

Inteligência Artificial

Aula 13
Profª Bianca Zadrozny
<http://www.ic.uff.br/~bianca/ia>

Aula 13 - 08/05/09

Raciocínio Probabilístico

Capítulo 14 – Russell & Norvig
Seções 14.1 a 14.2

Aula 13 - 08/05/09

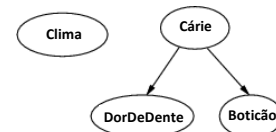
Redes Bayesianas

- Estrutura de dados para representar as dependências entre variáveis e fornecer uma especificação **concisa** de **qualquer** distribuição de probabilidade conjunta total.
- Sintaxe:
 - um conjunto de nós, um para cada variável aleatória
 - grafo direcionado e acíclico (seta = “influência direta”)
 - cada nó tem uma distribuição condicional $P(X_i | \text{Pais}(X_i))$ que quantifica o efeito dos pais sobre o nó
 - **tabela de probabilidade condicional (TPC)**

Aula 13 - 08/05/09

Exemplo

- A topologia de uma rede representa relações de independência condicional :



- *Clima* é independente de outras variáveis
- *DorDeDente* e *Boticão* são condicionalmente independentes dado *Cárie*
- Informalmente, a rede representa o fato de que *Cárie* é uma causa direta de *DorDeDente* e *Boticão*.

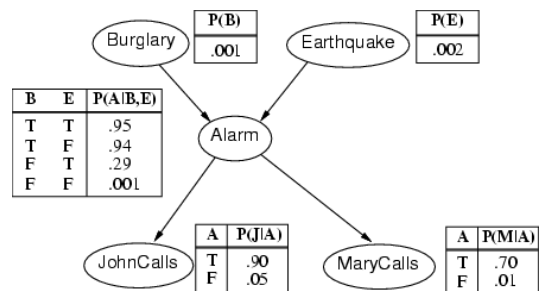
Aula 13 - 08/05/09

Exemplo

- “ *Estou no trabalho, meu vizinho João me liga pra dizer que o alarme da minha casa está tocando, mas a minha vizinha Maria não me liga. Às vezes o alarme toca por causa de terremotos leves. Houve um roubo na minha casa?* ”
- Variáveis: *Roubo (Burglary)*, *Terremoto (Earthquake)*, *Alarme (Alarm)*, *JoãoLiga (JohnCalls)*, *MariaLiga (MaryCalls)*
- A topologia da rede reflete conhecimento “causal”:
 - Um roubo pode ativar o alarme
 - Um terremoto pode ativar o alarme
 - O alarme faz Maria telefonar
 - O alarme faz João telefonar

Aula 13 - 08/05/09

Exemplo



Aula 13 - 08/05/09

Da topologia da rede

- Roubos e terremotos afetam diretamente a probabilidade do alarme tocar.
- Mas o fato de João e Maria telefonarem só depende do alarme.
- Desse modo, **a rede representa nossas suposições de que eles não percebem quaisquer roubos diretamente, não notam os terremotos e não verificam antes de ligar!**

Aula 13 - 08/05/09

As probabilidades...

- ... resumem um conjunto potencialmente infinito de circunstâncias (Maria ouve música alta, João liga quando ouve o telefone tocar; umidade, falta de energia, etc., podem interferir no alarme; João e Maria não estão em casa, etc.
 - “Preguiça e ignorância”

Aula 13 - 08/05/09

Tabelas de probabilidade condicional (TPC)

- Cada linha em uma TPC contém a probabilidade condicional de cada valor de nó para um **caso de condicionamento**;
 - um caso de condicionamento é uma combinação possível de valores para os nós superiores
- Cada linha requer um número p para $X_i = true$ (a probabilidade para $X_i = false$ é $1-p$)

Aula 13 - 08/05/09

Tabelas de probabilidade condicional (TPC)

- Um nó sem pais tem apenas uma linha: probabilidade a priori
- Em geral, uma tabela para uma variável booleana com k pais booleanos possui 2^k probabilidades
- Se cada variável não possui mais do que k pais, a rede completa será $O(n \cdot 2^k)$, para n = número de nós.
 - i.e., cresce linearmente com n , vs. $O(2^n)$ da distribuição total

Aula 13 - 08/05/09

Semântica das RB

- Duas maneiras equivalentes:
 - **Semântica global** (ou numérica): entender as redes como uma representação da distribuição de probabilidade conjunta.
 - Indica como construir uma rede.
 - **Semântica local** (ou topológica): visualizá-las como uma codificação de uma coleção de declarações de independência condicional.
 - Indica como fazer inferências com uma rede.

