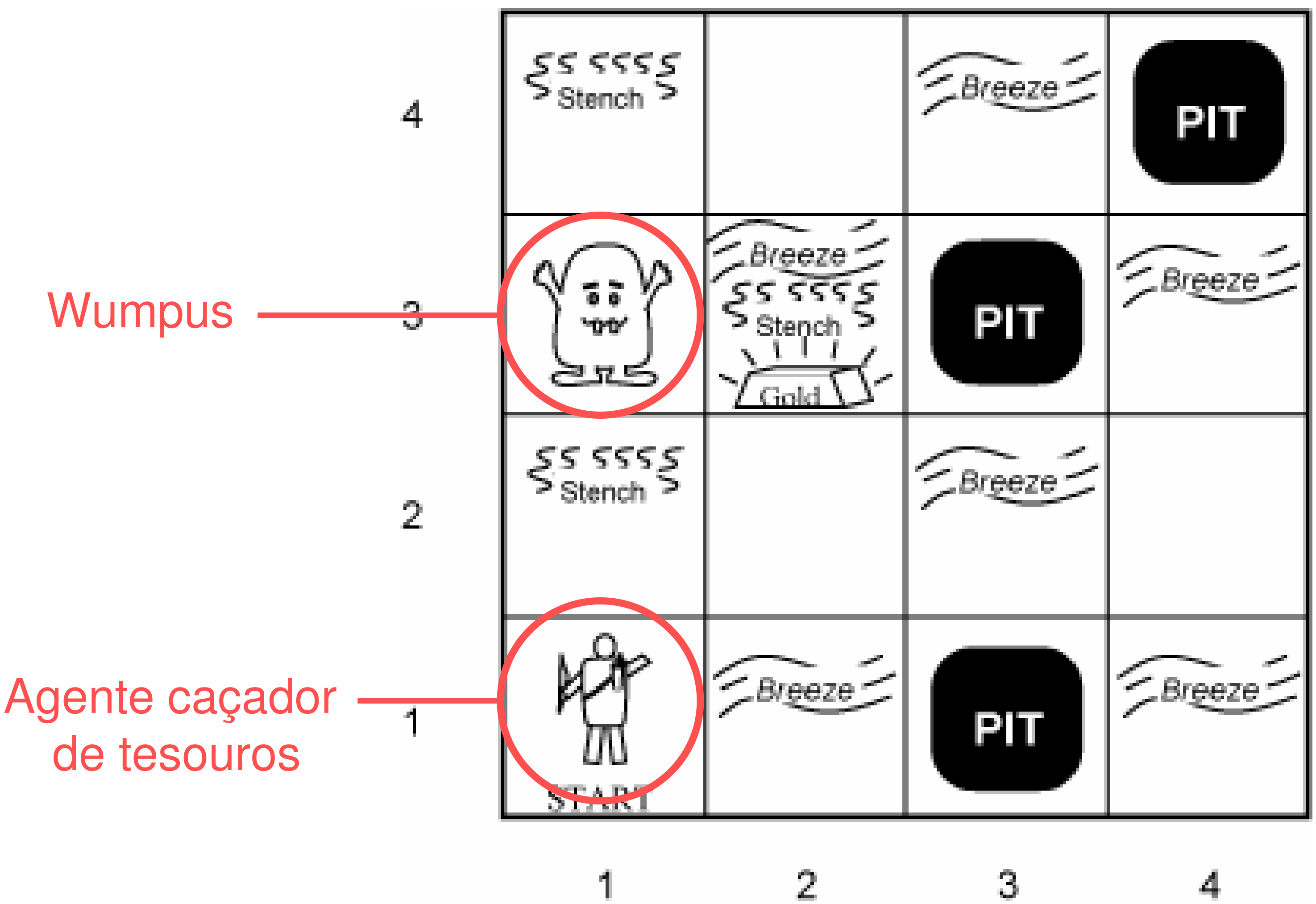
Inteligência Artificial

Prof. Tiago A. E. Ferreira

Aula 15 – Agentes que Raciocinam Logicamente

Bem-vindos ao “Mundo do Wumpus”



Codificação do Mundo do Wumpus

4	fedor		brisa	**B**
3	**W**	**O** fedor brisa, luz	**B**	brisa
2	fedor		brisa	
1	**A** início	brisa	**B**	brisa
	1	2	3	4

