Introdução ao tema das Redes Bayesianas

Seminário de Modelação

Cátia Azevedo

25/01/2013

FCUP
Departamento
Matemática
ENGENHARIA
MATEMÁTICA

Índice

- Introdução
- Redes Bayesianas
  - Aprendizagem Bayesiana
  - Teorema de Bayes
  - Distribuição de probabilidade conjunta
  - Condição de Markov
  - Exemplo Prático
- Projecto
  - Resumo

Introdução

- Métodos Bayesianos
  - Alguns domínios de problema são caracterizados pela presença de incerteza e pelo raciocínio baseado em crenças parciais a respeito do mundo.

  - Raciocínios adequados sobre esses domínios podem ser feitos através de métodos Bayesianos.

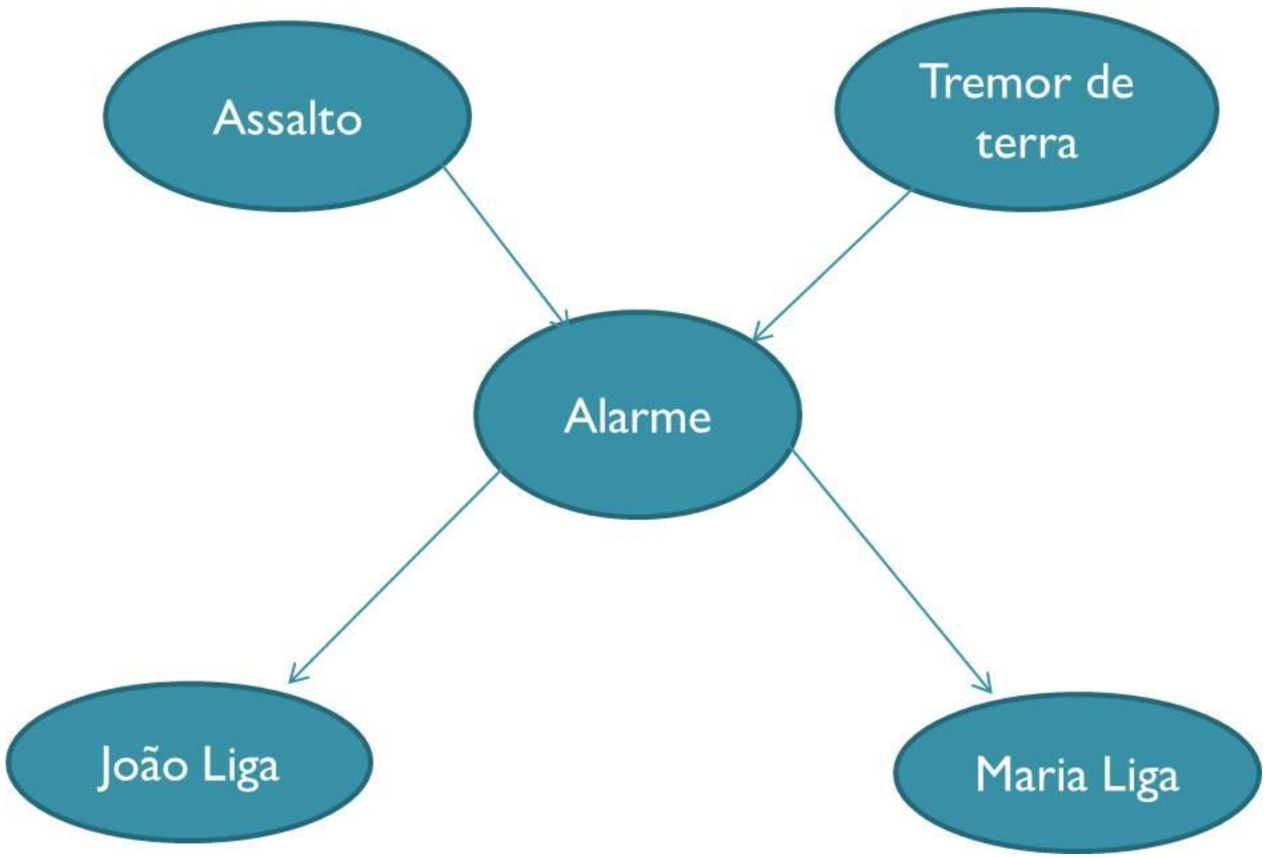
- Redes Bayesianas
  - Metodologia para a construção de sistemas que confiam no conhecimento probabilístico, e que trabalham com o conhecimento incerto e incompleto por meio da Teoria da Probabilidade Bayesiana;

