

Sistemas Inteligentes

Aula 21/10

Agentes Lógicos

Agente Baseado em Conhecimento

- **Agentes Baseados em Conhecimento** ou **Agentes Lógicos**.
- Podem lidar mais facilmente com **ambientes parcialmente observáveis**.
- O agente pode usar as suas percepções e conhecimento do mundo para **inferir aspectos ainda desconhecidos** do ambiente.
- São **flexíveis** e podem assumir novas tarefas na forma de objetivos explicitamente descritos.

Agente Baseado em Conhecimento

- O componente central de um agente baseado em conhecimento é sua **base de conhecimento**.
- A base de conhecimento é formada por um conjunto de **sentenças** expressas por meio de uma **linguagem lógica** de representação de conhecimento.
- Deve ser possível adicionar novas sentenças à base e consultar o que se conhece. Ambas as tarefas podem envolver **inferência** (derivação de novas sentenças a partir de sentenças antigas).

Agente Baseado em Conhecimento

- Agente genérico baseado em conhecimento:

```
função AGENTE-KB(percepção) retorna uma ação  
  persistente: KB, uma base de conhecimento  
                t, um contador, inicialmente igual a 0,  
                indicando tempo  
  TELL(KB, CRIAR-SENTENÇA-DE-PERCEPÇÃO(percepção, t))  
  ação ← ASK(KB, CRIAR-CONSULTA-DE-AÇÃO(t))  
  TELL(KB, CRIAR-SENTENÇA-DE-AÇÃO(ação, t))  
  t ← t + 1  
retornar ação
```

Agente Baseado em Conhecimento

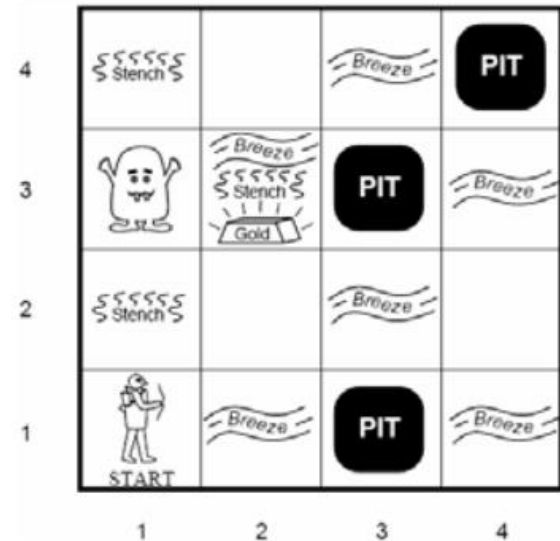
- Processo de execução de um agente baseado em conhecimento:
 - **(1)** Informa à base de conhecimento o que o agente está percebendo do ambiente.
 - **(2)** Pergunta à base de conhecimento qual a próxima ação que deve ser executada. Um extenso processo de **raciocínio lógico** é realizado sobre a base de conhecimento para que sejam decididas as ações que devem ser executadas.
 - **(3)** Realiza a ação escolhida e informa à base de conhecimento sobre a ação que está sendo realizada.

Agente Baseado em Conhecimento

- Porque utilizar uma linguagem lógica de representação de conhecimento?
 - **Facilita a criação dos agentes.** É possível dizer o que o agente sabe por meio de sentenças lógicas.
 - O agente pode **adicionar** novas sentenças a sua base de conhecimento enquanto ele explora o ambiente.
 - Abordagens de construção de sistemas
 - **Declarativa:** projetista informa sentenças uma a uma até que o agente saiba operar
 - **Procedural:** comportamentos desejados diretamente como código de programa

