

Inteligência Artificial

Lidando com incerteza

Profa Debora Medeiros

Representação de incerteza

- Muitas vezes não se tem acesso a toda verdade sobre o domínio
- Muitas asserções reais são incertas e/ou incompletas
 - Provavelmente irá chover
 - A rampa é levemente inclinada
 - O paciente deve estar doente

Lidar com incerteza

Exemplo

- Considerar **grau de crença** em sentenças relevantes
 - Teoria da probabilidade é uma das ferramentas
 - Ex.: Há 80% de chance do paciente ter cárie se ele está com dor de dente
 - Resume os casos em que todos os fatores necessários para que **cárie cause dor de dente** estão presentes
 - E também casos em que **paciente tem dor de dente e cárie, mas não há relação entre elas**
 - Outros 20% são os casos em que somos **“preguiçosos” ou “ignorantes”** demais para confirmar ou negar

Lidando com a incerteza

Probabilidade

Modela o **grau de crença** de um agente dadas as evidências disponíveis

“A₂₅ chegará a tempo com probabilidade 0.04”

“A₄₅ chegará a tempo com probabilidade 0.85”

“A₆₀ chegará a tempo com probabilidade 0.95”

Extensão da Lógica

- Graus de crença aplicados a proposições
 - Afirmações que tal situação está ocorrendo
- Elemento básico: **variável aleatória**
 - Análogo à lógica proposicional
 - **Mundos possíveis** são definidos pela **atribuição de valores às variáveis**.
 - Se refere a uma “parte” do mundo cujo “status” é inicialmente desconhecido
 - Ex.: Cárie
 - Pode se referir ao fato de siso inferior esquerdo de alguém ter cárie
 - Notação em letra maiúscula

Variável aleatória

- Cada variável aleatória tem um **domínio**
 - Valores que pode assumir
 - Ex.: $\text{dom}(\text{Cárie}) = \langle \text{verdadeiro}, \text{falso} \rangle$
 - Letras minúsculas para valores
- Proposição elementar
 - Afirma que uma variável aleatória tem um valor específico
 - Ex.: $\text{Cárie} = \text{verdadeiro}$
 - Notação simplificada cárie
 - Ex.: $\text{Cárie} = \text{falso}$
 - Notação simplificada $\sim \text{cárie}$

Variável aleatória

- Tipos:
 - Booleanas:
 - Domínio <verdadeiro, falso>
 - Discretas:
 - Incluem booleanas como caso especial
 - Ex.: $\text{dom}(\text{Tempo}) = \langle \text{ensolarado}, \text{chuvoso}, \text{nublado}, \text{nevoeiro} \rangle$
 - Ex.: nevoento é abreviatura de Tempo = nevoeiro
 - Contínuas:
 - Assumem valores reais
 - Ex.: intervalo $[0, 1]$

Conectivos lógicos

- Combinação de proposições elementares
 - Para formar **proposições complexas**
 - Conectivos lógicos padrão
 - Ex.: Cárie = verdadeiro \wedge DorDeDente = falso
 - Também pode ser escrito como cárie \wedge \sim dordedente



Eventos atômicos

- Especificação *completa* do estado do mundo sobre o qual se está inseguro
 - Atribuição de valores específicos a todas as variáveis do problema
 - Ex.: problema com apenas as variáveis booleanas
Cárie e DorDeDente
 - Há quatro eventos atômicos distintos



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

Eventos atômicos

- Algumas propriedades:
 - *Mutuamente exclusivos*
 - No máximo um deles ocorre em cada instante
 - Ex.: $\text{cárie} \wedge \text{dordedente}$ e $\text{cárie} \wedge \sim \text{dordedente}$ não podem ocorrer simultaneamente
 - *Conjunto de todos eventos atômicos é exaustivo*
 - Pelo menos um deles tem que ocorrer
 - Disjunção de todos eventos atômicos é verdadeira

Axiomas da probabilidade

- Todas probabilidades estão entre 0 e 1
 - Para qualquer proposição a , $0 \leq P(a) \leq 1$
- Proposições válidas têm probabilidade 1
 - Necessariamente verdadeiras
 - $P(\text{verdadeira}) = 1$
- Proposições não-satisfazíveis têm probabilidade 0
 - Necessariamente falsas
 - $P(\text{falsa}) = 0$

Axiomas da probabilidade

- Probabilidade da disjunção

- $P(a \vee b) = P(a) + P(b) - P(a \wedge b)$

- $P(a) + P(b)$ conta a intersecção duas vezes

- Usando os axiomas:

- $P(\sim a) = 1 - P(a)$

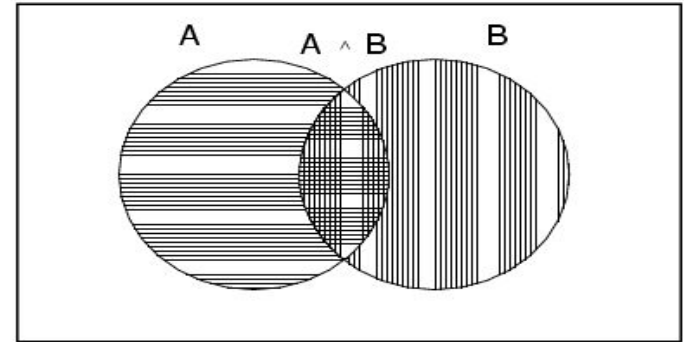
- $\sum_{i=1:n} P(D=d_i) = 1$

- Probabilidade proposição =
probabilidades dos eventos
atômicos em que é válida

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

- Permite calcular a probabilidade de qualquer proposição

True



Probabilidade a priori

- Probabilidade incondicional associada a uma proposição a
 - Grau de crença para a proposição na ausência de quaisquer outras informações
 - Crença a priori, antes da chegada de qualquer nova evidência
 - Representada por $P(a)$
 - Ex.: $P(\text{Cárie}=\text{verdadeiro}) = 0,1$ ou $P(\text{cárie}) = 0,1$

Probabilidade a priori

■ Exemplo:

Paciente	Teste	Doença
001	positivo	presente
002	negativo	presente
003	negativo	ausente
004	positivo	presente
005	positivo	ausente
006	positivo	presente
007	negativo	ausente
008	negativo	presente
009	positivo	ausente
010	positivo	presente

Probabilidade *a priori* pode ser estimada pela frequência

$$P(-) = 4/10 = 0,4$$

$$P(+) = 6/10 = 0,6$$

$$P(\text{presente}) = 6/10 = 0,6$$

$$P(\text{ausente}) = 4/10 = 0,4$$

Probabilidade

- Referindo às probabilidades de **todos os valores possíveis** de uma variável
 - Vetor de valores
 - Ex.: $\mathbf{P}(\text{Tempo}) = \langle 0,7, 0,2, 0,08, 0,02 \rangle$
 - $P(\text{Tempo} = \text{ensolarado}) = 0,7$
 - $P(\text{Tempo} = \text{chuvoso}) = 0,2$
 - $P(\text{Tempo} = \text{nublado}) = 0,08$
 - $P(\text{Tempo} = \text{nevoeiro}) = 0,02$

} Soma = 1

Define **distribuição de probabilidade** a priori
para a variável aleatória Tempo

Probabilidade

- Probabilidade conjunta

- Probabilidades de todas as combinações de valores de um conjunto de variáveis aleatórias
- Ex.: $\mathbf{P}(\text{Tempo}, \text{Cárie})$ pode ser representada por tabela 4 x 2 de probabilidades

Clima =	ensol	chuv	nubla	neve
Cárie = V	0.144	0.02	0.016	0.02
Cárie = F	0.576	0.08	0.064	0.08

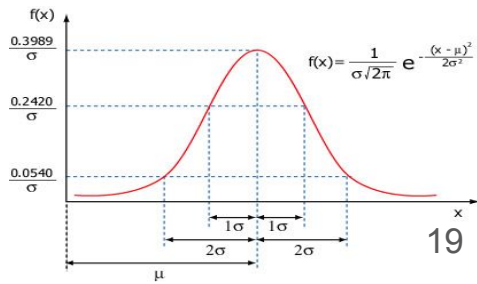
Qualquer questão sobre o domínio pode ser respondida pela distribuição de probabilidade conjunta

Probabilidade

- Distribuição de probabilidade conjunta total
 - Probabilidade conjunta que abrange conjunto completo de variáveis aleatórias do problema
 - Ex.: problema com variáveis Cárie, DorDeDente e Tempo $\Rightarrow \mathbf{P}(\text{Cárie}, \text{DorDeDente}, \text{Tempo})$
 - Essa distribuição conjunta pode ser representada por uma tabela 2 x 2 x 4
 - É especificação completa da incerteza sobre o problema em questão