Redes Bayesianas (RB)

Grafo acíclico dirigido (DAG) com as seguintes características:

- Os nós representam as variáveis aleatórias;
- As arestas ligam pares de variáveis, com dependência entre si;
- Nós não conectados representam variáveis independentes;
- Cada nó tem associado os estados da variável que representa e uma tabela de probabilidades condicionadas (calculadas pela fórmula de Bayes) que quantifica os efeitos que os 'pais' exercem sobre um nó (probabilidade do nó estar num estado especifico dado os estados dos seus 'pais');

Ilustração de uma Rede



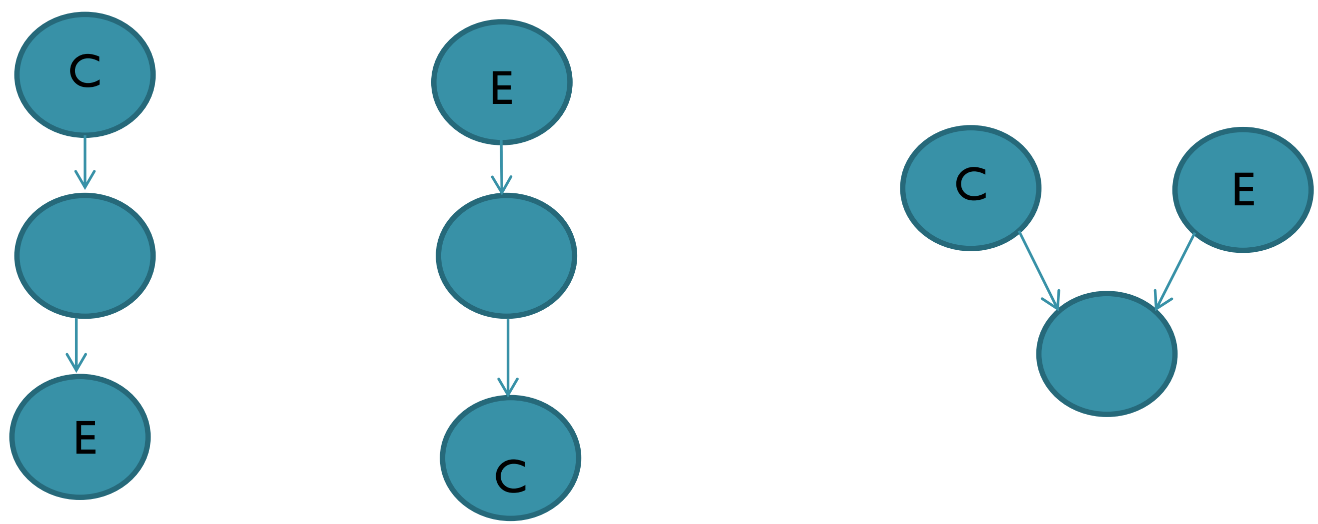
Aprendizagem Bayesiana

- É a forma de obter a representação interna da rede;
- É dividida em duas partes:

Aprendizagem da estrutura	Aprendizagem dos parâmetros
⬇	⬇
Variáveis e as suas relações	Distribuição de probabilidades
⬇	⬇
Consiste em entender a estrutura gráfica da rede, através de uma base de dados	Consiste na obtenção da tabela de probabilidades condicionadas de cada variável

Tipos de conhecimento

- Causal: das causas para os efeitos; (1)
- Diagnóstico: dos efeitos para as causas; (2)
- Intercausal: entre causas de um efeito comum; (3)



