

AGENTES BASEADOS EM CONHECIMENTO

LÓGICA DE PRIMEIRA ORDEM

Sumário

Agentes baseados em Conhecimento:

- ◇ O que é um agente baseado em conhecimento (ABC)
- ◇ Bases de conhecimento
- ◇ O mundo “Wumpus”
- ◇ Lógica em geral
- ◇ Revisão de lógica proposicional

Lógica de Primeira Ordem

- ◇ Sintaxe e semântica
- ◇ Manipulação de sentenças

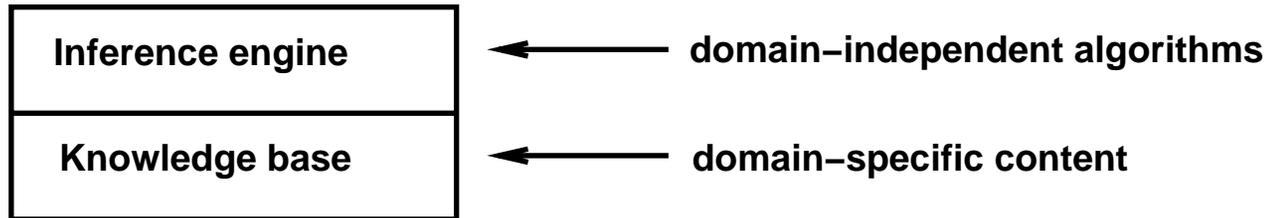
O que é um ABC

Até agora, ignoramos uma metodologia para descrever estados e ações: agentes apenas realizavam uma busca em um espaço de estados e ações pré-determinados.

Vamos agora estudar agentes com a capacidade de adaptar um modelo consistente do mundo à medida que se observam novos estados, e de se deduzir um curso de ações com base nesse modelo.

ABC = Base de conhecimento + mecanismo de inferência

Bases de conhecimento



Base de conhecimento (BC) = conjunto de sentenças em uma linguagem formal (linguagem de representação do conhecimento).

Linguagem formal = um conjunto de *strings* de símbolos de um *alfabeto* (TEORIA DA COMPUTAÇÃO)

Abordagem declarativa para construir um agente (ou outro sistema):

Informe o que se precisa saber; e depois **pergunte** o que fazer — respostas vem da BC.

ABCs: Níveis de Entendimento

Agentes podem ser entendidos ao nível de conhecimento

i.e., o que eles sabem, não interessando **como** foi implementado;

Ou ao nível lógico

i.e., como o conhecimento é codificado em sentenças e manipulado;

Ou ao nível de implementação

i.e., **estruturas** de dados na BC e **algoritmos** que os manipulam

Um agente simples baseado em conhecimento

```
function KB-AGENT(percept) returns an action  
  static: KB, a knowledge base  
           t, a counter, initially 0, indicating time  
  
  TELL(KB, MAKE-PERCEPT-SENTENCE(percept, t))  
  action ← ASK(KB, MAKE-ACTION-QUERY(t))  
  TELL(KB, MAKE-ACTION-SENTENCE(action, t))  
  t ← t + 1  
  return action
```

O agente deve ser capaz de:

Representar estados, ações, etc.

Incorporar novas observações

Atualizar representações internas do mundo

Deduzir propriedades ocultas do mundo

Deduzir ações apropriadas

Deve ser capaz \neq sempre seremos capazes de projetar um capaz ...

Descrição PAGE do mundo “Wumpus”

Percepções Brisa, Brilho, Cheiro

Ações Giro à esquerda, Giro à direita, Adiante, Agarre, Largue, Atire

Objetivos Levar ouro para o ponto inicial sem cair em buracos ou no quadrado do temível “Wumpus”

Ambiente

Quadrados adjacentes ao “Wumpus” têm mau cheiro

Quadrados adjacente a buracos têm brisa

Brilho é observado se e só se o ouro estiver no mesmo quadrado

Tiros matam o “Wumpus” se você estiver cara a cara com ele

Apenas uma bala

Ouro só pode ser agarrado se estiver no mesmo quadrado

O ouro largado cai no mesmo quadrado

Características do mundo Wumpus:

Mundo é determinístico?? Sim—resultados de ações são especificados com exatidão

Mundo é completamente acessível?? Não—percepção é local

Mundo é estático?? Sim—“Wumpus” e buracos não se movem

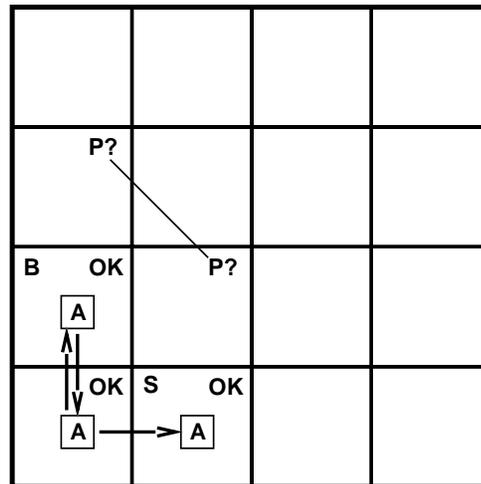
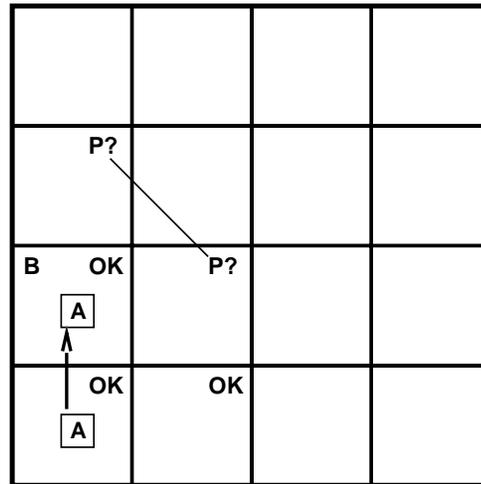
Mundo é discreto?? Sim

Explorando o mundo “Wumpus”

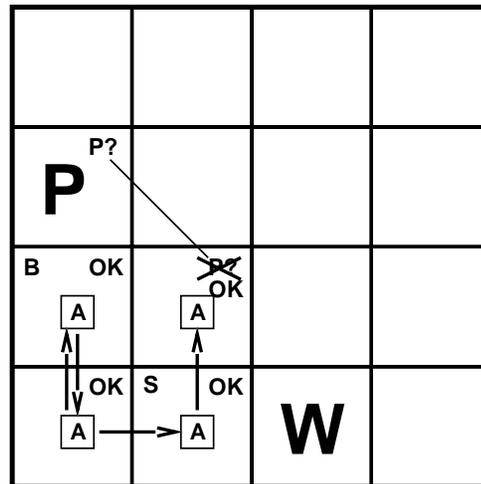
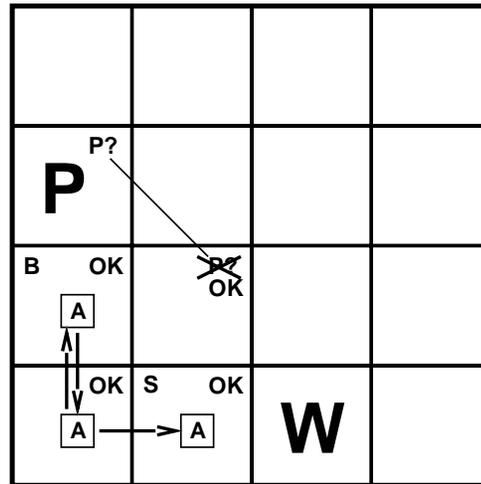
OK			
OK A	OK		

B OK A			
A ↑ OK	OK		

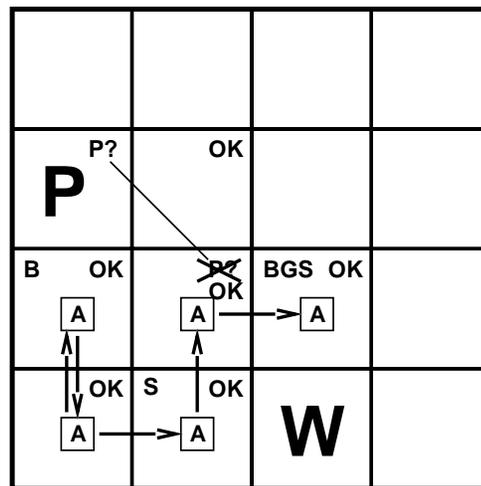
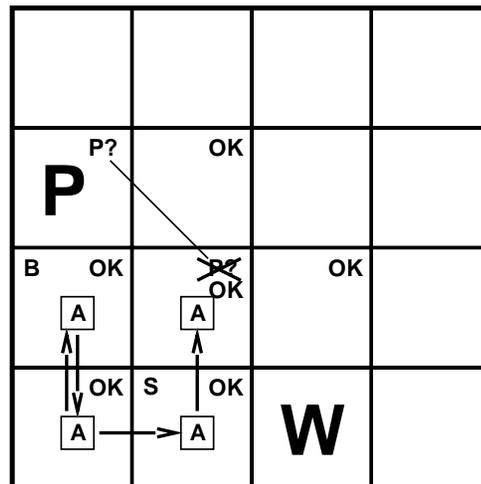
Explorando o mundo “Wumpus”