Aula 13 - 08/05/09

Semântica Global

Se uma rede bayesiana for uma representação da distribuição conjunta, então ela poderá ser utilizada para responder qualquer consulta efetuando-se o **produto** de todas as entradas conjuntas relevantes.



Aula 13 - 08/05/09

Semântica Global

A semântica global (ou numérica) define a distribuição de probabilidade total como o produto das distribuições condicionais locais:

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | \text{Pais}(X_i))$$



Aula 13 - 08/05/09

Semântica Global

A semântica global (ou numérica) define a distribuição de probabilidade total como o produto das distribuições condicionais locais:

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | \text{Pais}(X_i))$$



e.g., $P(j \wedge m \wedge a \wedge \neg b \wedge \neg e)$
 $= P(j | a) P(m | a) P(a | \neg b, \neg e) P(\neg b) P(\neg e)$
 $= 0.9 \times 0.7 \times 0.001 \times 0.999 \times 0.998$
 $= 0.00063$

Aula 13 - 08/05/09

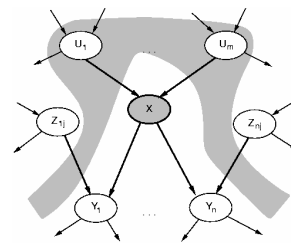
Semântica local

- **Semântica local** (topológica): cada nó é condicionalmente independente de seus não-descendentes, dados seus pais. Ex. **J** é independente de **B** e **E**.



Aula 13 - 08/05/09

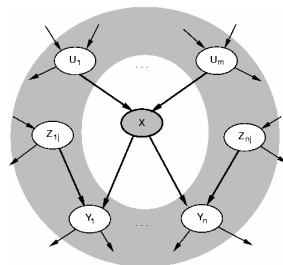
Semântica local



Um nó X é condicionalmente independente de seus não descendentes (Z_j) dados seus pais (U_j).

Aula 13 - 08/05/09

Semântica local



Um nó X é condicionalmente independente de todos os outros nós dada a sua **cobertura de Markov** (=pais, filhos e pais dos filhos)

Aula 13 - 08/05/09

Semântica local e global

- A partir dessas asserções sobre a independência condicional e das TPCs, a distribuição conjunta pode ser reconstruída.
 - Deste modo a semântica numérica e topológica são equivalentes.

Aula 13 - 08/05/09

Construindo uma rede Bayesiana

1. Escolha uma ordem para as variáveis aleatórias X_1, \dots, X_n
2. Para $i = 1$ a n
 - adicione X_i à rede
 - selecione pais para X_1, \dots, X_{i-1} tais que

$$\mathbf{P}(X_i | \text{Pais}(X_i)) = \mathbf{P}(X_i | X_1, \dots, X_{i-1})$$

Esta escolha de pais garante a semântica global:

$$\begin{aligned} \mathbf{P}(X_1, \dots, X_n) &= \prod_{i=1}^n \mathbf{P}(X_i | X_1, \dots, X_{i-1}) \text{ (regra da cadeia)} \\ &= \prod_{i=1}^n \mathbf{P}(X_i | \text{Pais}(X_i)) \text{ (por construção)} \end{aligned}$$

Aula 13 - 08/05/09

Ordem para as variáveis

- A ordem correta em que os nós devem ser adicionados consiste em adicionar primeiro as “causas de raiz”, depois as variáveis que elas influenciam e assim por diante, até chegarmos às folhas, que não tem nenhuma influência causal direta sobre as outras variáveis.
- E se escolhermos a ordem “errada”?