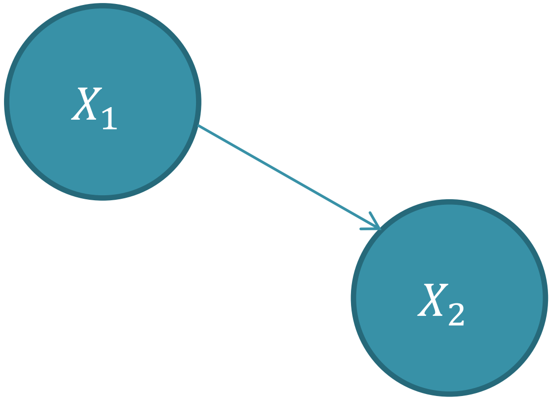
(1) (2) (3)

Exemplo:

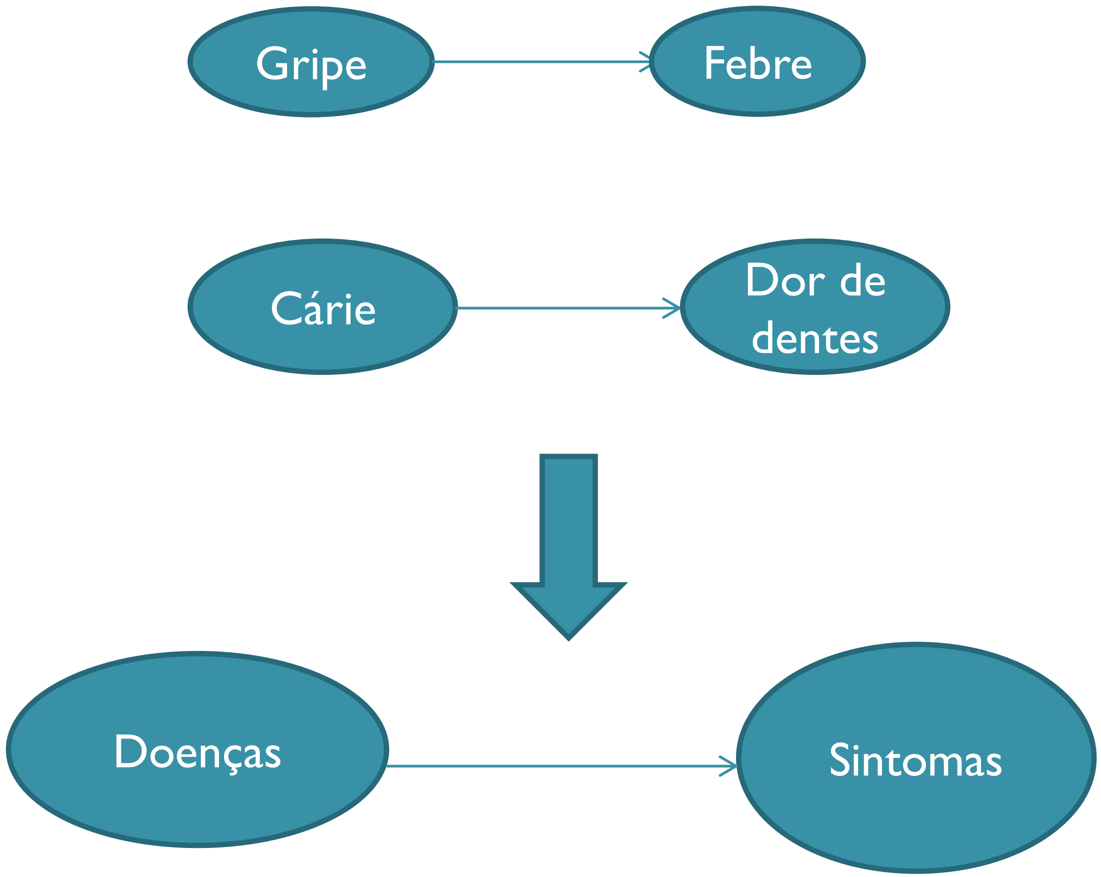
O mais usual é a Rede Bayesiana representar relações de causalidade entre variáveis de um sistema.