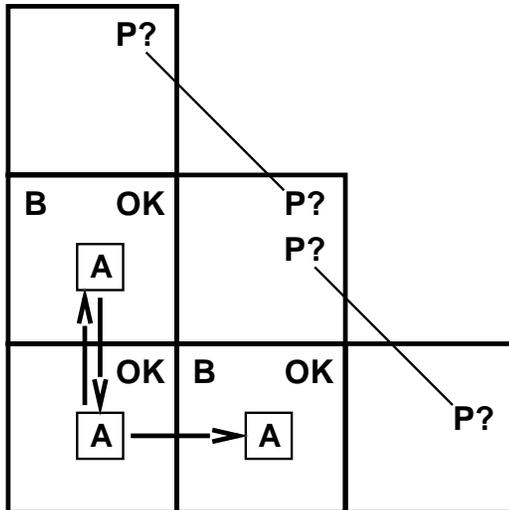
Explorando o mundo “Wumpus”



Explorando o mundo “Wumpus”



Outros pontos críticos

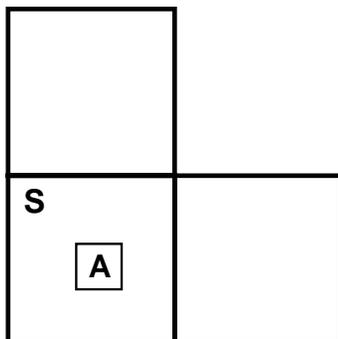


Brisa em (1,2) e (2,1)
 \Rightarrow nenhuma ação segura

Assumindo buracos distribuídos uniformemente,
 (2,2) é provavelmente um buraco

Cheiro em (1,1)
 \Rightarrow não pode se mover

Possa usar uma estratégia coercitiva:
 atiro para a frente
 “Wumpus” estava lá \Rightarrow morto \Rightarrow seguro
 “Wumpus” não estava lá \Rightarrow seguro



Lógica em geral

Uma **lógica** é uma linguagem formal para representar informação a partir da qual podem ser tiradas conclusões.

A sintaxe define as sentenças da linguagem.

A semântica define o “significado” das sentenças; i.e., define a veracidade de uma sentença em um mundo

Por exemplo, na linguagem aritmética usual

$x + 2 \geq y$ é uma sentença; $x^2 + y >$ não é uma sentença

$x + 2 \geq y$ é verdadeiro se o número $x + 2$ não for menor que o número y

$x + 2 \geq y$ é verdadeiro em um mundo onde $x = 7$, $y = 1$

$x + 2 \geq y$ é falso em um mundo onde $x = 0$, $y = 6$

Tipos de lógica

Lógicas são caracterizadas pelo que elas definem como “primitivas”

Compromisso **ontológico**: o que existe—fatos? objetos? tempo? crenças?
Ontologia é o que se assume como parte do mundo do agente.

Compromisso **epistemológico**: que estados de conhecimento? Epistemologia é o que se assume como possíveis estados de conhecimento do agente.

Language	Ontological Commitment	Epistemological Commitment
Propositional logic	facts	true/false/unknown
First-order logic	facts, objects, relations	true/false/unknown
Temporal logic	facts, objects, relations, times	true/false/unknown
Probability theory	facts	degree of belief 0...1
Fuzzy logic	degree of truth	degree of belief 0...1

Implicação Lógica

$$BC \models \alpha$$

A base de conhecimento BC implica ou vincula (*entails*) uma sentença α se e só se α é verdadeira em todos os mundos onde BC é verdadeira.

Diz-se então que α é uma implicação de (ou está vinculada à — *entailed by*) BC .