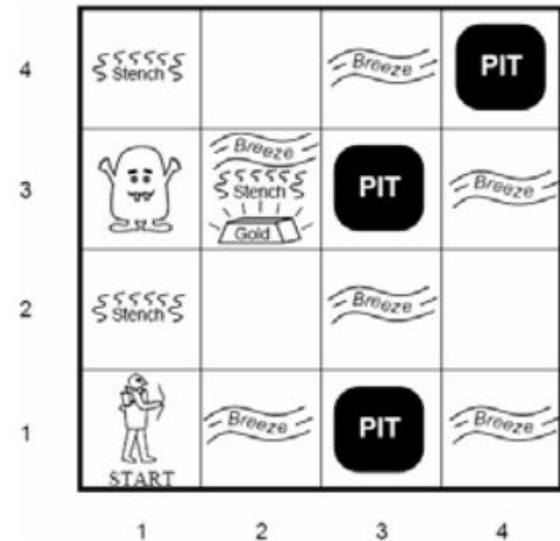
O Mundo de Wumpus

- **O ambiente contém:**
 - Salas conectadas por passagens;
 - Ouro em alguma sala;
 - Poços sem fundo nos quais cairá qualquer um que passar pela sala, exceto o Wumpus;
 - Wumpus: monstro que devora qualquer guerreiro que entrar em sua sala. O Wumpus pode ser morto pelo agente, mas o agente só tem uma flecha.



O Mundo de Wumpus

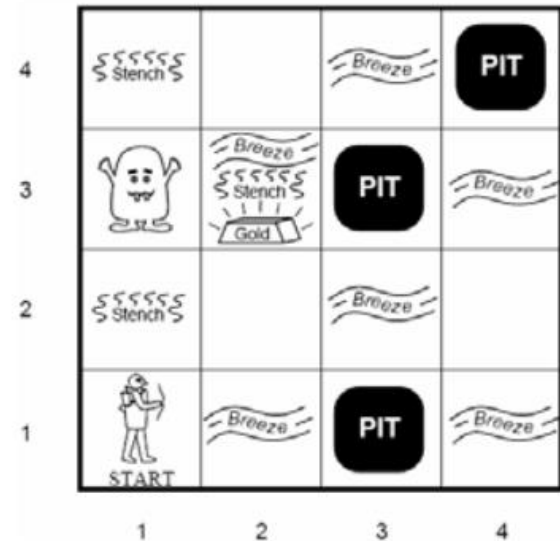
- **Medida de desempenho:** +1.000 por pegar ouro, -1.000 se cair em um poço ou for devorado pelo Wumpus, -1 para cada ação executada, -10 pelo uso da flecha.
- **Ambiente:** malha 4x4 de salas. O agente sempre começa no quadrado identificado como [1,1] voltado para à direita. As posições do Wumpus, ouro e poços são escolhidas aleatoriamente.
- **Ações possíveis:** O agente pode mover-se para frente, virar à esquerda, virar à direita, agarrar um objeto e atirar a flecha.



O Mundo de Wumpus

- **Sensores:**

- Em quadrados adjacentes ao Wumpus, exceto diagonal, o agente sente o **fedor** do Wumpus;
- Em quadrados adjacentes a um poço, exceto diagonal, o agente sente uma **brisa**;
- Quadrados onde existe ouro o agente percebe o **brilho** do ouro;
- Ao caminhar contra uma parede o agente sente um **impacto**;
- Quando o Wumpus morre o agente ouve um **grito**;



[fedor, brisa, brilho, impacto, grito]

O Mundo de Wumpus

- **Passo 1:**

- **Sensores:**

- [nada, nada, nada, nada, nada]

- **Conclusão:**

- [1,2] e [2,1] são seguros

- **Movimento escolhido:**

- [2,1]

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 OK	2,2	3,2	4,2
1,1 A OK	2,1 OK	3,1	4,1

O Mundo de Wumpus

- **Passo 2:**

- **Sensores:**

- [nada, brisa, nada, nada, nada]

- **Conclusão:**

- Há poço em [2,2], [3,1] ou ambos

- **Movimento escolhido:**

- [1,1] e depois [1,2]

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 OK	2,2 P?	3,2	4,2
1,1 V OK	2,1 A B OK	3,1 P?	4,1

O Mundo de Wumpus

- **Passo 3:**

- **Sensores:**

- [fedor, nada, nada, nada]

- **Conclusão:**

- Há Wumpus em [1,3] ou [2,2]

- Wumpus não pode estar em [2,2]

- Wumpus em [1,3]

- Não existe poço em [2,2]

- Poço em [3,1]

- [2,2] é seguro

- **Movimento escolhido:**

- [2,2]