Seja X_1 e X_2 duas variáveis de um sistema em estudo, a figura pode representar uma Rede Bayesiana em que

X_1 é a 'causa' de X_2.



Em medicina



Teorema de Bayes

- Mostra a relação entre uma probabilidade condicionada e a sua inversa;
- Dados dois eventos *A* e *B*,

$$P(A|\mathrm{B}) = \frac{P(B|\mathrm{A})P(A)}{P(B)}$$

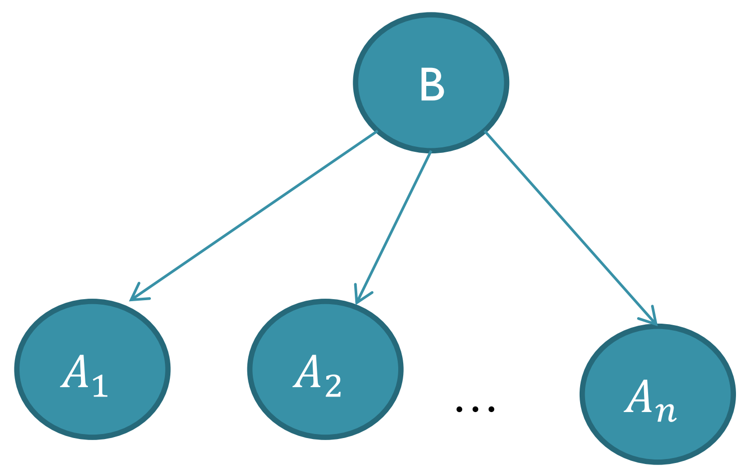
- $P(A)$ e $P(B)$ são as probabilidade *à priori* de A e de B, respectivamente;
- $P(A|\mathrm{B})$ e $P(B|\mathrm{A})$ é a probabilidade condicionada (ou *à posteriori*) de A e B, respectivamente;

Consideremos uma partição de um espaço amostral S, um conjunto de eventos $A_1, A_2, A_3, \ldots, A_n$ independentes entre si e que a sua união é S, e um outro evento B.

Então para qualquer i,

$$P(A_i|B) = \frac{P(B|A_i).P(A_i)}{P(B)} = \frac{P(B|A_i).P(A_i)}{\sum_i P(B|A_i).P(A_i)}$$

A interpretação do teorema de Bayes consiste em considerar os eventos A_i como 'efeitos' do evento B, sendo atribuído probabilidades deste evento atuar na ocorrência de B.



Antes da formalização do problema as probabilidades dos eventos A_i são calculadas e designadas por probabilidades *a priori.*

Sabendo que um evento B ocorreu estas probabilidades são revistas pela fórmula de Bayes e então resulta nas probabilidades *a posteriori* $P(A_i|B)$, para $i = 1,2, \dots, n$.

Distribuição de Probabilidade Conjunta

- Distribuição de probabilidade para um conjunto de duas ou mais variáveis aleatórias.
- Podemos ver uma rede Bayesiana como uma representação da distribuição de probabilidade conjunta.
- A distribuição de probabilidade conjunta de um conjunto de variáveis $X_1, X_2, \ldots, X_n$ é definida como $P(X_1 \wedge X_2 \wedge \cdots \wedge X_n)$, ou $P(X_1, X_2, \ldots, X_n)$,
- Toda a entrada na distribuição de probabilidade conjunta total pode ser calculada a partir de informações armazenadas na rede.

Condição de Markov

- Propriedade Markoviana:
  - A definição desta propriedade é que os estados anteriores são irrelevantes para a predição dos estados seguintes, desde que o estado actual seja conhecido.

- Cadeia de Markov:
  - É uma sequência $X_1, X_2, \ldots, X_n$ de variáveis aleatórias. Se a distribuição de probabilidade condicional de X_{n+1} é uma função apenas de X_n então:

$$P(X_{n+1}=\mathcal{X} \mid X_1, X_2, \ldots, X_n) = P(X_{n+1}=\mathcal{X} \mid X_n)$$

  onde x é algum estado do processo.

