

Inteligência Artificial

Aula 11
 Prof^a Bianca Zadrozny
<http://www.ic.uff.br/~bianca/ia>

Aula 11 - 17/09/10

Incerteza

Capítulo 13 – Russell & Norvig
 Seções 13.1 a 13.4

Aula 11 - 17/09/10

Incerteza

- Seja a ação A_t = sair para o aeroporto t minutos antes do voo.
- A_t me levará ao aeroporto a tempo?
- Dificuldades de saber o resultado da ação:
 - Estados parcialmente observáveis
 - Estados das estradas, trânsito, etc.
 - Sensores ruidosos
 - Relatórios de trânsito
 - Incerteza quanto ao efeito das ações
 - Acidentes, pneu furado, etc.
 - Grande complexidade em prever e modelar o trânsito

Aula 11 - 17/09/10

Incerteza

- Um procedimento puramente lógico não é muito útil nesse caso, porque:
 1. Arriscaria deduzir algo potencialmente falso
 - “ A_{45} me levará a tempo ao aeroporto”
 2. Levaria a conclusões fracas para tomada de decisões
 - “ A_{45} me levará a tempo ao aeroporto, se nenhum acidente ocorrer na ponte, se não chover, se nenhum pneu furar, etc.”
 3. Levaria a conclusões que não práticas
 1. “ A_{1440} me levará a tempo ao aeroporto”

Aula 11 - 17/09/10

Lidando com a incerteza

- Probabilidade
 - Modela o **grau de crença** de um agente dadas as evidências disponíveis
 - “ A_{25} chegará a tempo com probabilidade 0.04”
 - “ A_{45} chegará a tempo com probabilidade 0.85”
 - “ A_{60} chegará a tempo com probabilidade 0.95”

Aula 11 - 17/09/10

Probabilidade

- A probabilidade proporciona um meio para resumir a incerteza que vem de:
 - Preguiça = falha em enumerar todas as possíveis exceções à regra
 - Ignorância = falta de conhecimento sobre fatos relevantes, condições iniciais

Aula 11 - 17/09/10

Probabilidade

- Probabilidade subjetiva ou bayesiana
 - Estabelece o estado de crença do agente em uma sentença, dadas as evidências.
 - Muda quando novas evidências chegam
 - $P(A_{25} | \text{nenhum acidente}) = 0.06$
 - $P(A_{25} | \text{nenhum acidente, 5 a.m.}) = 0.15$
- As sentenças são verdadeiras ou falsas.
 - O que muda é o grau de crença do agente na sentença.
 - Atribuir probabilidade 0 a uma sentença significa acreditar que ela é falsa com certeza absoluta.
 - Atribuir probabilidade 1 a uma sentença significa acreditar que ela é verdadeira com certeza absoluta.

Aula 11 - 17/09/10

Decisões sob incerteza

- Suponha o seguinte conjunto de crenças:

$P(A_{25} \text{ chega a tempo} \dots)$	= 0.04
$P(A_{90} \text{ chega a tempo} \dots)$	= 0.70
$P(A_{120} \text{ chega a tempo} \dots)$	= 0.95
$P(A_{1440} \text{ chega a tempo} \dots)$	= 0.9999
- Que ação o agente deve tomar?
 - Depende de suas **preferências** sob perder o voo versus o tempo esperando no aeroporto.
 - Teoria da utilidade = representação de preferências
 - Teoria da decisão = teoria da probabilidade + teoria da utilidade

Aula 11 - 17/09/10

Introdução à probabilidade

- Elemento básico: variável aleatória
 - Análogo à lógica proposicional
 - Mundos possíveis são definidos pela atribuição de valores às variáveis.
 - Cada variável aleatória tem um **domínio** que determina seus valores possíveis.
 - Tipos de domínio
 - **Booleano**, ex.: *Cárie* possui valores em <verdadeiro,falso>
 - **Discreto**, ex.: *Clima* possui valores em <ensolarado, chuvoso, nublado, neve>
 - **Contínuo**, ex.: *Temperatura*

Aula 11 - 17/09/10

Introdução à probabilidade

- Proposições elementares
 - São construídas através da atribuição de valores a variáveis.
 - Ex.: *Clima* = ensolarado, *Cárie* = falso (abreviado como \neg *cárie*)
- Proposições complexas
 - São formadas a partir de proposições elementares e conectivos lógicos padrão
 - Ex.: *Clima* = ensolarado \vee *Cárie* = falso

Aula 11 - 17/09/10

Introdução à probabilidade

- Evento atômico
 - Especificação completa do estado do mundo sobre o qual o agente está incerto.
 - Uma atribuição de valores a TODAS as variáveis das quais o mundo é formado.
 - Eventos atômicos são mutuamente exclusivos e exaustivos.