1,4	2,4	3,4	4,4
1,3 W!	2,3	3,3	4,3
1,2 A S OK	2,2 OK	3,2	4,2
1,1 V OK	2,1 B V OK	3,1 P!	4,1

Lógica

- A base de conhecimento de um agente é formada por um conjunto de sentenças expressas por meio de uma **linguagem lógica de representação de conhecimento**.
- O conceito de lógica por Aristóteles. “É o conhecimento das formas gerais e regras gerais do pensamento correto e verdadeiro, independentemente dos conteúdos pensados”
 - Silogismo: “Todo homem é mortal”
 - “Sócrates é um homem”
 - “Logo, Sócrates é mortal”
- Todo X é Y. Z é X. Portanto, Z é Y.

Tipos de Lógica

- **Lógica proposicional:** (ou lógica Booleana) lógica que representa a estrutura de sentenças usando conectivos como: “e”, “ou” e “não”.
- **Lógica de predicados:** lógica que representa a estrutura de sentenças usando conectivos como: “alguns”, “todos” e “nenhum”.
- **Lógica multivalorada:** estende os tradicionais valores verdadeiro/falso para incluir outros valores como “possível” ou um número infinito de “graus de verdade”, representados, por exemplo, por um número real entre 0 e 1.
- Entre outras:
 - **Lógica modal**
 - **Lógica temporal**
 - **Lógica paraconsistente**
 - ...

Conceitos de Lógica

- **Sintaxe:** especifica todas as sentenças que são bem-formadas.
 - Exemplo na aritmética: “ $x+y=4$ ”, “ $x4y+=$ ”.
- **Semântica:** Especifica o significado das sentenças. A verdade de cada sentença com relação a cada “mundo possível”.
 - Exemplo: a sentença “ $x+y=4$ ” é verdadeira em um mundo no qual $x=2$ e $y=2$, mas é falsa em um mundo em que $x=1$ e $y=1$.

Conceitos de Lógica

- **Modelo:** designação de um “mundo possível”.
 - A frase “ m satisfaz α ” (ou “ m é um modelo de α ”) indica que a sentença α é verdadeira no modelo m .
- **Consequência lógica:** utilizada quando uma sentença decorre logicamente de outra.
 - Notação matemática: $\alpha \models \beta$
(a sentença α tem como consequência lógica a sentença β).
 - Pode ser aplicada para derivar conclusões, ou seja, para conduzir inferência lógica.

Consequência lógica no Mundo de Wumpus

- **Base de conhecimento:**

Nada em [1,1];

Brisa em [2,1];

Regras do mundo de Wumpus;

- **Interesse do agente:**

Saber se os quadrados [1,2],

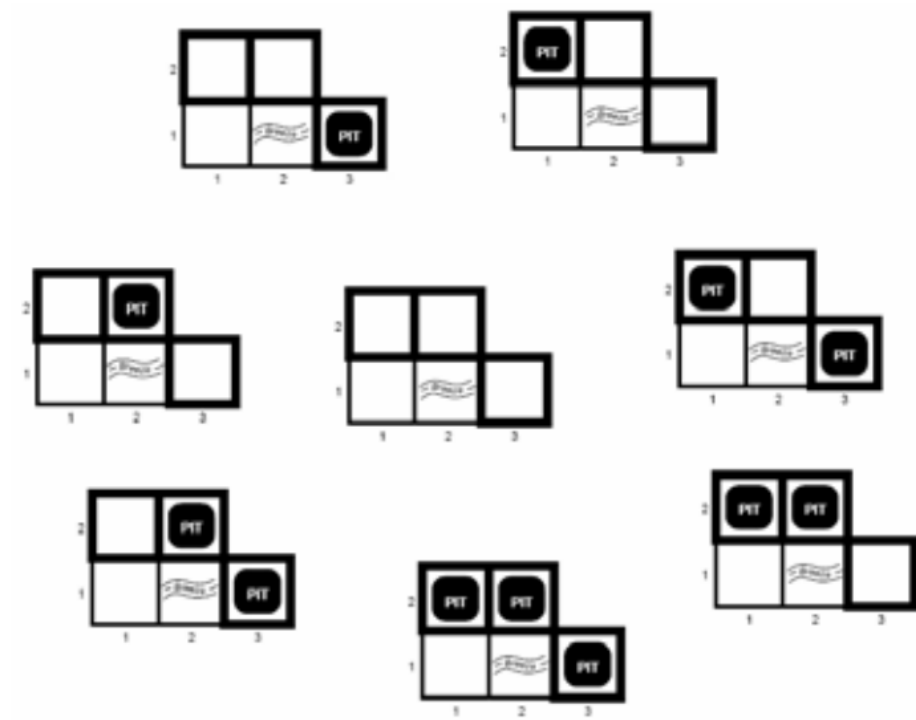
[2,2] e [3,1] contêm poços.