- Condição de Markov:
  - Suponha a distribuição de probabilidade conjunta das variáveis aleatórias em um conjunto de nós V num DAG, $G = (V, E)$. Então diz-se que (G, P) satisfazem as condições de Markov se cada variável $X \in V$, X é condicionalmente independente dos nós não descendentes dados seus 'pais'.

Concluindo, a condição de Markov afirma que as variáveis não descendentes não fornecem informações adicionais sobre a variável em questão.

Se uma rede Bayesiana satisfaz a condição de Markov, então a sua distribuição de probabilidade conjunta é igual ao produto das probabilidades condicionais de todos os nós, dados os valores dos seus 'pais'.

Sendo assim, podemos definir a distribuição de probabilidade conjunta como

$$P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P(X_i | pa(X_i))$$

Sendo, X_i um nó da rede e $pa(X_i)$ os seus 'pais'.

Exemplo Prático

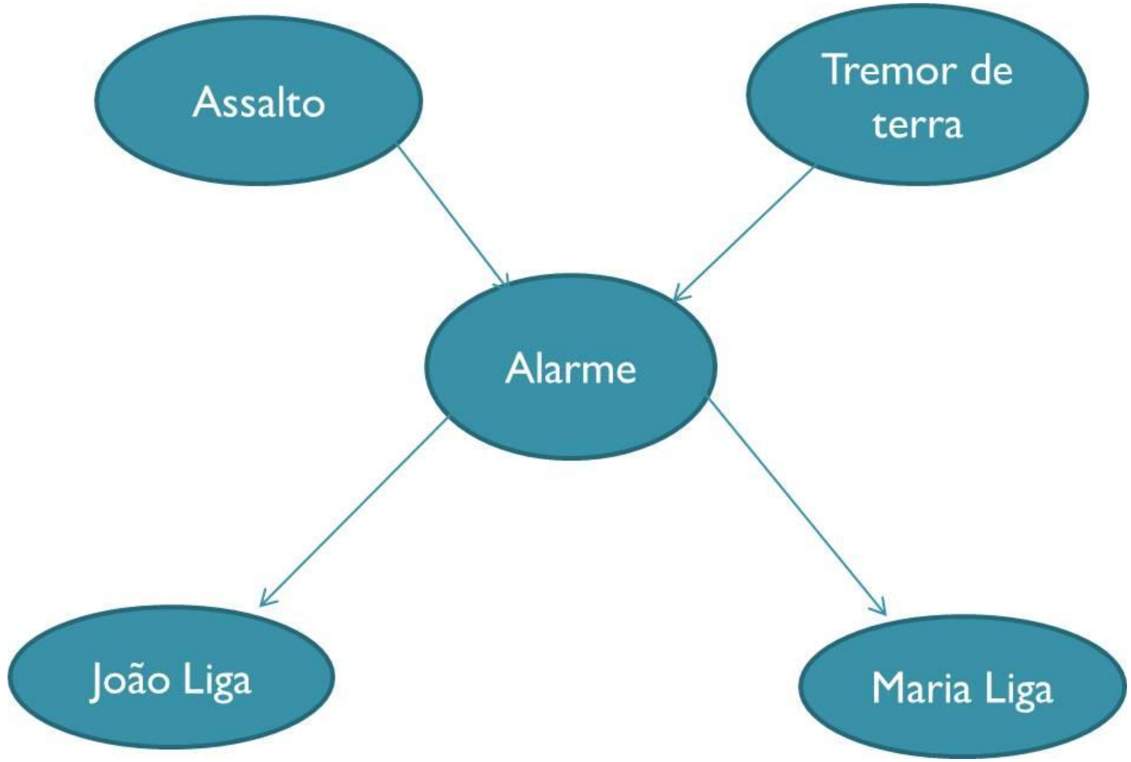
Situação:

- Sistema de alarme instalado em casa
  - É activado quando há um assalto;
  - Também responde quando existem pequenos tremores de terra;
- Dois vizinhos que prometem telefonar para o emprego quando ouvirem o alarme
  - Maria: costuma ouvir musica muito alta e por vezes não ouve o alarme;
  - João: telefona sempre que ouve o alarme mas por vezes confunde o alarme com o toque do telefone.