Probabilidade

- Para variáveis contínuas, não é possível representar distribuição por tabela
 - Existem infinitos valores possíveis
 - Em geral, se define a probabilidade de uma variável aleatória assumir algum valor x como uma função parametrizada de x
 - Ex.: $P(X=x) = U[18, 26](x)$
 - Expressa que x está distribuída uniformemente entre 18 e 26 graus
 - Distribuições de probabilidade para variáveis contínuas = funções de densidade de probabilidade



Probabilidade condicional ou “a posteriori”

- Obtenção de probabilidades condicionais ou posteriores
 - $P(a | b)$, em que a e b são proposições
 - . A probabilidade de a , dado o que sabemos de b
 - . Ex.: $P(\text{cárie} | \text{dordedente}) = 0,8$
 - Se observarmos paciente com dor de dente e não houver outra informação, a probabilidade dele ter cárie é 80%
 - Se sabemos que cárie é verdade
 - . $P(\text{cárie} | \text{dordedente}, \text{cárie}) = 1$
 - Se evidência é irrelevante, pode simplificar
 - . $P(\text{cárie} | \text{dordedente}, \text{ensolarado}) = P(\text{cárie} | \text{dordedente}) = 0,8$
 - Distribuição condicional
 - . $P(Y|X)$ fornece o valor de $P(Y=y_i | X=x_i)$ para cada valor de i e j possíveis.
 - Probabilidades condicionais podem ser definidas em termos de probabilidades incondicionais

Probabilidade condicional

$$P(a|b) = \frac{P(a \wedge b)}{P(b)}$$

- Reescrevendo, chega a **regra do produto**:

$$P(a \wedge b) = P(a|b) P(b)$$

Para a e b serem verdadeiras, b deve ser verdadeira e a deve ser verdadeira, dado b

Também pode ser escrito:

$$P(a \wedge b) = P(b|a) P(a)$$

Probabilidade condicional

- Pode-se usar também a notação $\mathbf{P}(X | Y)$
 - Fornece os valores de $\mathbf{P}(X=x_i | Y=y_j)$ para cada i, j possível
 - Ex.: $\mathbf{P}(\text{Cárie} | \text{DorDeDente})$
 - Pode usar teorema anterior também
 - $\mathbf{P}(\text{Tempo}, \text{Cárie}) = \mathbf{P}(\text{Tempo} | \text{Cárie}) \mathbf{P}(\text{Cárie})$
 - Conjunto de 4 x 2 equações

Inferência probabilística

- Computação de probabilidades posteriores para proposições de consulta a partir de evidências observadas
- Inferência com uso de distribuições conjuntas totais
 - Base de conhecimento a partir da qual são derivadas respostas para todas as consultas

Inferência probabilística

- Base de conhecimento = distribuição conjunta total
 - Exemplo: problema com variáveis `DorDeDente`, `Cárie` e `Boticão`
 - Distribuição conjunta total:

	<i>toothache</i>		\neg <i>toothache</i>	
	<i>catch</i>	\neg <i>catch</i>	<i>catch</i>	\neg <i>catch</i>
<i>cavity</i>	.108	.012	.072	.008
\neg <i>cavity</i>	.016	.064	.144	.576

Inferência probabilística

- Probabilidade de proposição

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

- Soma dos eventos atômicos em que é verdadeira

- $P(\text{cárie} \vee \text{dordedente}) = 0,108 + 0,012 + 0,072 + 0,008 + 0,016 + 0,008 + 0,016 + 0,064 = 0,28$

	<i>toothache</i>		\neg <i>toothache</i>	
	<i>catch</i>	\neg <i>catch</i>	<i>catch</i>	\neg <i>catch</i>
<i>cavity</i>	.108	.012	.072	.008
\neg <i>cavity</i>	.016	.064	.144	.576

Inferência probabilística

- Probabilidade de uma variável = **probabilidade marginal**

$$\begin{aligned} - P(\text{cárie}) &= 0,108 + 0,012 + 0,072 + 0,008 \\ &= 0,2 \end{aligned}$$

	toothache		\neg toothache	
	catch	\neg catch	catch	\neg catch
cavity	.108	.012	.072	.008
\neg cavity	.016	.064	.144	.576