Aula 11 - 17/09/10

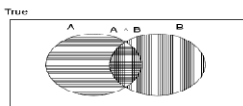
Evento atômico: exemplo

- Se o mundo consistir somente de 2 variáveis booleanas (*Cárie* e *DorDeDente*), então há 4 eventos atômicos distintos:
 - *Cárie* = verdadeiro \wedge *DorDeDente* = verdadeiro
 - *Cárie* = verdadeiro \wedge *DorDeDente* = falso
 - *Cárie* = falso \wedge *DorDeDente* = verdadeiro
 - *Cárie* = falso \wedge *DorDeDente* = falso

Aula 11 - 17/09/10

Axiomas da Probabilidade

- Para quaisquer proposições A, B
 - $0 \leq P(A) \leq 1$
 - $P(\text{verdade}) = 1$ e $P(\text{falso}) = 0$
 - (proposições neces. verdadeiras -- válidas -- prob=1 e proposições neces. falsas -- não satisfáveis -- prob.=0)
 - $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$



Aula 11 - 17/09/10

Probabilidade

- A probabilidade de uma proposição é igual à soma das probabilidades dos eventos atômicos em que ela é válida:

$$P(a) = \sum_{e_i \in e(a)} P(e_i)$$

- Essa equação permite calcular a probabilidade de qualquer proposição dada uma distribuição conjunta total que especifique todos os eventos atômicos.

Aula 11 - 17/09/10

Probabilidade incondicional ou “a priori”

- É o grau de crença em uma proposição na ausência de outras informações.
 - Exemplos:
 - $P(\text{Cárie} = \text{verdadeiro}) = 0.1$
 - $P(\text{Clima} = \text{ensolarado}) = 0.72$
- Distribuição de probabilidades
 - Dá probabilidades a todos os valores possíveis de uma variável aleatória.
 - $P(\text{Clima}) = \langle 0.72, 0.1, 0.08, 0.1 \rangle$ (normalizado, i.e., soma da 1)

Aula 11 - 17/09/10

Distribuição de Probabilidade Conjunta

- Probabilidades de todas as combinações de valores de um conjunto de variáveis aleatórias.

$P(\text{Clima}, \text{Cárie}) =$ tabela 4×2 de valores:

<i>Clima</i> =	ensolarado	chuvoso	nublado	neve
<i>Cárie</i> = verdadeiro	0.144	0.02	0.016	0.02
<i>Cárie</i> = falso	0.576	0.08	0.064	0.08

- Uma distribuição conjunta total especifica a probabilidade de qualquer evento atômico.
 - Qualquer probabilidade nesse domínio pode ser calculada a partir da distribuição conjunta total.

Aula 11 - 17/09/10

Probabilidade condicional ou “a posteriori”

- É o grau de crença em uma proposição dada a presença de evidências (valores de variáveis aleatórias conhecidos).
 - Exemplos:
 - $P(\text{Cárie} = \text{verdadeiro} \mid \text{DorDeDente} = \text{verdadeiro}) = 0.8$
 - $P(\text{Cárie} = \text{verdadeiro} \mid \text{DorDeDente} = \text{verdadeiro}, \text{Cárie} = \text{verdadeiro}) = 1$
 - $P(\text{Cárie} = \text{verdadeiro} \mid \text{DorDeDente} = \text{verdadeiro}, \text{Ensolarado} = \text{verdadeiro}) = P(\text{Cárie} = \text{verdadeiro} \mid \text{DorDeDente}) = 0.8$
- Distribuição condicional
 - $P(Y|X)$ fornece o valor de $P(Y=y_i \mid X=x_i)$ para cada valor de i e j possíveis.