Temos as variáveis:

- Assalto (AS)
- Tremor de terra (T)
- Alarme (A)
- Maria Liga (ML)
- João liga (JL)

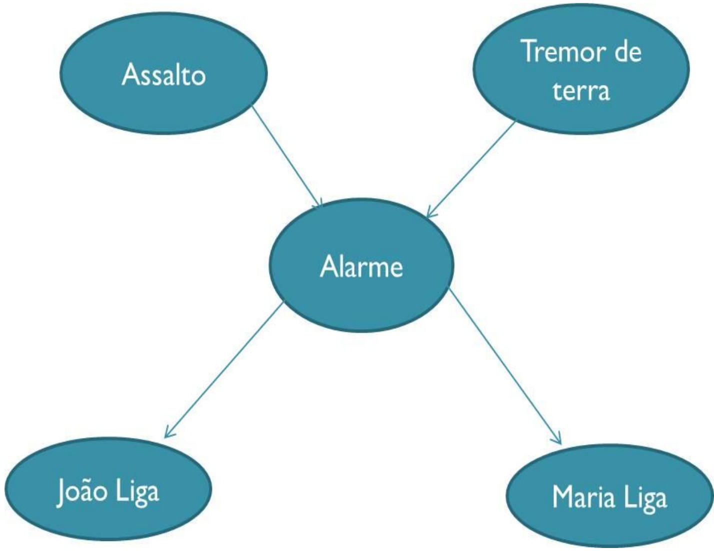
E a Rede Bayesiana é apresentada de seguinte forma:



Uma vez definida a topologia da rede, é necessário determinar as probabilidades dos nós que têm dependências diretas e utilizar estas para determinar as probabilidades que desejamos.

É necessário então calcular as tabelas de probabilidade condicionada de cada nó (variável) dado os seus ‘pais’.

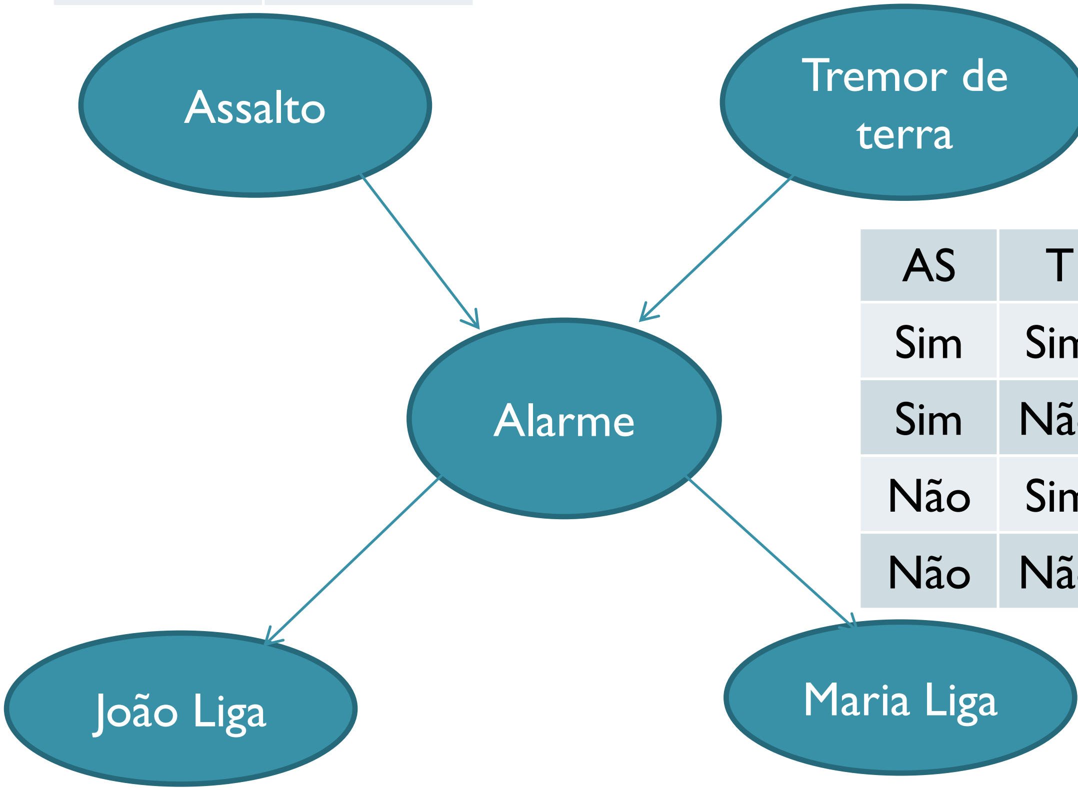
Construção das tabelas de Probabilidade condicionada