Processo chamado marginalização ou totalização de variáveis

Inferência probabilística

- Calculando probabilidades condicionais:
 - Transformar em probabilidades não condicionais

$$P(a|b) = \frac{P(a \wedge b)}{P(b)}$$

- Depois, calcular como nos casos anteriores

$$\text{Ex.: } P(\text{cárie} | \text{dordedente}) = \frac{P(\text{cárie} \wedge \text{dordedente})}{P(\text{dordedente})}$$

$$= \frac{0,108 + 0,012}{0,108 + 0,012 + 0,016 + 0,064} = 0,6$$

	toothache		\neg toothache	
	catch	\neg catch	catch	\neg catch
cavity	.108	.012	.072	.008
\neg cavity	.016	.064	.144	.576

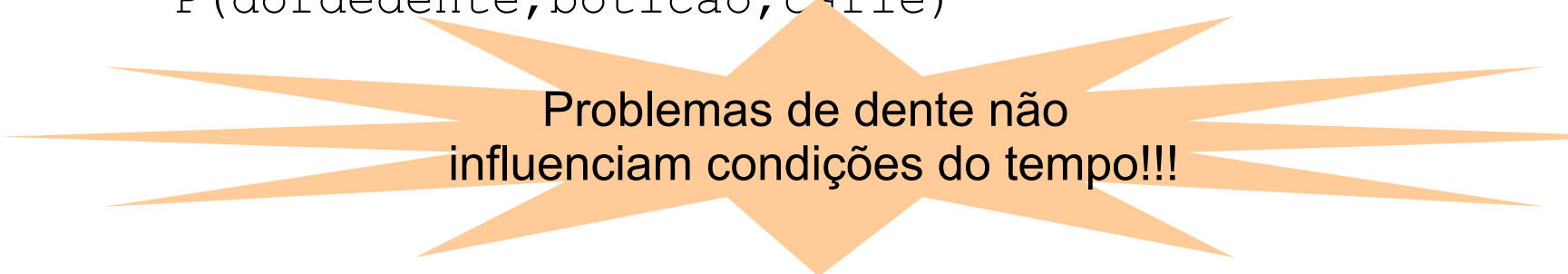
Inferência probabilística

- Distribuição conjunta total tabular **não é ferramenta prática** para construir sistemas de raciocínio
 - Tamanho escala **exponencialmente** com o número de variáveis
 - Ex.: Adicionar variável T_{tempo} ao problema
 - Distribuição conjunta total =
 $P(DorDeDente, Boticão, Cárie, T_{tempo})$
 - Tabela tem 32 entradas
 - Pois T_{tempo} tem quatro valores
 - Dariam quatro das tabelas anteriores, uma para cada tipo de T_{tempo}

Independência

- Ex.:

- $P(\text{dordedente}, \text{boticão}, \text{cárie}, \text{nublado}) =$
 $P(\text{nublado} | \text{dordedente}, \text{boticão}, \text{cárie})$
 $P(\text{dordedente}, \text{boticão}, \text{cárie})$

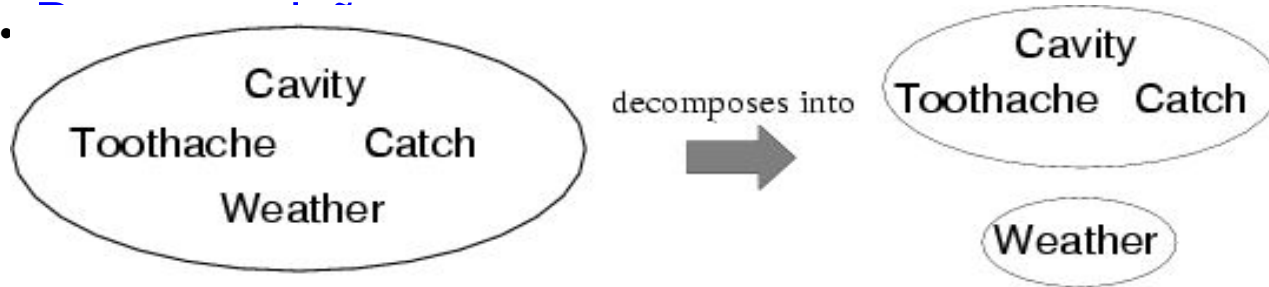


Problemas de dente não
influenciam condições do tempo!!!

$$\Rightarrow P(\text{nublado} | \text{dordedente}, \text{boticão}, \text{cárie}) = P(\text{nublado})$$