Por exemplo, a BC que contém “o Flamengo venceu” e “o Vasco empatou” implica “O Flamengo venceu ou o Vasco empatou”

Modelos

Estudiosos de lógica pensam tipicamente em termos de modelos, que são mundos formalmente estruturados, com respeito aos quais veracidade pode ser avaliada

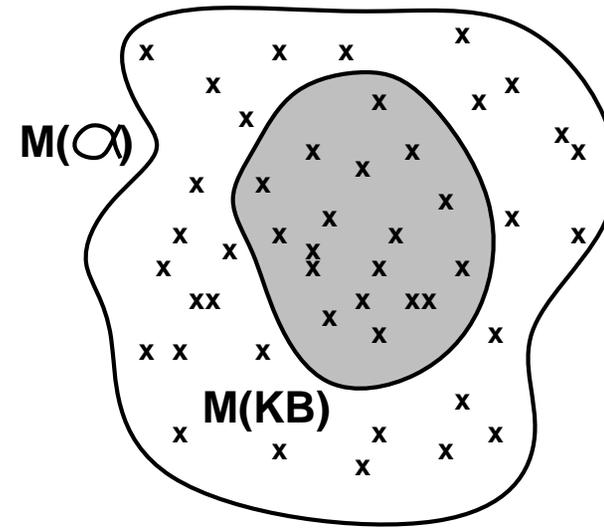
Dizemos que m é um modelo de uma sentença α se α é verdade em m

$M(\alpha)$ é o conjunto de todos os modelos de α

Então $BC \models \alpha$ se e só se $M(BC) \subseteq M(\alpha)$

Por exemplo $BC = \text{Palmeiras venceu e Corinthians venceu}$

$\alpha = \text{Palmeiras venceu}$



Inferência

$BC \vdash_i \alpha$ = sentença α pode ser derivada de BC por um procedimento i

Correção: i é correto (*sound*) se

sempre que $BC \vdash_i \alpha$, também for verdade que $BC \models \alpha$

Um procedimento correto não **inventa** coisas.

Completeza: i é completo se

sempre que $BC \models \alpha$, também for verdade que $BC \vdash_i \alpha$

Um procedimento completo **prova** qualquer sentença implicada pela BC.

Definiremos depois uma lógica (lógica de primeiro ordem) que é expressiva o bastante para dizer quase qualquer coisa de interesse, e para a qual existe um procedimento de inferência **correto e completo**.

O procedimento responderá a qualquer pergunta cuja resposta segue do que é conhecido pelo BC .

Lógica proposicional: Sintaxe

A lógica proposicional é a lógica mais simples—ilustra idéias básicas

Os símbolos proposicionais P_1, P_2 etc são sentenças

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

Se S_1 e S_2 são sentenças, $S_1 \wedge S_2$ é uma sentença

Se S_1 e S_2 são sentenças, $S_1 \vee S_2$ é uma sentença

Se S_1 e S_2 são sentenças, $S_1 \Rightarrow S_2$ é uma sentença

Se S_1 e S_2 são sentenças, $S_1 \Leftrightarrow S_2$ é uma sentença

Lógica proposicional: Semântica

Cada modelo especifica verdadeiro/falso para cada símbolo proposicional

Por exemplo

A	B	C
<i>Verdadeiro</i>	<i>Falso</i>	<i>Falso</i>

Regras para avaliar a verdade com respeito a um modelo m :

$\neg S$	é verdadeiro iff	S	é falso
$S_1 \wedge S_2$	é verdadeiro iff	S_1	é verdadeiro <u>e</u> S_2 é verdadeiro
$S_1 \vee S_2$	é verdadeiro iff	S_1	é verdadeiro <u>ou</u> S_2 é verdadeiro
$S_1 \Rightarrow S_2$	é verdadeiro iff	S_1	é falso <u>ou</u> S_2 é verdadeiro
i.e.,	é falso iff	S_1	é verdadeiro <u>e</u> S_2 é falso
$S_1 \Leftrightarrow S_2$	é verdadeiro iff	$S_1 \Rightarrow S_2$	é verdadeiro <u>e</u> $S_2 \Rightarrow S_1$ é verdadeiro

Inferência proposicional: Método da enumeração

Seja $\alpha = A \vee B$ e $KB = (A \vee C) \wedge (B \vee \neg C)$

É o caso que $KB \models \alpha$?