Aula 11 - 17/09/10

Probabilidade Condicional

- Pode ser definida em termos de probabilidades a priori: $P(a \mid b) = P(a \wedge b) / P(b)$ se $P(b) > 0$
- Regra do produto dá uma definição alternativa: $P(a \wedge b) = P(a \mid b) P(b) = P(b \mid a) P(a)$
- Isso pode ser generalizado para distribuições totais: e.g. $P(\text{Clima}, \text{Cárie}) = P(\text{Clima} \mid \text{Cárie}) P(\text{Cárie})$ (que é um conjunto de 4×2 equações, não uma multiplicação matricial.)
- Regra da cadeia é obtida a partir de aplicações sucessivas da regra do produto:

$$P(x_1, \dots, x_n) = P(x_1) P(x_2 \mid x_1) P(x_3 \mid x_1, x_2) \dots P(x_n \mid x_1, \dots, x_{n-1})$$

$$= \prod_{i=1}^n P(x_i \mid x_1, \dots, x_{i-1})$$

Aula 11 - 17/09/10

Inferência Probabilística

- **Inferência probabilística:** a computação a partir de evidências observadas de probabilidades posteriores para proposições de consulta.
- **Inferência com o uso de distribuições conjuntas totais:** base de conhecimento a partir da qual são derivadas respostas para todas as consultas.

Aula 11 - 17/09/10

Exemplo: Inferência Probabilística

- Suponha um domínio com a seguinte distribuição conjunta total:

	<i>dordedente</i>		<i>-dordedente</i>	
	<i>boticão</i>	<i>-boticão</i>	<i>boticão</i>	<i>-boticão</i>
<i>cárie</i>	.108	.012	.072	.008
<i>-cárie</i>	.016	.064	.144	.576

- Para qualquer proposição α , $P(\alpha)$ é a soma dos eventos atômicos w onde α ocorre: $P(\alpha) = \sum_{w:w|\models\alpha} P(w)$

Aula 11 - 17/09/10

Exemplo: Inferência Probabilística

- Suponha um domínio com a seguinte distribuição conjunta total:

	<i>dordedente</i>		<i>-dordedente</i>	
	<i>boticão</i>	<i>-boticão</i>	<i>boticão</i>	<i>-boticão</i>
<i>cárie</i>	.108	.012	.072	.008
<i>-cárie</i>	.016	.064	.144	.576

- Para qualquer proposição α , $P(\alpha)$ é a soma dos eventos atômicos w onde α ocorre: $P(\alpha) = \sum_{w:w|\models\alpha} P(w)$

$$P(\textit{dordedente}) = 0.108 + 0.012 + 0.016 + 0.064 = 0.2$$

Aula 11 - 17/09/10

Exemplo: Inferência Probabilística

- Suponha um domínio com a seguinte distribuição conjunta total:

	<i>dordedente</i>		<i>-dordedente</i>	
	<i>boticão</i>	<i>-boticão</i>	<i>boticão</i>	<i>-boticão</i>
<i>cárie</i>	.108	.012	.072	.008
<i>-cárie</i>	.016	.064	.144	.576

- Para qualquer proposição α , $P(\alpha)$ é a soma dos eventos atômicos w onde α ocorre: $P(\alpha) = \sum_{w:w|\models\alpha} P(w)$

$$P(\textit{dordedente} \vee \textit{cárie}) = 0.108 + 0.012 + 0.016 + 0.064 + 0.072 + 0.008 = 0.28$$

Aula 11 - 17/09/10

Exemplo: Inferência Probabilística

- Podemos calcular probabilidades condicionais:

	<i>dordedente</i>		<i>-dordedente</i>	
	<i>boticão</i>	<i>-boticão</i>	<i>boticão</i>	<i>-boticão</i>
<i>cárie</i>	.108	.012	.072	.008
<i>-cárie</i>	.016	.064	.144	.576

$$\begin{aligned}
 P(\textit{-cárie} | \textit{dordedente}) &= \frac{P(\textit{-cárie} \wedge \textit{dordedente})}{P(\textit{dordedente})} \\
 &= \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064} \\
 &= 0.4
 \end{aligned}$$

Aula 11 - 17/09/10

Normalização

	<i>dordedente</i>		<i>-dordedente</i>	
	<i>boticão</i>	<i>-boticão</i>	<i>boticão</i>	<i>-boticão</i>
<i>cárie</i>	.108	.012	.072	.008
<i>-cárie</i>	.016	.064	.144	.576

- O denominador pode ser visto como uma constante de normalização α

$$\begin{aligned}
 P(\textit{Cárie} | \textit{dordedente}) &= \alpha P(\textit{Cárie}, \textit{dordedente}) \\
 &= \alpha [P(\textit{Cárie}, \textit{dordedente}, \textit{boticão}) + P(\textit{Cárie}, \textit{dordedente}, \textit{-boticão})] \\
 &= \alpha [<0.108, 0.016> + <0.012, 0.064>] \\
 &= \alpha [<0.12, 0.08>] \\
 &= <0.6, 0.4>
 \end{aligned}$$

Aula 11 - 17/09/10