- **Possíveis modelos:**

$$2^3=8$$

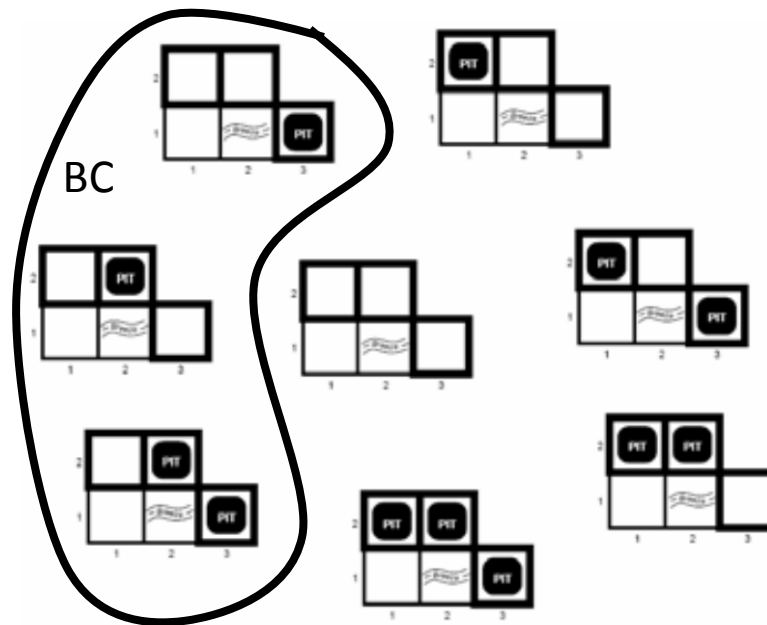
1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 OK	2,2 P?	3,2	4,2
1,1 V OK	2,1 <input type="checkbox"/> A B OK	3,1 P?	4,1

Possíveis Modelos



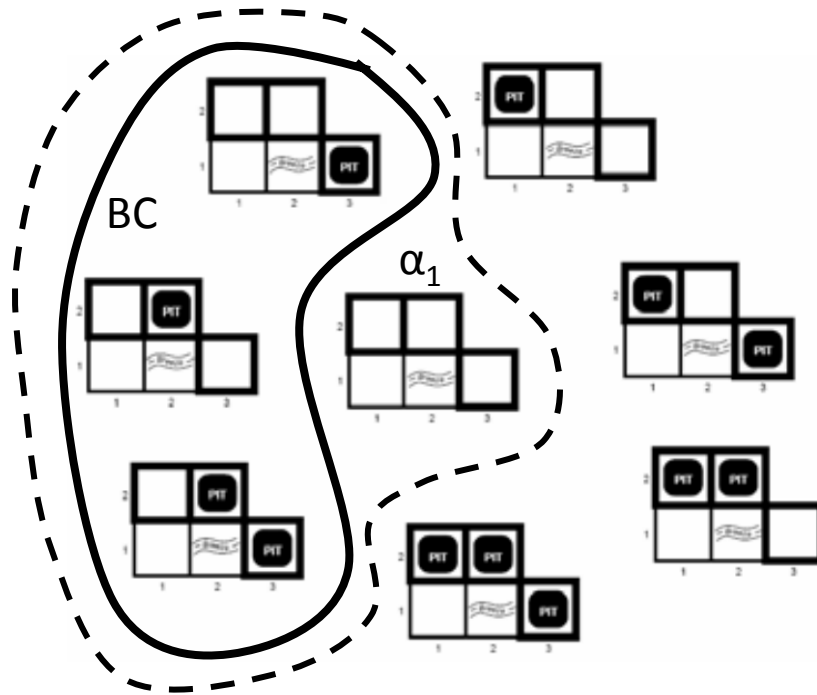
Consequência lógica no Mundo de Wumpus

- A base de conhecimento (BC ou KB) é falsa em modelos que contradizem o que o agente sabe. Nesse caso, há apenas 3 modelos em que a base de conhecimento é verdadeira:



Consequência lógica no Mundo de Wumpus

- Considerando a possível conclusão:
 - α_1 = “não existe nenhum poço em [1,2]”



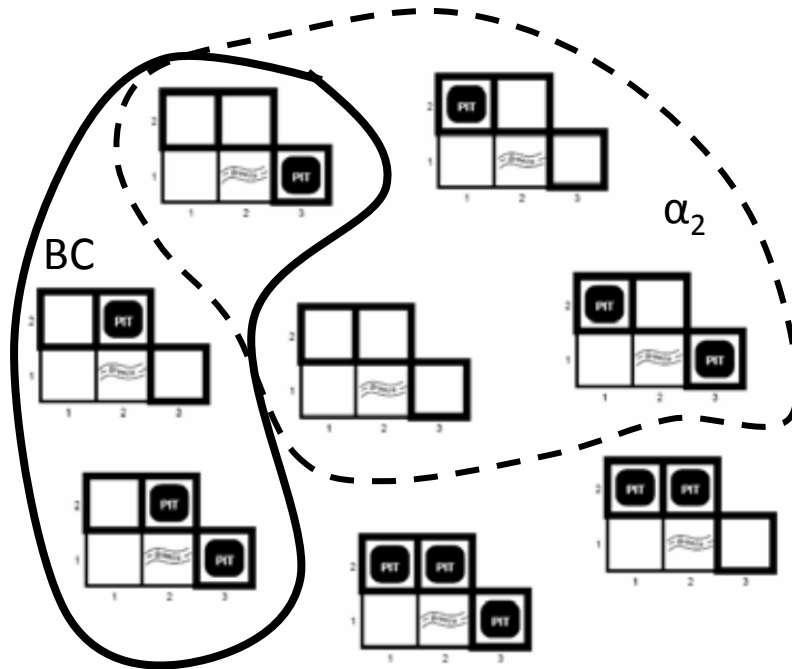
Em todo modelo no qual BC é verdadeira, α_1 também é verdadeira.

Consequentemente
 $BC \models \alpha_1$

não existe nenhum poço em [1,2]

Consequência lógica no Mundo de Wumpus

- Considerando a possível conclusão:
 - α_2 = “não existe nenhum poço em [2,2]”



Em alguns modelos nos quais BC é verdadeira, α_2 é falsa.

Consequentemente

$$BC \not\models \alpha_2$$

não pode concluir que não existe poço em [2,2]

Inferência Lógica

- **O exemplo anterior:**
 - Ilustra a **consequência lógica**.
 - Mostra como a consequência lógica pode ser aplicada para produzir **inferência lógica** (derivar conclusões).
 - O algoritmo ilustrado no exemplo se chama **verificação de modelos (model checking)**. Ele numera todos os possíveis modelos para checar se α é verdade em todos os modelos onde BC é verdade.

Agente Baseado em Conhecimento

- Como representar a base de conhecimento do agente?
 - Lógica Proposicional
 - Lógica de Primeira Ordem
 - Outras linguagens lógicas

Lógica Proposicional

- Lógica simples.
- As sentenças são formadas por conectivos como: “e”, “ou”, “então”.
- É necessário definir:
 - **Sintaxe** (sentenças válidas).
 - **Semântica** (modo pelo qual a verdade das sentenças é determinada).
 - **Consequência lógica** (relação entre uma sentença e outra que decorre dela).
 - **Algoritmo para inferência lógica.**