Confira todos os possíveis modelos— α será verdadeira quando KB for verdadeira

<i>A</i>	<i>B</i>	<i>C</i>	$A \vee C$	$B \vee \neg C$	<i>KB</i>	α
<i>False</i>	<i>False</i>	<i>False</i>				
<i>False</i>	<i>False</i>	<i>True</i>				
<i>False</i>	<i>True</i>	<i>False</i>				
<i>False</i>	<i>True</i>	<i>True</i>				
<i>True</i>	<i>False</i>	<i>False</i>				
<i>True</i>	<i>False</i>	<i>True</i>				
<i>True</i>	<i>True</i>	<i>False</i>				
<i>True</i>	<i>True</i>	<i>True</i>				

Formas normais

Outros enfoques para inferência usam operações sintáticas em sentenças, frequentemente expressas em formas padronizadas

Forma Normal Conjuntiva (CNF—universal)

conjunção de disjunções de literais
cláusulas

Por exemplo, $(A \vee \neg B) \wedge (B \vee \neg C \vee \neg D)$

Forma Normal Disjuntiva (DNF—universal)

disjunção de disjunções de literais
cláusulas

Por exemplo, $(A \wedge B) \vee (A \wedge \neg C) \vee (A \wedge \neg D) \vee (\neg B \wedge \neg C) \vee (\neg B \wedge \neg D)$

Forma de Horn (restrita)

conjunção de cláusulas Horn (cláusulas com ≤ 1 literais positivos)

Por exemplo, $(A \vee \neg B) \wedge (B \vee \neg C \vee \neg D)$

Comumente escrito como um conjunto de implicações:

$B \Rightarrow A$ e $(C \wedge D) \Rightarrow B$

Validez e Satisfatibilidade

Uma sentença é válida se for verdadeira em todos os modelos

Por exemplo, $A \vee \neg A$, $A \Rightarrow A$, $(A \wedge (A \Rightarrow B)) \Rightarrow B$

Validez relaciona-se com inferência pelo Teorema da Dedução:

$BC \models \alpha$ se e só se $(BC \Rightarrow \alpha)$ é válido

Uma sentença é satisfatível se for verdadeira em algum modelo

Por exemplo, $A \vee B$, C

Uma sentença é não-satisfatível se não for verdadeira em nenhum modelo

Por exemplo, $A \wedge \neg A$

Satisfatibilidade relaciona-se com inferência por:

$BC \models \alpha$ se e só se $(BC \wedge \neg \alpha)$ é não-satisfatível

i.e., prova-se α por *reductio ad absurdum*

Métodos de prova

Os métodos de prova se dividem em dois tipos básicos:

Verificação de modelo

enumeração de tabela-verdade (correta e completa para lógica proposicional)

busca heurística em espaço de modelos (correto mas incompleto)

Por exemplo, o algoritmo GSAT (Ex. 6.15)

Aplicação de regras de inferência

Geração correta de sentenças novas a partir das velhas

Prova = uma sucessão de aplicações de regras de inferência

Possa usar regras de inferência como operadores em um algoritmo de busca

Regras de inferência para lógica proposicional

Resolução (para CNF): completa para lógica proposicional

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

Modus Ponens (para Forma de Horn): completa para BCs Horn

$$\frac{\alpha_1, \dots, \alpha_n, \quad \alpha_1 \wedge \dots \wedge \alpha_n \Rightarrow \beta}{\beta}$$

Pode ser usado com forward chaining ou backward chaining

Outras regras de inferência ...

Sumário

Agentes lógicos aplicam inferência a uma base de conhecimento para obter informação nova e tomar decisões

Conceitos básicos de lógica:

–sintaxe: estrutura formal das sentenças

–semântica: veracidade das sentenças com respeito a modelos

–implicação: verdade consequente de uma sentença dada uma outra

–inferência: sentenças derivadas de outras sentenças

–correção: derivações produzem apenas sentenças implicadas –completeza: derivações podem produzir todas as sentenças implicadas

O mundo “Wumpus” requer a habilidade para representar informação parcial ou oculta, argumentação através de casos, etc.

Lógica proposicional é suficiente para algumas destas tarefas

O método da tabela-verdade é correto e completo para lógica proposicional

Sintaxe de LPO: Elementos básicos

Constantes	<i>KingJohn, 2, UCB, ...</i>
Predicados	<i>Brother, \}, ...</i>
Funções	<i>Sqrt, LeftLegOf, ...</i>
Variáveis	<i>x, y, a, b, ...</i>
Conectivos	$\wedge \vee \neg \Rightarrow \Leftrightarrow$
Igualdade	$=$
Quantificadores	$\forall \exists$

Termos e Sentenças Atômicas

Um objeto é uma entidade no mundo. Um termo é uma expressão lógica que se refere a um objeto. Uma sentença atômica expressa fatos.

