriedades do mundo (não observáveis) a partir de suas sequências de percepção
- Formular os seus objetivos
- Inferir ações apropriadas (deve saber o que as suas várias ações fazem em diferentes situações do mundo)

5

Agentes Baseados em Conhecimento

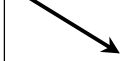
- **Base de Conhecimento** (*Knowledge Base - KB*): principal componente de um agente baseado em conhecimento
  - conjunto de representações de **fatos sobre o mundo**. Cada representação individual é chamada de **sentença**.
  - sentenças são representadas em uma **linguagem de representação de conhecimento**

6

Agentes Baseados em Conhecimento

- Tarefa **Tell**: adiciona novas sentenças à Base de Conhecimento.
- Tarefa **Ask**: pergunta à Base de Conhecimento o que é conhecido.

Quando alguém faz uma pergunta **Ask** para a Base de Conhecimento a resposta deve ser *derivada* a partir do que foi previamente adicionado (através de **Tell**) à Base de Conhecimento



Responsabilidade do Mecanismo de Inferência

7

Agente-KB genérico

Função KB-agente (*percepção*) **devolve** *ação*

estático: *KB*, *t*

Informa(*KB*, Gera-sentença(*percepção*,*t*))

ação ⟵ Pergunta(*KB*, Gera-query(*t*))

Informa(*KB*, Gera-sentença(*ação*,*t*))

t ⟵ *t* + 1

devolve *ação*

8

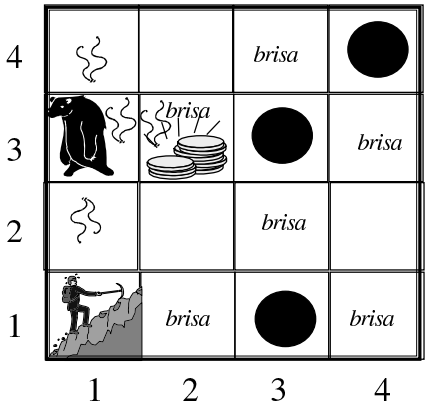
Três níveis de descrição de um agente baseado em conhecimento:

- **Nível do conhecimento ou epistemológico:** descrição das categorias de conhecimento que o agente possue e como ele usa o conhecimento para raciocinar (*o que* ele sabe e *como* usa o seu conhecimento)
- **Nível lógico:** descreve como o conhecimento é representado formalmente e descreve o raciocínio do agente através de mecanismos lógicos de provas de teoremas (mecanismos de propósito geral)
- **Nível de implementação:** descreve a Base de Conhecimento como estruturas de dados e os algoritmos que as manipulem (estruturas computacionais)

Sistemas Baseados em Conhecimento

Mecanismo de Inferência	←	algoritmos independentes do domínio (objetivo de IA)
Base de Conhecimento	←	conhecimento específico do domínio (ontologias)

O mundo de Wumpus



11

Ambiente

- quadrados adjacentes ao wumpus *cheiram* mal
- quadrados adjacentes ao abismo possuem uma *brisa*
- um quadrado *brilha* sss contém o ouro
- o agente só possue uma flexa
- *atirar* a flexa mata o wumpus se o agente estiver de frente para ele (linha ou coluna)
- *pegar* o ouro só é possível se o agente estiver no mesmo quadrado
- *soltar* deixa o ouro no mesmo quadrado em que o agente estiver

12

O mundo de Wumpus

- **percepção**: brisa, cheiro, brilho, escuta, choque (não percebe sua localização)
- **ação**: escala, vira para direita, vira para esquerda, vai para frente, segura, solta, atira
- **objetivo**: encontrar o ouro e trazê-lo para o início o mais rápido possível sem entrar num quadrado com abismo ou com wumpus
  - O agente morre se entra em um quadrado que possue um wumpus vivo ou um abismo

O mundo de Wumpus - ambiente

Classe de ambientes para a qual queremos construir um agente-KB:

- grade 4x4 cercada de paredes
- o agente sempre inicia no quadrado (1,1) virado para a direita
- as localizações do ouro e do wumpus são aleatórias com uma distribuição uniforme
- a menos do quadrado de entrada, todos os outros podem conter um abismo, com probabilidade 0,2

Mundo de wumpus

- Determinístico?
- Acessível?
- Estático?
- Discreto?