A - Agente
W - Wumpus
B - Buraco
O - Ouro

O Mundo do Wumpus: formulação do problema

- Ambiente:
  - agente, Wumpus, cavernas, buracos, ouro
- Estado inicial:
  - agente na caverna (1,1) com apenas uma flecha
  - Wumpus e buracos em cavernas quaisquer
- Objetivos:
  - pegar a barra de ouro e voltar à caverna (1,1) com vida
- Percepções:
  - fedor, brisa, luz, choque (contra a parede da caverna) e grito do Wumpus
- Ações:
  - *avançar* para próxima caverna
  - *girar* 90 graus à *direita* ou à *esquerda*
  - *pegar* um objeto na mesma caverna que o agente
  - *atirar* na direção para onde o agente está olhando (a flecha pára quando encontra uma parede ou mata o Wumpus)
  - *sair* da caverna

Raciocinando e Agindo no Mundo do Wumpus

- Conhecimento do agente:

  (a) no início do jogo, depois de receber sua primeira percepção , e

  (b) depois do 1o movimento, com a seqüência de percepções [nada,brisa,nada,nada,nada]

V - caverna visitada

4				
3				
2	ok			
1	**A** ok	ok		
	1	2	3	4

4				
3				
2	ok	B?		
1	**V** ok	b **A** ok	B?	
	1	2	3	4

Raciocinando e Agindo no Mundo do Wumpus

- Estando em (2,2), o agente move-se para (2,3) e encontra o ouro!!!

4				
3	W!			
2	f A ok	ok		
1	V ok	b V ok	B!	
	1	2	3	4

4		B?		
3	W!	A f b l	B?	
2	f V ok	V ok		
1	V ok	b V ok	B!	
	1	2	3	4

Representação de Conhecimento usando a Lógica

Representação: Semântica

- Uma *sentença lógica* não significa nada por si só...

- É necessário estabelecer a *correspondência* entre *fatos* e *sentenças*, fixando seu significado através de uma <u>*interpretação*</u> da sentença.

- Exemplo:

  "O Papa já está no Rio"

  - mensagem secreta trocada entre dois agentes do FBI que significa que os documentos sobre as armas atômicas da Rússia (o Papa) foram entregues ao Pentágono (o Rio) a salvo (já está).

Interpretação e validade

- Interpretação
  - uma sentença é ***verdadeira*** sob uma dada interpretação se o "estado do mundo" (*state of affairs*) que ela representa se verifica.
  - *Valor verdade* depende da *interpretação* da frase + do *estado atual* do mundo.
- Exemplo
  - "O papa está no Rio" pode ser verdade na interpretação anteriormente dada se de fato, no mundo do FBI, tais documentos foram recebidos pelo Pentágono a salvo.
  - "O papa está no Rio" , sob a interpretação "Papa = João Paulo II", "Rio = Cidade do Rio de Janeiro", "está = verbo estar", é falsa pois ele está em Roma.
- Nesta ótica, uma sentença pode ser:
  - válida, satisfazível ou insatisfazível.

Sentença Satisfazível

- Uma sentença é válida (tautologia)
  - se ela é verdade sob todas as possíveis interpretações em todos os mundos possíveis.
    - Ex. “existe um buraco em (1,2), ou não existe um buraco em (1,2)” é sempre verdade, independente da interpretação e do valor-verdade de “existe um buraco em (1,2)” em qualquer mundo possível.
- Uma sentença é satisfazível
  - se existe alguma interpretação em algum mundo sob a qual ela é verdade .
  - Caso contrário, a sentença é dita **não satisfazível** ou **insatisfazível**.
    - Exemplo: sentenças contraditórias são insatisfazíveis quando a contradição não depende da interpretação dos símbolos:
    - “existe um buraco em (1,2), e não existe um buraco em (1,2)”

Raciocínio: Inferência em Computadores

- Computadores têm conhecimento limitado sobre o mundo:
  - não sabem que interpretação foi dada às sentenças na Base de Conhecimento (BC), e
  - não sabem nada sobre o mundo, apenas o que existe na BC.
- Então, como responder à pergunta
  - "Está OK mover o agente para (2,2)?"
- Verificando se a sentença abaixo é *implicada* a partir da BC
  - "(2,2) está OK".
- O procedimento de inferência deve mostrar que a sentença abaixo é <u>*válida*</u>
  - "Se a BC é verdade, então (2,2) está OK"

Propriedades da inferência

- A inferência pode ter várias propriedades...
  - Corretude, completude, composicionalidade, monotonicidade, localidade, eficiência, etc.
- Corretude (sound)
  - gera *apenas* sentenças válidas
- Completude
  - gera *todas* as sentenças válidas
- Composicionalidade:
  - o significado de uma sentença é uma função do significado de suas partes.
    - B1-2 = existe um buraco na caverna (1,2)
    - B2-3 = existe um buraco na caverna (2,3)
    - B1-2 e B2-3 = hoje é feriado, dia do funcionário público