Sentença atômica = $predicate(term_1, \dots, term_n)$
ou $term_1 = term_2$

Termo = $function(term_1, \dots, term_n)$
ou *constant* ou *variable*

Por exemplo, $Brother(KingJohn, RichardTheLionheart)$
 $(Length(LeftLegOf(Richard)))$

Sentenças atômicas podem ser combinadas à conectivos, formando sentenças complexas:

$\neg S, S_1 \wedge S_2, S_1 \vee S_2, S_1 \Rightarrow S_2, S_1 \Leftrightarrow S_2$

Por exemplo $Sibling(KingJohn, Richard) \Rightarrow Sibling(Richard, KingJohn)$
 $>(1, 2) \vee \leq(1, 2)$

Verdade em lógica de primeiro-ordem

Sentenças são verdadeiras com respeito a um modelo e uma interpretação. Um modelo contém objetos e relações entre eles

Uma interpretação especifica referências para

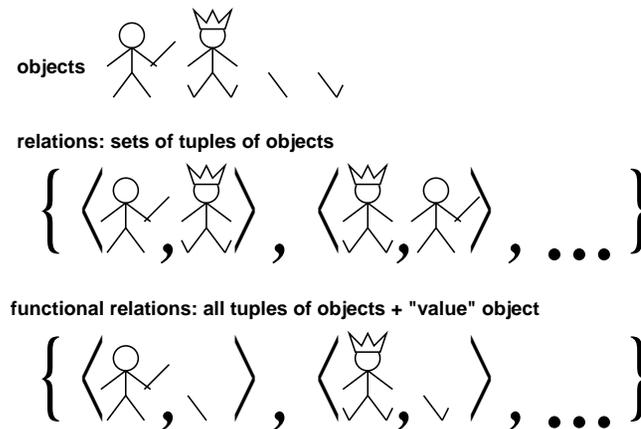
símbolos constantes → objetos

símbolos de predicados → relações

símbolos funcionais → relações funcionais

Uma sentença atômica $predicate(term_1, \dots, term_n)$ é verdadeira se os objetos referidos por $term_1, \dots, term_n$ estão na relação referida por $predicate$

Exemplo de modelo:



Quantificação universal

\forall *<variaveis>* *<sentenca>*

Todo mundo na Unicamp é inteligente:

$\forall x \text{ At}(x, \text{Unicamp}) \Rightarrow \text{Clever}(x)$

$\forall x P$ é equivalente à conjunção de instâncias de P

$\text{At}(\text{KingJohn}, \text{Unicamp}) \Rightarrow \text{Clever}(\text{KingJohn})$

$\wedge \text{At}(\text{Richard}, \text{Unicamp}) \Rightarrow \text{Clever}(\text{Richard})$

$\wedge \text{At}(\text{Unicamp}, \text{Unicamp}) \Rightarrow \text{Clever}(\text{Unicamp})$

$\wedge \dots$

Tipicamente, \Rightarrow é o conectivo principal a se usar com \forall .

Engano comum: usar \wedge como conectivo principal para \forall :

$\forall x \text{ At}(x, \text{Unicamp}) \wedge \text{Clever}(x)$

significa “ Todo mundo está na Unicamp e todo mundo é inteligente”

Quantificação existencial

$\exists \langle \text{variaveis} \rangle \langle \text{sentenca} \rangle$

Alguém em Stanford é inteligente:

$\exists x \text{ At}(x, \text{Stanford}) \wedge \text{Clever}(x)$

$\exists x P$ é equivalente a uma disjunção de instâncias de P

$\text{At}(\text{King John}, \text{Stanford}) \wedge \text{Clever}(\text{King John})$

$\vee \text{At}(\text{Richard}, \text{Stanford}) \wedge \text{Clever}(\text{Richard})$

$\vee \text{At}(\text{Stanford}, \text{Stanford}) \wedge \text{Clever}(\text{Stanford})$

$\vee \dots$

Tipicamente, \wedge é o conectivo principal a se usar com \exists .