Independência

- Há equação semelhante para toda entrada:
 - $P(\text{DorDeDente}, \text{Boticão}, \text{Cárie}, \text{Tempo}) = P(\text{Tempo}) P(\text{DorDeDente}, \text{Boticão}, \text{Cárie})$
 - ⇒ Tabela de 32 elementos pode ser construída a partir de uma tabela de 8 elementos e uma tabela de 4 elementos



Empregou-se propriedade de **independência**, também chamada **independência marginal** e **independência absoluta**

Independência

- Independência entre duas proposições a e b :
 - $P(a|b) = P(a)$ ou
 - $P(b|a) = P(b)$ ou
 - $P(a \wedge b) = P(a) P(b)$
- Independência entre duas variáveis X e Y :
 - $P(X|Y) = P(X)$ ou
 - $P(Y|X) = P(Y)$ ou
 - $P(X,Y) = P(X) P(Y)$

Regra de Bayes

- Regra do produto:

- $P(a \wedge b) = P(a|b) P(b)$

- $P(a \wedge b) = P(b|a) P(a)$

- Igualando as equações e dividindo por $P(a)$:

-Regra (teorema) de Bayes:

$$P(b|a) = \frac{P(a|b) P(b)}{P(a)}$$

Base de todos sistemas modernos de IA para inferência probabilística.

Útil para diagnóstico:

$$P(\text{Causa}|\text{Efeito}) = P(\text{Efeito}|\text{Causa}) P(\text{Causa}) / P(\text{Efeito})$$

Exemplo

- Diagnóstico médico

- Frequentemente se tem probabilidades condicionais sobre causas e queremos derivar um diagnóstico
 - Ex.: médico sabe que:
 - Meningite faz paciente ter rigidez no pescoço durante 50% do tempo
 - Probabilidade a priori de paciente ter meningite: $1/50000$
 - Probabilidade a priori de paciente ter rigidez no pescoço: $1/20$
- $\Rightarrow P(\text{meningite} | \text{rigidez}) = 0,5 \times \frac{1/50000}{1/20} = 0,0002$

**Direção de
diagnóstico**

**Direção
causal**

Regra de Bayes

- Combinação de evidências:
 - Aumenta número de cálculos em distribuição conjunta
 - Uso de independência pode diminuir
 - Em muitos casos, de exponencial a linear

Regra da Cadeia

Temos que $P(A,B) = P(A|B)P(B)$

Estendendo para várias variáveis

$$P(A,B,C) = P(A|B,C)P(B,C)$$

$$P(A,B,C) = P(A|B,C)P(B|C)P(C)$$

Forma geral

$$P(A_1, A_2, \dots, A_n) = P(A_1|A_2, \dots, A_n)P(A_2|A_3, \dots, A_n) \dots P(A_{n-1}|A_n)P(A_n)$$

Calcular a distribuição de probabilidade conjunta

Simple se as variáveis forem independentes

Independência

Recaptulando

$$P(A|B) = P(A)$$

$$P(B|A) = P(B)$$

$$P(A \wedge B) = P(A)P(B)$$

Impacto na regra da cadeia

$$P(A_1, A_2, \dots, A_n) = P(A_1)P(A_2) \dots P(A_n)$$

$$P(A_1, A_2, \dots, A_n) = \prod_{i=1}^n P(A_i)$$

Independência condicional

Independência é difícil de satisfazer

Duas variáveis são independentes condicionalmente **dada uma terceira variável**, se o conhecimento de uma não afeta a probabilidade da outra se a terceira é conhecida

$$P(A|B,C) = P(A|C)$$

$$P(B|A,C) = P(B|C)$$

Impacto na regra da cadeia

$$P(A_1, A_2, \dots, A_n) = P(A_1|A_n)P(A_2|A_n) \dots P(A_{n-1}|A_n)P(A_n)$$

$$P(A_1, A_2, \dots, A_n) = P(A_n) \prod_{i=1}^{n-1} P(A_i|A_n)$$

Bibliografia

- . G. Bittencourt
 - Capítulo 2
- . S. O. Rezende
 - Capítulos 1 e 2
- . Russel e Norvig
 - Capítulo 13
- . Slides
 - . Ana Carolina Lorena, Unifesp
 - . Ronaldo Prati, UFABC
 - . Richard Khoury, University of Waterloo
 - Profa Anne Magaly Canuto, UFRN
 - Profa Josiane
 - Profa Solange O. Rezende, ICMC-USP
 - Prof Guilherme Bittencourt, UFSC
 - Prof Ricardo Campello, ICMC-USP