Mundo de wumpus

- Determinístico? Sim: efeitos de ações totamente especificadas.
- Acessível? Não. O agente possue percepção local
- Estático? Sim: wumpus e abismos não se locomovem (em outras versões desse jogo o wumpus se locomove)
- Discreto? Sim

Representação de Conhecimento (RC)

- Principal problema no desenvolvimento de um agente inteligente:
  - ***como representar o conhecimento que o agente possue sobre o domínio do problema?***
  - ***como raciocinar usando esse conhecimento para responder perguntas ou tomar decisões inteligentes?***
- **Lógica**: uma das mais antigas linguagens de Representação de Conhecimento estudadas em IA.
  - Base de muitos sistemas que tiveram a lógica como inspiração: p. ex., Sistema Especialista baseado em regras e linguagem de programação PROLOG

Agente baseado em conhecimento

- Deve:
  - Representar conhecimento sobre o mundo em uma ***linguagem formal*** (KB)
  - Raciocinar sobre o mundo usando ***inferências*** na linguagem (sobre a KB)
  - Decidir que ação tomar ***inferindo*** que a ação selecionada é a melhor

Linguagens de representação

- Problema ⟶ deve-se fazer um compromisso entre:
  - uma linguagem que é expressiva o suficiente para representar os objetos, relações e propriedades em um domínio;
  - permitir uma maneira tratável (eficiente) de raciocinar e responder questões sobre informações implícitas em tempo razoável
- **Lógica** é uma linguagem de propósito geral para descrever o que é verdade e o que é falso no mundo e que possue procedimentos mecânicos que podem operar sobre sentenças da linguagem para realizar o *raciocínio* no computador

Lógica

- linguagem de propósito geral para descrever o que é verdade e o que é falso no mundo

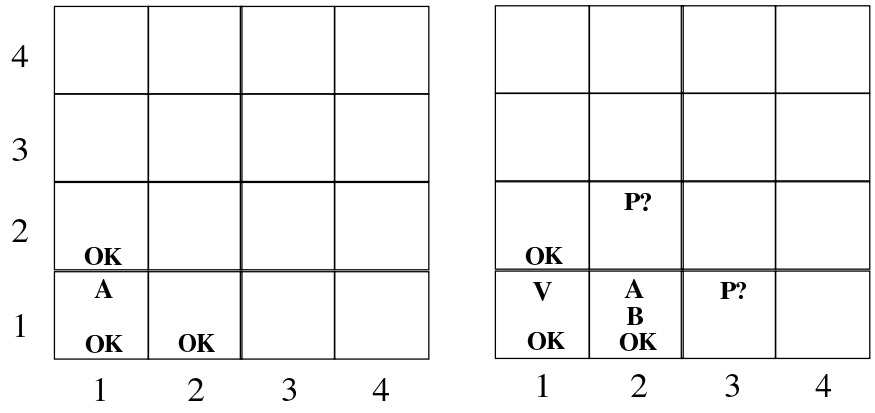
+

- procedimentos mecânicos que podem operar sobre sentenças da linguagem para realizar o *raciocínio*

KB não(brisa(2,1)) e não(cheiro(1,2))	*deriva*	KB + não(abismo(2,2)) e não(wumpus(2,2))

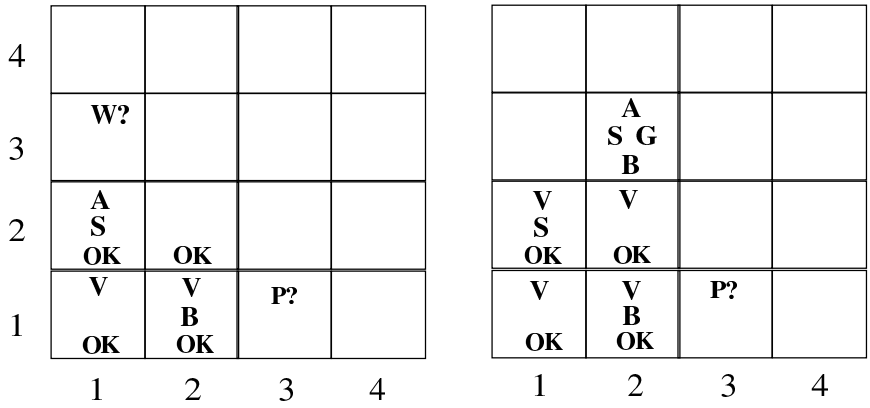
Agente Lógico no Mundo do Wumpus

Percepção (i,j) ≡ [<*Stench*>,<*Breeze*>,<*Glitter*>,<*Bump*>,<*Scream*>]