Aula 13 - 08/05/09

Exemplo

- Assumindo a ordem: M, J, A, B, E

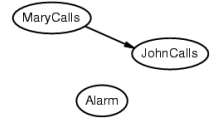


$$\mathbf{P}(J | M) = \mathbf{P}(J)?$$

Aula 13 - 08/05/09

Exemplo

- Assumindo a ordem: M, J, A, B, E



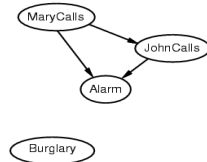
$$\mathbf{P}(J | M) = \mathbf{P}(J)? \text{ Não}$$

$$\mathbf{P}(A | J, M) = \mathbf{P}(A | J)? \quad \mathbf{P}(A | J, M) = \mathbf{P}(A)?$$

Aula 13 - 08/05/09

Exemplo

- Assumindo a ordem: M, J, A, B, E



$$\mathbf{P}(J | M) = \mathbf{P}(J)? \text{ Não}$$

$$\mathbf{P}(A | J, M) = \mathbf{P}(A | J)? \quad \mathbf{P}(A | J, M) = \mathbf{P}(A)? \text{ Não}$$

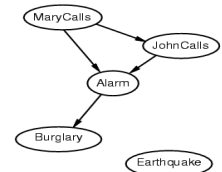
$$\mathbf{P}(B | A, J, M) = \mathbf{P}(B | A)?$$

$$\mathbf{P}(B | A, J, M) = \mathbf{P}(B)?$$

Aula 13 - 08/05/09

Exemplo

- Assumindo a ordem: M, J, A, B, E



$$\mathbf{P}(J | M) = \mathbf{P}(J)? \text{ Não}$$

$$\mathbf{P}(A | J, M) = \mathbf{P}(A | J)? \quad \mathbf{P}(A | J, M) = \mathbf{P}(A)? \text{ Não}$$

$$\mathbf{P}(B | A, J, M) = \mathbf{P}(B | A)? \text{ Sim}$$

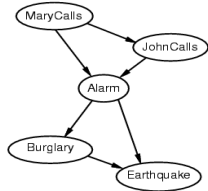
$$\mathbf{P}(B | A, J, M) = \mathbf{P}(B)? \text{ Não}$$

$$\mathbf{P}(E | B, A, J, M) = \mathbf{P}(E | A)?$$

$$\mathbf{P}(E | B, A, J, M) = \mathbf{P}(E | A, B)?$$

Aula 13 - 08/05/09

Exemplo



$P(J | M) = P(J)$? Não

$P(A | J, M) = P(A | J)$? $P(A | J, M) = P(A)$? Não

$P(B | A, J, M) = P(B | A)$? Sim

$P(B | A, J, M) = P(B)$? Não

$P(E | B, A, J, M) = P(E | A)$? Não

$P(E | B, A, J, M) = P(E | A, B)$? Sim

Aula 13 - 08/05/09

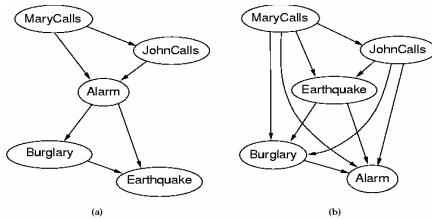
Exemplo



- A rede resultante terá dois vínculos a mais que a rede original e exigirá outras probabilidades para serem especificadas
- Alguns dos vínculos apresentam relacionamentos ténues que exigem julgamentos de probabilidade difíceis e antinaturais (prob de *Terremoto*, dados *Roubo* e *Alarme*)
- (Em geral) é melhor pensar de *causas* para *efeitos* (modelo causal) e não do contrário (modelo de diagnóstico)

Aula 13 - 08/05/09

Exemplo



- Uma ordenação de nós ruim: *MaryCalls*, *JohnCalls*, *Earthquake*, *Burglary* e *Alarm*

- Entretanto, todas as três redes devem representar a mesma distribuição conjunta. As duas últimas só não expressam todas as independências condicionais

Aula 13 - 08/05/09