Lógica: Inferência

- Uma Lógica é dita monotônica quando
  - Tudo que era verdade continua sendo depois de uma inferência
  - se BC1 |= a então (BC1 U BC2) |= a
  - todas as sentenças implicadas pela BC original são ainda implicadas pela BC aumentada pelas novas sentenças inferidas
    - e.g., Lógica Proposicional e de Primeira Ordem.
    - contra-exemplo: Teoria da Probabilidade
- Localidade:
  - regras são ditas locais porque sua premissa só necessita ser comparada com uma pequena porção da BC.
  - Localidade só é possível devido à monotonicidade, uma vez que esta garante que o resto da BC não vai afetar a corretude da inferência.

Lógica: Inferência

- *Localidade* **e** *composicionalidade* **são centrais na construção de sistemas por possibilitar** *modularidade*.
- Modularidade **favorece a** *reusabilidade* **e a** *extensibilidade* **do sistema.**

Formalização de Agentes baseados em Lógica Proposicional

Validade de sentenças

- A validade pode ser verificada de duas maneiras
  - Tabelas-Verdade
  - Regras de inferência
- Tabelas-Verdade
  - ex. Validade de $((P \vee H) \wedge \neg H) \Rightarrow P$?

P	**H**	$(P \vee H)$	$((P \vee H) \wedge \neg H)$	$((P \vee H) \wedge \neg H) \Rightarrow P$
F	F	F	F	T
F	T	T	F	T
T	F	T	T	T
T	T	T	F	T

Lógica Proposicional: Regras de Inferência

- Modus Ponens: $\dfrac{\alpha \Rightarrow \beta, \alpha}{\beta}$
- E-eliminação: $\dfrac{\alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_n}{\alpha_i}$
- E-introdução: $\dfrac{\alpha_1, \alpha_2, \ldots, \alpha_n}{\alpha_1 \wedge \alpha_2 \wedge \ldots \wedge \alpha_n}$
- Ou-introdução: $\dfrac{\alpha_i}{\alpha_1 \vee \alpha_2 \vee \ldots \vee \alpha_n}$
- Eliminação de dupla negação: $\dfrac{\neg\neg\alpha}{\alpha}$
- Resolução unidade: $\dfrac{\alpha \vee \beta, \neg \beta}{\alpha}$
- Resolução: $\dfrac{\alpha \vee \beta, \neg\beta \vee \gamma}{\alpha \vee \gamma} \Leftrightarrow \dfrac{\neg\alpha \Rightarrow \beta, \beta \Rightarrow \gamma}{\neg\alpha \Rightarrow \gamma}$

α/β diz que a sentença β pode ser derivada de α por inferência.

Validade de sentenças

- Regras de inferência:
  - capturam padrões de inferências (sintáticos!!!)
  - sempre que algum fato na BC casar com o padrão acima da linha, a regra de inferência conclui o padrão abaixo da linha.
  - uma regra de inferência é correta (preserva a verdade) se a conclusão é verdade em todos os casos onde as premissas são verdadeiras.
- $((P \lor H) \land \neg H) \Rightarrow P$?

$((P \land \neg H) \lor (H \land \neg H)) \Rightarrow P$

$((P \land \neg H) \lor false) \Rightarrow P$

$(P \land \neg H) \Rightarrow P$

$\neg(P \land \neg H) \lor P$

$\neg P \lor H \lor P$

$True \lor H$

$True$

Lógica Proposicional: Modelo

- Qualquer mundo no qual uma sentença é verdade sob uma dada interpretação é chamado de modelo (**da sentença sob essa interpretação).**
  - Ex: o Mundo de Wumpus é um modelo da sentença "B1-2" sob a interpretação de que existe um buraco na caverna (1,2).
  - Podem existir muitos modelos para "B1-2", basta que eles tenham um buraco em (1,2).
- Modelos são muito importantes em lógica porque:
  - uma sentença ***a*** é implicada por uma BC (***BC |= a***) se os modelos da BC são também modelos de ***a***.
  - assim, sempre que a ***BC*** for verdade, ***a*** também será verdade.