Agente Lógico no Mundo do Wumpus

Percepção (i,j) ≡ [<*Stench*>,<*Breeze*>,<*Glitter*>,<*Bump*>,<*Scream*>]



Os dois objetivos de IA

① construir artefatos (produtos) capazes de desempenhar tarefas que, quando executadas por humanos, demandam inteligência

② chegar a um entendimento científico dos princípios que fundamentam o comportamento inteligente, conforme manifestado em humanos, animais e especificamente nos produtos do primeiro objetivo da IA.

Raciocínio correto sobre uma ***representação correta***

- Podemos arguir que a melhor maneira de se obter comportamento inteligente será como um produto de um ***raciocínio correto*** sobre uma ***representação correta***
- A Lógica serve para esse propósito e portanto é um instrumento para se atingir o objetivo ② da Inteligência Artificial

Tipos de Lógica

Lógicas são caracterizadas pelas suas "primitivas":

- **Compromisso ontológico**: *o que existe - fatos? objetos? tempo? crenças?*
- **Compromisso epistemológico**: *quais estados do conhecimento ela admite?*

Tipos de lógicas

Linguagem	Compromisso Ontológico	Compromisso Epistemológico
Lógica Proposicional	fatos	verdade/falso/ desconhecido
Lógica de Primeira Ordem	fatos, objetos, relações	verdade/falso/ desconhecido
Lógica Temporal	fatos, objetos, relações, tempo	verdade/falso/ desconhecido
Teoria da Probabilidade	fatos	escala de crença 0 .. 1

Lógica

- Sistema formal onde fórmulas ou sentenças possuem valores Verdadeiro ou Falso
- Uma lógica inclue:
  - **Sintaxe**: especifica os símbolos na linguagem e como eles podem ser combinados para formar sentenças.
  - **Semântica**: define o "significado" de sentenças, isto é, define a verdade de uma sentença no mundo
  - **Procedimento de Inferência**: método mecânico para computar (derivar) novas sentenças (verdades) a partir de outras

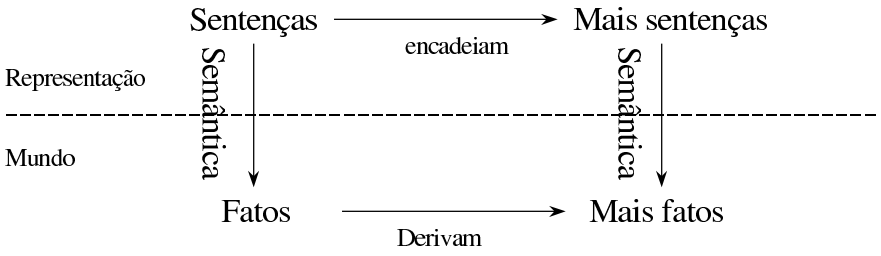
Linguagem da aritmética

- $x + 2 > y$ é uma sentença
- $x + y >$ não é uma sentença
- $x + 2 \geq y$ é verdade sse o número $x + 2$ não for menor que o número y
- $x + 2 > y$ é verdade no mundo onde $x = 7, y = 1$
- $x + 2 > y$ é falso no mundo onde $x = 0$, $y = 6$

Lógica

- **Fatos** são declarações sobre o mundo que podem ser verdadeiras ou falsas.
- **Representação** é uma expressão (sentença) de uma linguagem que pode ser codificada em um programa de computador e descreve os objetos e relações do mundo
- É preciso ter certeza que a representação é consistente com a realidade

Correspondência entre a representação e o mundo



Encadeamento de Sentenças

- Geração de novas sentenças <u>necessariamente verdadeiras</u> a partir de sentenças prévias também verdadeiras (*KB*)
- encadeamento entre a *KB* e uma sentença:

$$KB \models \alpha$$

KB <u>encadeia</u> a sentença α sse

α é verdade em todos os mundos onde KB é verdade

Mecanismo de Inferência

- Um agente pode raciocinar sobre a sua Base de Conhecimento (KB) de duas maneiras:
  - dada uma *KB,* o ***mecanismo de inferência*** pode gerar novas sentenças α encadeadas pela *KB,* ou seja, $KB \models \alpha$

ou

  - dada uma *KB* e uma sentença α, o mecanismo de inferência pode dizer se α é encadeada pela *KB.* Exemplo: a posição (2,3) é segura? (α ou $\neg \alpha$?)