Sintaxe em Lógica Proposicional

- A sintaxe da lógica proposicional define as sentenças permitidas. É formada por:
 - **Símbolos:** nomes em letras maiúsculas (P, Q, R, ...) que podem assumir verdadeiro e falso;
 - **Sentenças atômicas:** constituídas por elementos sintáticos indivisíveis (símbolo proposicional);
 - **Sentenças complexas:** são construídas a partir de sentenças mais simples com a utilização de conectivos lógicos: \neg (não), \wedge (e), \vee (ou), \Rightarrow (implicação: então), \Leftrightarrow (dupla implicação: se e somente se)
 - Sentença cujo principal conectivo é \wedge : conjunção
 - Sentença cujo principal conectivo é \vee : disjunção

Gramática da Lógica Proposicional

- **Sentença** \rightarrow SentençaAtômica | SentençaComplexa
- **SentençaAtômica** \rightarrow Verdadeiro | Falso | Símbolo
- **Símbolo** \rightarrow P | Q | R | ...
- **SentençaComplexa** \rightarrow \neg Sentença
| (Sentença \wedge Sentença)
| (Sentença \vee Sentença)
| (Sentença \Rightarrow Sentença)
| (Sentença \Leftrightarrow Sentença)

Exemplos de Sentenças Válidas

- P
- Verdadeiro
- $P \wedge Q$
- $(P \vee Q) \Rightarrow S$
- $(P \wedge Q) \vee R \Rightarrow S$
- $\neg(P \vee Q)$
- $\neg(P \vee Q) \Rightarrow R \wedge S$

Implicação Lógica (\Rightarrow)

- $P \Rightarrow Q$
 - Se P é verdade então Q também é verdade.
 - Exemplo:
 - Se está chovendo então as ruas estão molhadas.

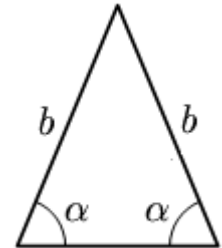
Equivalência Lógica (\Leftrightarrow)

- $P \Leftrightarrow Q$

- Se P é verdade então Q também é verdade. Se Q é verdade então P também é verdade.

- Exemplo:

- Se dois lados de um triângulo são iguais então os dois ângulos da base do triângulo são iguais (triângulo isósceles).



- A equivalência pode ser substituída por duas sentenças de implicação: $(P \Rightarrow Q) \wedge (Q \Rightarrow P)$

Semântica em Lógica Proposicional

- Descreve como calcular o valor verdade de qualquer sentença com base em um mesmo **modelo**. É necessário definir como calcular a verdade de sentenças atômicas e como calcular a verdade de sentenças formadas com cada um dos cinco conectivos (\neg , \wedge , \vee , \Rightarrow , \Leftrightarrow).
- **Sentenças atômicas:**
 - Verdadeiro é verdadeiro e falso é falso em todo modelo.
 - O valor-verdade de todos os outros símbolos proposicionais deve ser especificado diretamente no modelo.
- **Sentenças complexas:**
 - As regras em cada conectivo são resumidas em uma **tabela-verdade**.

Tabela-verdade para os Conectivos

- Para os cinco conectivos lógicos apresentados, teremos:

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \Rightarrow Q$	$P \Leftrightarrow Q$
F	F	V	F	F	V	V
F	V	V	F	V	V *	F
V	F	F	F	V	F	F
V	V	F	V	V	V	V

- (*) Lógica proposicional não exige relação de causa e efeito entre P e Q. Deve-se entender esta relação como “se P é verdadeira, então Q é verdadeira. Caso contrário, não estou fazendo nenhuma afirmação”. Exemplo:
 - “5 é ímpar implica que Tóquio é capital do Japão” (V)
 - “5 é par implica que João é inteligente” (V) (independente se João é ou não inteligente)

Exemplo: Mundo de Wumpus

- Vocabulário de símbolos proposicionais:
 - Seja $P_{i,j}$ verdadeiro se existe poço em $[i,j]$
 - Seja $B_{i,j}$ verdadeiro se existe brisa em $[i,j]$

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 OK	2,2 P?	3,2	4,2
1,1 V OK	2,1 A B OK	3,1 P?	4,1

Exemplo: Mundo de Wumpus

Base de Conhecimento:

R1: $\neg P_{1,1}$



Não há poço em [1,1].