Complexidade

- Checar se um conjunto de sentenças é satisfazível é um problema NP-completo
  - tabela verdade para uma sentença envolvendo n símbolos tem 2^n colunas (exponencial!)
- Cláusulas de Horn
  - Classe de sentenças úteis que permitem inferência em tempo polinomial
  - $P1 \wedge P2 \wedge P3 \wedge ... \wedge Pn \Rightarrow Q$
  - é usada em Prolog (linguagem em paradigma lógico)
- 2 casos especiais das cláusulas de Horn
  - se Q é falso, $\neg P1 \vee \neg P2 \vee \neg P3 \vee ... \vee \neg Pn$
  - se n = 1 e P1 = verdadeiro, temos Q (um fato)
    - *Verdadeiro* $\Rightarrow$ *Q* é idêntico a *Q* (já que a sentença como um todo tem que ser V).

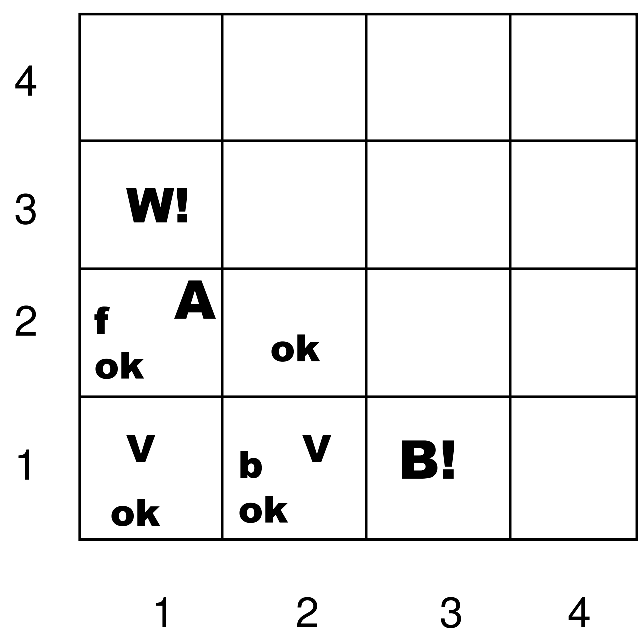
Um Agente-BC para o Mundo do Wumpus

- A *Base de Conhecimento* consiste em:
  - sentenças representando as percepções do agente
  - sentenças válidas implicadas a partir das sentenças das percepções
  - regras utilizadas para implicar novas sentenças a partir das sentenças existentes
- Símbolos:
  - Ax-y significa que "o agente está na caverna (x,y)"
  - Bx-y significa que "existe um buraco na caverna (x,y)"
  - Wx-y significa que "o Wumpus está na caverna (x,y)"
  - Ox-y significa que "o ouro está na caverna (x,y)"
  - bx-y significa que "existe brisa na caverna (x,y)"
  - fx-y significa que "existe fedor na caverna (x,y)"
  - lx-y significa que "existe luz na caverna (x,y)"

Base de Conhecimento para o Mundo do Wumpus

- Com base nas percepções do estado abaixo, a BC deverá conter as seguintes sentenças:

$\neg\, f1\text{–}1$	$\neg\, b1\text{–}1$
$\neg\, f2\text{–}1$	$b2\text{–}1$
$f1\text{–}2$	$\neg\, b1\text{–}2$



V - caverna visitada

Base de Conhecimento para o Mundo do Wumpus

- O agente também tem algum conhecimento prévio sobre o ambiente, *e.g.:*
  - <u>se</u> uma caverna não tem *fedor*, <u>então</u> o Wumpus não está nessa caverna, nem está em nenhuma caverna adjacente a ela.
- O agente terá uma <u>regra</u> para <u>cada</u> caverna no seu ambiente

R1: $\neg f1{-}1 \Rightarrow \neg W1{-}1 \wedge \neg W1{-}2 \wedge \neg W2{-}1$

R2: $\neg f2{-}1 \Rightarrow \neg W1{-}1 \wedge \neg W2{-}1 \wedge \neg W2{-}2 \wedge \neg W3{-}1$

R3: $\neg f1{-}2 \Rightarrow \neg W1{-}1 \wedge \neg W1{-}2 \wedge \neg W2{-}2 \wedge \neg W1{-}3$

- O agente também deve saber que, se existe *fedor* em (1,2), então deve haver um Wumpus em (1,2) ou em alguma caverna adjacente a ela:

R4: $f1{-}2 \Rightarrow W1{-}3 \vee W1{-}2 \vee W2{-}2 \vee W1{-}1$

Como Encontrar o Wumpus - Inferência!

- O Wumpus está em (1,3). Como provar isto?
- **O agente precisa mostrar que** $BC \Rightarrow W1\text{–}3$ **é uma sentença** válida:

$$2^{12} = 4096$$

(1) construindo a Tabela-Verdade para a sentença

- existem 12 símbolos proposicionais na BC, então a Tabela-Verdade terá 12 colunas

(2) usando regras de inferência!