- Assalto, e Tremor de Terra
  - Estas variáveis não possuem ‘pais’, logo, a probabilidade destes correspondem a probabilidade à priori, calculada anteriormente.
- Alarme
  - Esta variável tem como ‘pais’ as variáveis assalto e tremor de terra.
- João liga e Maria liga
  - São variáveis que têm como ‘pai’ a variável alarme, dependem apenas desta.

	P(AS)
Sim	0.001
Não	0.999

	P(T)
Sim	0.002
Não	0.998



		P(A)	
AS	T	Sim	Não
Sim	Sim	0.95	0.05
Sim	Não	0.95	0.05
Não	Sim	0.29	0.61
Não	Não	0.001	0.999

A \ P(JL)	Sim	Não
Sim	0.90	0.10
Não	0.05	0.95

A \ P(ML)	Sim	Não
Sim	0.70	0.30
Não	0.01	0.99

Com as tabelas de probabilidade condicionada definidas, é possível então obter a distribuição de probabilidade conjunta.

<u>Exemplo</u>: Suponhamos que queremos saber qual a probabilidade de ter tocado o alarme, sabendo não houve tremor de terra, nem assalto e que a Maria e o João ligaram.

$$P\begin{pmatrix} \textit{não haver tremor, nao haver assalto,} \\ \textit{o alarme tocou, a maria ligou e o joão ligou} \end{pmatrix} =$$

$$= P(\neg T \wedge \neg AS \wedge A \wedge ML \wedge JL) = P(\neg T)P(\neg AS)P(A|\neg T \wedge \neg AS)P(ML)P(JL) =$$

$$= 0.98 \times 0.999 \times 0.001 \times 0.7 \times 0.9 = 0.00062$$

Projecto 2012/2013

O método apresentado vai ser aplicado de forma inovadora no âmbito do projecto:

Diagnóstico epidemiológico e previsão de cárie dentária com recurso a Redes Bayesianas.

Orientadora: Sónia Gouveia (IEETA, Universidade de Aveiro)

Co-orientadores: João Nuno Tavares (DM e CMUP, Universidade do Porto) e Álvaro Azevedo (FMDUP, Universidade do Porto)

Os resultados serão contextualizados no diagnóstico da cárie dentária com a ajuda do especialista da FMUP.

Resumo do Projecto

- A cárie dentária é uma doença infecciosa de origem bacteriana;
- O diagnóstico desta doença ainda representa um dos principais desafios em medicina dentária;
- Nas últimas décadas o estudo das causas desta doença tem sofrido alterações, principalmente em comunidades mais desenvolvidas;

- No estudo do trabalho será determinado um indicador de cárie a partir de rastreios com diferentes métodos e critérios epidemiológicos, em duas comunidades escolares (rural e urbana);

- Simultaneamente, será avaliada a presença de fatores de risco da doença de modo a permitir uma previsão sobre a probabilidade da presença da doença com necessidade de intervenção clínica.

Referências

- Kajerlff U. (2005), *Probabilistic Networks – na introduction to Bayesian Networks and inflence Diagrams*, Aslborg University;

- Barber D. (2012), *Bayesian Reasoning and Machine Learning*, Cambridge University Press.

- Mago,V. K., Prasad, B., Bhatia, A., and Mago, A., *A decision making system for the treatment of dental caries*, in Soft computing applications in business, Prasad B. (Ed.) Vol. 230. Springer, Berlin, pp. 231–242, 2008.