Engano comum: usar \Rightarrow como conectivo principal para \exists :

$\exists x \text{ At}(x, \text{Stanford}) \Rightarrow \text{Clever}(x)$

é verdade se existir alguém que não está em Stanford (expanda como uma disjunção e verifique)!

Propriedades de quantificadores

$\forall x \forall y$ é igual a $\forall y \forall x$ (Por que??)

$\exists x \exists y$ é igual a $\exists y \exists x$ (Por que??)

$\exists x \forall y$ não é igual a $\forall y \exists x$

$\exists x \forall y \text{ Loves}(x, y)$

“Há uma pessoa que ama todo o mundo”

$\forall y \exists x \text{ Loves}(x, y)$

“Todo mundo é amado por pelo menos uma pessoa”

Dualidade dos quantificadores: cada um pode ser expresso usando-se o outro

$\forall x \text{ Likes}(x, \text{IceCream}) \quad \neg \exists x \neg \text{Likes}(x, \text{IceCream})$

$\exists x \text{ Likes}(x, \text{Broccoli}) \quad \neg \forall x \neg \text{Likes}(x, \text{Broccoli})$

Manipulando orações

Irmãos são irmãos ou irmãs

$$\forall x, y \text{ Brother}(x, y) \Leftrightarrow \text{Sibling}(x, y).$$

Fraternalidade é reflexiva

$$\forall x, y \text{ Sibling}(x, y) \Leftrightarrow \text{Sibling}(y, x)$$

A mãe da pessoa é o genitor fêmea da pessoa

$$\forall x, y \text{ Mother}(x, y) \Leftrightarrow (\text{Female}(x) \wedge \text{Parent}(x, y))$$

Um primo em primeiro grau é um filho de irmão de um pai

$$\forall x, y \text{ FirstCousin}(x, y) \Leftrightarrow \exists p, ps \text{ Parent}(p, x) \wedge \text{Sibling}(ps, p) \wedge \text{Parent}(ps, y)$$

Igualdade

$term_1 = term_2$ é verdade sob uma determinada interpretação se e só se $term_1$ e $term_2$ se referem ao mesmo objeto

Por exemplo, $1 = 2$ e $\forall x \times (Sqrt(x), Sqrt(x)) = x$ são satisfatíveis
 $2 = 2$ é válido

Por exemplo, definição correta de *Sibling* em termos de *Parent*:

$$\forall x, y \text{ Sibling}(x, y) \Leftrightarrow [\neg(x = y) \wedge \exists m, f \neg(m = f) \wedge \\ \text{Parent}(m, x) \wedge \text{Parent}(f, x) \wedge \text{Parent}(m, y) \wedge \text{Parent}(f, y)]$$

Interagindo com BCs em LPO

Primeiro passo: definir percepções e ações.

Suponha um agente no mundo “Wumpus” usando uma BC baseada em LPO, que percebe um cheiro e uma brisa (mas nenhum brilho) em $t = 5$:

$TELL(KB, Percept([Smell, Breeze, None], 5))$
 $ASK(KB, \exists a \text{ Action}(a, 5))$

I.e., a BC requer alguma ação particular em $t = 5$?

Resposta: *Sim*, $\{a/Shoot\}$ \leftarrow substituição (*binding list*)

Dado uma sentença S e uma substituição σ , $S\sigma$ denota o resultado de “ligar” σ a S ; por exemplo,

$S = Cleverer(x, y)$

$\sigma = \{x/Hillary, y/Bill\}$

$S\sigma = Cleverer(Hillary, Bill)$

$ASK(KB, S)$ devolve algum/todos σ tais que $KB \models S\sigma$

Um Agente Reflexo

“Percepção”

$\forall b, g, t \text{ Percept}([Smell, b, g], t) \Rightarrow Smell(t)$

$\forall s, b, t \text{ Percept}([s, b, Glitter], t) \Rightarrow AtGold(t)$

Reflexo: $\forall t \text{ AtGold}(t) \Rightarrow \text{Action}(Grab, t)$

Reflexo com estado interno: já tenho o ouro?

$\forall t \text{ AtGold}(t) \wedge \neg Holding(Gold, t) \Rightarrow \text{Action}(Grab, t)$

$Holding(Gold, t)$ não pode ser observado

\Rightarrow manter rastro de mudança é essencial

Fatos são válidos em situações, e não eternamente!