Como Encontrar o Wumpus - Inferência!

- Inicialmente, vamos mostrar que o Wumpus não está em nenhuma outra caverna, e então concluir, por eliminação, que ele está em (1,3).

1. Aplicando Modus Ponens a $\neg f1-1$ e R1, obtemos:
   $\neg W1-1 \wedge \neg W1-2 \wedge \neg W2-1$

2. Aplicando E-eliminação a (1), obtemos três sentenças isoladas:
   $\neg W1-1 \quad \neg W1-2 \quad \neg W2-1$

3. Aplicando Modus Ponens a $\neg f2-1$ e R2, e em seguida aplicando E-eliminação obtemos:
   $\neg W1-1 \quad \neg W2-1 \quad \neg W2-2 \quad \neg W3-1$

4. Aplicando Modus Ponens a $f1-2$ e R4, obtemos:
   $W1-3 \vee W1-2 \vee W2\text{-}2 \vee W1-1$

Como Encontrar o Wumpus - Inferência!

5. Aplicando Resolução Unidade, onde α é $W1-3 \vee W1-2 \vee W2-2$ e β é $W1-1$ obtemos (do passo 2, temos $\neg W1-1$):

$W1-3 \vee W1-2 \vee W2-2$

6. Aplicando Resolução Unidade, onde α é $W1-3 \vee W1-2$ e β é $W2-2$ obtemos:

$W1-3 \vee W1-2$

7. Aplicando Resolução Unidade, onde α é $W1-3$ e β é $W1-2$ obtemos:

***W**1–3 !!!*

Transformando Conhecimento em Ações

- **O conhecimento inferido deve ser usado para auxiliar o agente a realizar ações.**
- **definir regras que relacionem o estado atual do mundo às ações que o agente pode realizar.**
- Ações:
  - avançar para próxima caverna,
  - girar 90 graus à direita ou à esquerda,
  - pegar um objeto na mesma caverna que o agente,
  - atirar na direção para onde o agente está olhando (a flecha para quando encontra uma parede ou mata o Wumpus),
  - sair da caverna.

Transformando Conhecimento em Ações

- Exemplo de Regra:
  - o agente está na caverna (1,1) virado para a direita, e
  - o Wumpus está na caverna (2,1), então:

    $A1\text{-}1 \wedge Dir \wedge W2\text{-}1 \Rightarrow \neg\, avançar$
- Com essas regras, o agente pode então perguntar à BC que ação ele deve realizar:
  - devo avançar?
  - devo girar para a esquerda?
  - devo atirar?, etc

Agente-BC com Lógica Proposicional

função Agente-BC-Proposicional(*percepção*)
retorna uma *ação*

Tell(*BC*, Percepções-Sentença(*percepção*,*t*))

para cada *ação* **em** *lista de possíveis ações* **faça**
se Ask(*BC*, Pergunta-Ação(*t*,*ação*))
então
t <- t + 1
retorna *ação*

Problemas com o Agente Proposicional

- Lógica Proposicional
  - é capaz de fazer inferências que resultam em ações.
  - Contudo, esta lógica é "fraca", não sendo capaz de lidar com domínios simples como o Mundo de Wumpus...
- Problema: existem proposições demais a considerar
  - ex.: a regra: "não avance se o Wumpus estiver em frente a você" só pode ser representada com um conjunto de 64 regras.
  - Assim, serão necessárias milhares de regras para definir um agente eficiente, e o processo de inferência ficará muito lento.
- Outro problema: domínios dinâmicos!

Problemas com o Agente Proposicional

- **Quando o agente faz seu primeiro movimento, a proposição A(1,1) torna-se falsa, e A(2,1) torna-se verdadeira.**
  - não podemos apenas "apagar" A(1,1) porque o agente precisa saber onde esteve antes.
- **Uma solução é usar símbolos diferentes para a localização do agente a cada tempo *t*, contudo...**
  - isso requer regras dependentes do tempo!
  - a BC tem que ser "reescrita" a cada tempo *t.*
- **Se o agente executar 100 passos, a BC terá 6400 regras apenas para dizer que ele não deve avançar quando o Wumpus estiver em frente a ele.**

Uma Solução: Lógica de Primeira Ordem

- Construir agentes baseados em Lógica de Primeira Ordem.
- Essa lógica representa objetos e relações entre objetos, além das proposições.
- As 6400 regras do agente proposicional serão reduzidas para 1.