R2: $B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$



Um quadrado tem uma brisa se e somente se existe um poço em um quadrado vizinho (todos os quadrados devem ser declarados).

R3: $B_{2,1} \Leftrightarrow (P_{1,1} \vee P_{2,2} \vee P_{3,1})$

R4: $\neg B_{1,1}$



Percepções adquiridas pelo agente do mundo em que ele se encontra.

R5: $B_{2,1}$

1,2 OK	2,2 P?	3,2
1,1 V OK	2,1 A B OK	3,1 P?

Inferência - Mundo de Wumpus

- **Inferência:** derivação de novas sentenças a partir de sentenças antigas.
- **Objetivo:** decidir se $BC \models \alpha$ para alguma sentença α . Exemplos:
 $P_{1,2}$? $P_{2,2}$?
- **Algoritmo:** enumerar todos os modelos e verificar se α é verdadeira em todo modelo no qual BC é verdadeira.
 - Símbolos proposicionais relevantes:
 $B_{1,1}, B_{2,1}, P_{1,1}, P_{1,2}, P_{2,1}, P_{2,2}, P_{3,1}$
 - 7 símbolos $\rightarrow 2^7=128$ modelos possíveis

Inferência - Mundo de Wumpus

Base de Conhecimento:

R1: $\neg P_{1,1}$



Não há poço em [1,1].

R2: $B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$



Um quadrado tem uma brisa se e somente se existe um poço em um quadrado vizinho (todos os quadrados devem ser declarados).

R3: $B_{2,1} \Leftrightarrow (P_{1,1} \vee P_{2,2} \vee P_{3,1})$

R4: $\neg B_{1,1}$



Percepções adquiridas pelo agente do mundo em que ele se encontra (exemplo específico).

R5: $B_{2,1}$

1,2 OK	2,2 P?	3,2
1,1 V OK	2,1 A B OK	3,1 P?

Tabela Verdade – Mundo de Wumpus

$B_{1,1}$	$B_{2,1}$	$P_{1,1}$	$P_{1,2}$	$P_{2,1}$	$P_{2,2}$	$P_{3,1}$	R1	R2	R3	R4	R5	BC
F	F	F	F	F	F	F	V	V	V	V	F	F
F	F	F	F	F	F	V	V	V	F	V	F	F
:	:	:	:	:	:	:	:	:	:	:	:	:
F	V	F	F	F	F	F	V	V	F	V	V	F
F	V	F	F	F	F	V	V	V	V	V	V	V
F	V	F	F	F	V	F	V	V	V	V	V	V
F	V	F	F	V	F	F	V	F	F	V	V	F
:	:	:	:	:	:	:	:	:	:	:	:	:
V	V	V	V	V	V	V	F	V	V	F	V	F

- Em três desses modelos toda a base de conhecimento é verdadeira.
- Nesses três modelos, $\neg P_{1,2}$ é verdadeira. Dessa maneira conclui-se que **não existe poço em [1,2]**.
- $P_{2,2}$ é verdadeira em dois dos três modelos e falsa em um. Assim, **não podemos dizer ainda se existe um poço em [2,2]**.

Equivalência

$(\alpha \wedge \beta) \equiv (\beta \wedge \alpha)$ comutatividade de \wedge

$(\alpha \vee \beta) \equiv (\beta \vee \alpha)$ comutatividade de \vee

$(\alpha \wedge \beta) \wedge \gamma \equiv \alpha \wedge (\beta \wedge \gamma)$ associatividade de \wedge

$(\alpha \vee \beta) \vee \gamma \equiv \alpha \vee (\beta \vee \gamma)$ associatividade de \vee

$\neg\neg\alpha \equiv \alpha$ eliminação de dupla negação

$(\alpha \Rightarrow \beta) \equiv (\neg\beta \Rightarrow \neg\alpha)$ contraposição

$(\alpha \Rightarrow \beta) \equiv (\neg\alpha \vee \beta)$ eliminação de implicação

$(\alpha \Leftrightarrow \beta) \equiv ((\alpha \Rightarrow \beta) \wedge (\beta \Rightarrow \alpha))$ eliminação de bicondicional