Mantendo rastro de mudanças

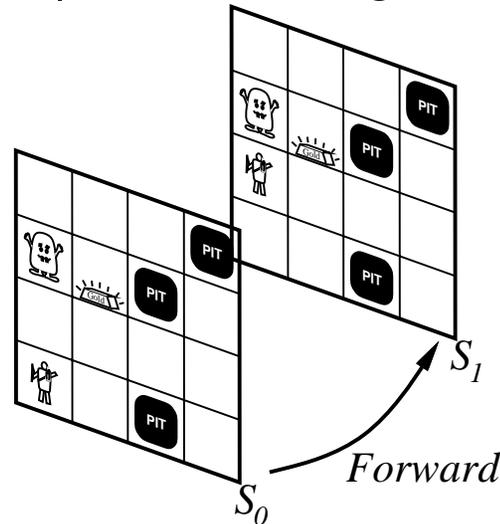
Cálculo de Situações é um modo de representar mudanças em LPO: Adiciona um “argumento de situação” para cada predicado temporário

Por exemplo, *Now* em *Holding(Gold, Now)* denota uma situação.

Situações são ligadas pela função *Result*

$Result(a, s)$ é a situação que resulta ao se fazer a em s

Coisas que não mudam não precisam do argumento de situação.



Descrevendo ações (I)

“Axioma do Efeito” — descreve mudanças devido à ação

$$\forall s \text{ AtGold}(s) \Rightarrow \text{Holding}(\text{Gold}, \text{Result}(\text{Grab}, s))$$

“Axioma do Frame” — descreve permanências devido à ação

$$\forall s \text{ HaveArrow}(s) \Rightarrow \text{HaveArrow}(\text{Result}(\text{Grab}, s))$$

Problema do Frame: ache um modo elegante para lidar com permanências

Problema da Qualificação: descrições verdadeiras de ações reais requerem infinito detalhamento — e se ouro for escorregadio ou estiver colado no chão ou ... É difícil garantir a efetividade das ações.

Problema de Ramificação: ações reais têm muitas consequências secundárias — poeira das brisas cai no ouro, luvas se desgastam, ... É difícil considerar corretamente a importância das consequências.

Descrivendo ações (II)

Axiomas sucessor-estado diminuem o problema do frame representacional

Cada axioma é “sobre” um predicado (não uma ação *per se*):

$$\begin{aligned} P \text{ verdadeiro depois} &\Leftrightarrow [\text{uma ação fez } P \text{ verdadeiro} \\ &\vee P \text{ já é verdadeiro e nenhuma ação fez } P \text{ falso}] \end{aligned}$$

Para segurar o ouro:

$$\begin{aligned} \forall a, s \text{ Holding}(\text{Gold}, \text{Result}(a, s)) &\Leftrightarrow \\ &[(a = \text{Grab} \wedge \text{AtGold}(s)) \\ &\vee (\text{Holding}(\text{Gold}, s) \wedge a \neq \text{Release})] \end{aligned}$$

Um único axioma para cada predicado “modificável”.

Deduzindo propriedades escondidas

Propriedades das localizações :

$$\forall l, t \text{ At}(Agent, l, t) \wedge \text{Smelt}(t) \Rightarrow \text{Smelly}(l)$$

$$\forall l, t \text{ At}(Agent, l, t) \wedge \text{Breeze}(t) \Rightarrow \text{Breezy}(l)$$

Localizações próximas a buracos têm brisa:

Regra de diagnóstico — deduza causas de efeitos

$$\forall y \text{ Breezy}(y) \Rightarrow \exists x \text{ Pit}(x) \wedge \text{Adjacent}(x, y)$$

Regra causal — deduza efeitos de causas

$$\forall x, y \text{ Pit}(x) \wedge \text{Adjacent}(x, y) \Rightarrow \text{Breezy}(y)$$

Nenhum destes é completo — por exemplo., a regra causal não diz se posições distantes de buracos podem ter brisa

Definição para o predicado *Breezy*:

$$\forall y \text{ Breezy}(y) \Leftrightarrow [\exists x \text{ Pit}(x) \wedge \text{Adjacent}(x, y)]$$

Resumo

Lógica de primeiro-ordem:

- objetos e relações são primitivas semânticas
 - sintaxe: constantes, funções, predicados, igualdade, quantificadores

Poder expressivo aumentado: suficiente para definir mundo “Wumpus”

Cálculo de situações:

- convenções para descrever ações e mudança em LPO
 - pode formular planejando como conclusão em uma BC com cálculo de situações