Propriedades de um mecanismo de inferência

- Se o mecanismo somente gera (infere) <u>todas</u> as sentenças encadeadas pela KB, então ele é **completo** ou **preserva a verdade**
- Se as sentenças inferidas pelo mecanismo de inferência correspondem aos fatos que derivam dos fatos iniciais sobre o mundo (sentenças iniciais da BC)
  então ele é **correto** (**sound**)
- melhor **correto** do que **completo !!**

Lógica Proposicional

- É a lógica mais simples e serve para ilustrar as idéias básicas
- Define-se um conjunto de **símbolos** como *P* e *Q* , e uma semântica para cada um desses símbolos. Por exemplo:
  - P significa "Está quente"
  - Q significa "Está úmido"
  - R significa "Está chovendo"

Lógica Proposicional - objetos

- São objetos da linguagem :
  - átomos/proposições (P, Q, homem-rico, … etc)
  - conectivos lógicos:
    - $\neg$ negação
    - $\wedge$ conjunção (AND)
    - $\vee$ disjunção (OR)
    - $\rightarrow$ implicação (ou condicional)
    - $\leftrightarrow$ equivalência (bicondicional)

Lógica Proposicional - sentenças

- Uma **sentença,** também chamada de fórmula ou fórmula bem formada (*well formed formula - wff*), é definida como:
  - Um átomo é uma sentença
  - Negação de uma sentença é uma sentença
  - Conjunção, disjunção e implicação de 2 sentenças é uma sentença
- **Literal**: um átomo ou um átomo negado

Lógica Proposicional

Se S é uma sentença, $\neg$S é uma sentença

Se S1 e S2 são sentenças, S1 $\wedge$ S2 é uma sentença

Se S1 e S2 são sentenças, S1 $\vee$ S2 é uma sentença

Se S1 e S2 são sentenças, S1 $\rightarrow$ S2 é uma sentença

LP - Gramática BNF

Sentença => SentençaAtômica | SentençaComplexa

SentençaAtômica => *Verdade* | *Falso* | P | Q | R | …

SentençaComplexa => (Sentença) |
Sentença Conectivo Sentença |
¬ Sentença

Conectivo => ∧ | ∨ | → | ↔

Ordem de **p**recedência dos operadores: ¬, ∧, ∨, →, ↔

Exemplos de fórmulas

homem_rico

homem_pobre

homem_rico → ¬ homem_pobre

homem_rico ∧ quer_lua_homem →
tem_lua_homem

homem_pobre ∧ quer_lua_homem →
¬ tem_lua_homem

Semântica do Cálculo Proposicional

- Significado de uma sentença através da noção de interpretações
- *Interpretação de uma sentença*: uma atribuição de valores verdade para cada átomo (proposição) da fórmula
- Exemplo:

  Seja a fórmula: $(A \vee B) \rightarrow (C \vee \neg D)$

  e uma interpretação:

  A:= V B:= F C:= F D:= V

Semântica do Cálculo Proposicional

- A verdade da fórmula geral é calculada baseada nas leis da Lógica Booleana: como computar a verdade de sentenças envolvendo conectivos ($\neg$, $\wedge$, $\vee$, $\rightarrow$) dados os valores verdade de seus constituintes ?
- Dada uma fórmula contendo n átomos, existem 2^n interpretações possíveis (valores V/F dos átomos individuais)
- Pode-se escolher qualquer interpretação para os átomos de uma KB porém, a verdade de todas as sentenças que envolvem esses átomos é DETERMINADA pelas regras da Lógica de Boole

Semântica do Cálculo Proposicional

Regras para avaliação da verdade de uma sentença com relação a um modelo M:

$\neg$S	é V sse	S	é F		
S1 $\wedge$ S2	é V sse	S1	é V e	S2	é V
S1 $\vee$ S2	é V sse	S1	é V ou	S2	é V
S1 $\rightarrow$S2	é V sse	S1	é F ou	S2	é V
isto é,	é F sse	S1	é V e	S2	é F
S1 $\leftrightarrow$ S2	é V sse	S1 $\rightarrow$ S2	é V e	S2 $\rightarrow$ S1	é V