$\neg(\alpha \wedge \beta) \equiv (\neg\alpha \vee \neg\beta)$ de Morgan

$\neg(\alpha \vee \beta) \equiv (\neg\alpha \wedge \neg\beta)$ de Morgan

$(\alpha \wedge (\beta \vee \gamma)) \equiv ((\alpha \wedge \beta) \vee (\alpha \wedge \gamma))$ distributividade de \wedge sobre \vee

$(\alpha \vee (\beta \wedge \gamma)) \equiv ((\alpha \vee \beta) \wedge (\alpha \vee \gamma))$ distributividade de \vee sobre \wedge

Padrões de Raciocínio em Lógica Proposicional

- **Modus Ponens** (modo que afirma): A partir de uma implicação (\Rightarrow ou \rightarrow) e a premissa da implicação, pode-se inferir a conclusão.
- **Eliminação de E**: De uma conjunção, pode-se inferir qualquer um dos conjuntores.
- **Resolução Unitária**: De uma disjunção, se um dos disjuntores é falso, então pode-se inferir que o outro é verdadeiro.

$$\frac{\alpha \rightarrow \beta, \alpha}{\beta}$$

$$\frac{\alpha \wedge \beta}{\alpha}$$

$$\frac{\alpha \vee \beta, \neg \beta}{\alpha}$$

De Volta ao Mundo de Wumpus

Base de Conhecimento:

R1: $\neg P_{1,1}$



Não há poço em [1,1].

R2: $B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$

R3: $B_{2,1} \Leftrightarrow (P_{1,1} \vee P_{2,2} \vee P_{3,1})$



Um quadrado tem uma brisa se e somente se existe um poço em um quadrado vizinho (todos os quadrados devem ser declarados).

R4: $\neg B_{1,1}$

R5: $B_{2,1}$



Percepções adquiridas pelo agente do mundo em que ele se encontra (exemplo específico).

1,2 OK	2,2 P?	3,2
1,1 V OK	2,1 A B OK	3,1 P?

Provando $\neg P_{1,2}$ em Wumpus

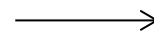
- Eliminação bicondicional em **R2**:

$$\mathbf{R2: } B_{1,1} \Leftrightarrow (P_{1,2} \vee P_{2,1})$$

$$\mathbf{R6: } (B_{1,1} \Rightarrow (P_{1,2} \vee P_{2,1})) \wedge ((P_{1,2} \vee P_{2,1}) \Rightarrow B_{1,1})$$

- Eliminação de “e” em **R6**:

$$\mathbf{R7: } (P_{1,2} \vee P_{2,1}) \Rightarrow B_{1,1}$$



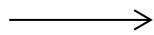
De uma conjunção, pode-se inferir qualquer um dos conjutores.

- Contraposição em **R7**:

$$\mathbf{R8: } \neg B_{1,1} \Rightarrow \neg(P_{1,2} \vee P_{2,1})$$

- Modus Ponens (**R4** + **R8**)

$$\mathbf{R4: } \neg B_{1,1}$$



A partir de uma implicação e a premissa da implicação, pode-se inferir a conclusão.

$$\mathbf{R9: } \neg(P_{1,2} \vee P_{2,1})$$

- Regra de Morgan em **R9**:

$$\mathbf{R10: } \neg P_{1,2} \wedge \neg P_{2,1}$$

- Eliminação de “e” em **R10**: $\neg P_{1,2}$

1,2 OK	2,2 P?	3,2
1,1 V OK	2,1 A B OK	3,1 P?

Prova Lógica

- A aplicação de uma sequência de regras de inferências para derivar uma conclusão é chamado de **prova lógica**.
- A aplicação de inferências lógicas é uma **alternativa à enumeração de modelos** vista anteriormente.
- Como saber quais regras de inferência devem ser utilizadas?

Limitações da Lógica Proposicional

- A lógica proposicional é **simples demais** para representar alguns problemas do mundo real.
- Em problemas complexos pode ser necessária a utilização de um **número muito grande de sentenças** para a criação de um agente realmente inteligente.