COMP0271: Inteligência Artificial

Raciocínio probabilístico

Professor: Hendrik Macedo

Universidade Federal de Sergipe, Brasil

Incerteza

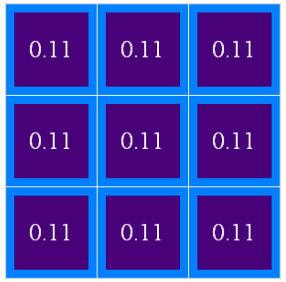
- Situação geral:

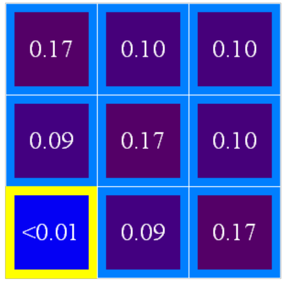
  - **Variáveis observadas (evidência):** agente sabe algumas coisas do mundo (ex: sintomas, valor de sensores, etc.)

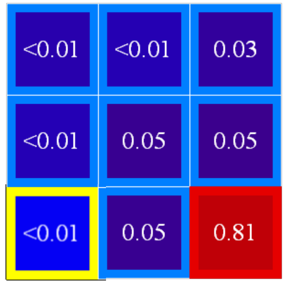
  - **Variáveis não observadas**: agente precisa raciocinar sobre outros aspectos (ex: onde o fantasma deve estar ou que doença o paciente tem, etc.)

  - **Modelo**: agente sabe algo sobre como as variáveis observadas se relacionam com as não observadas.

- **Raciocínio probabilístico** nos fornece um framework para lidar com nossas crenças (*beliefs*) e nosso conhecimento (*knowledge*).







Variáveis Aleatórias

- Um aspecto do mundo sobre o qual pode existir incerteza
  - R = Está chovendo?
  - T = Quente ou frio?
  - D = Distância até o trabalho?
  - L = Onde está o fantasma?
- Domínios
  - R : {true, false}   (usualmente, {+r, -r})
  - T : {hot, cold}
  - D : $[0, \infty)$
  - L : posições possíveis, {(0,0), (0,1), …}

Distribuição de probabilidade

- Temperatura:

$$P(T)$$

T	P
hot	0.5
cold	0.5

- Tempo (clima):

$$P(W)$$

W	P
sun	0.6
rain	0.1
fog	0.3
meteor	0.0

Distribuição de probabilidade

- Variáveis aleatórias não observadas possuem distribuição

$$P(T)$$

T	P
hot	0.5
cold	0.5

$$P(W)$$

W	P
sun	0.6
rain	0.1
fog	0.3
meteor	0.0

Notação compacta:

$$P(hot) = P(T = hot),$$
$$P(cold) = P(T = cold),$$
$$P(rain) = P(W = rain),$$
$$\ldots$$

- Uma distribuição é uma TABELA de probabilidade de valores

- Uma probabilidade (letra minúscula por conveção) é um valor

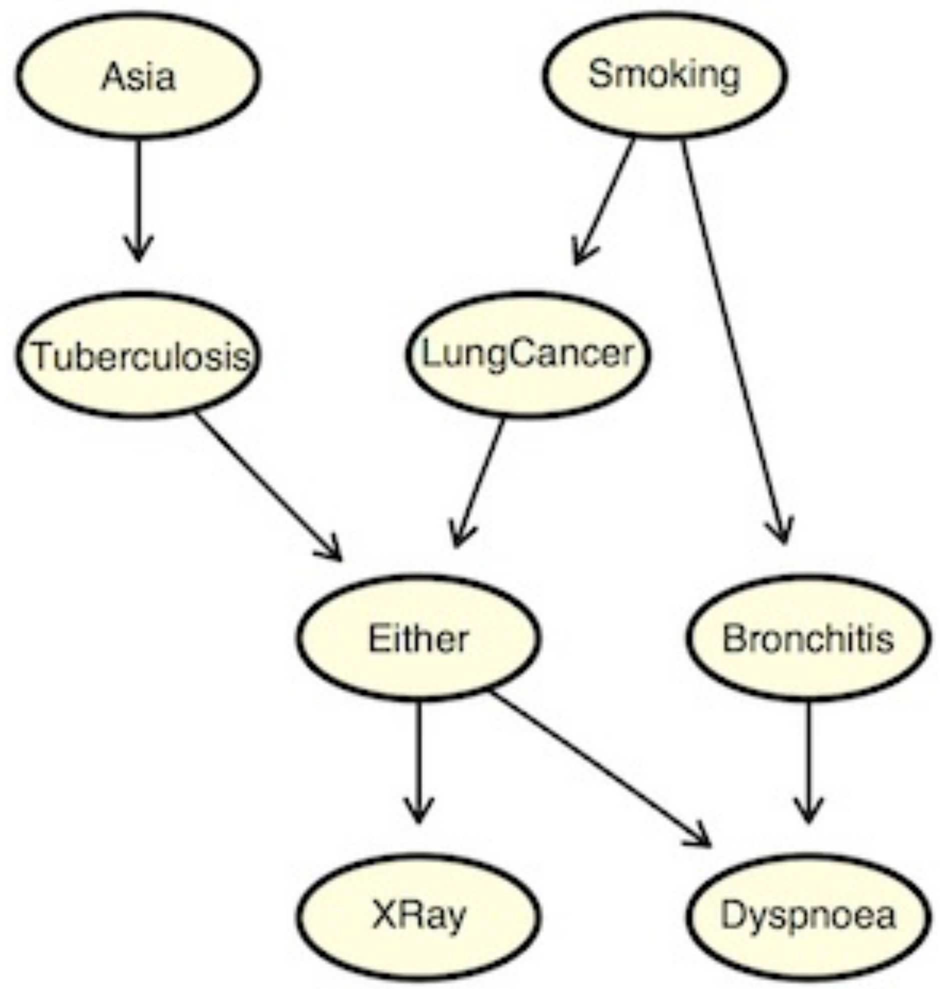
$$P(W = rain) = 0.1$$

$$\forall x \;\; P(X = x) \geq 0 \qquad \sum_x P(X = x) = 1$$

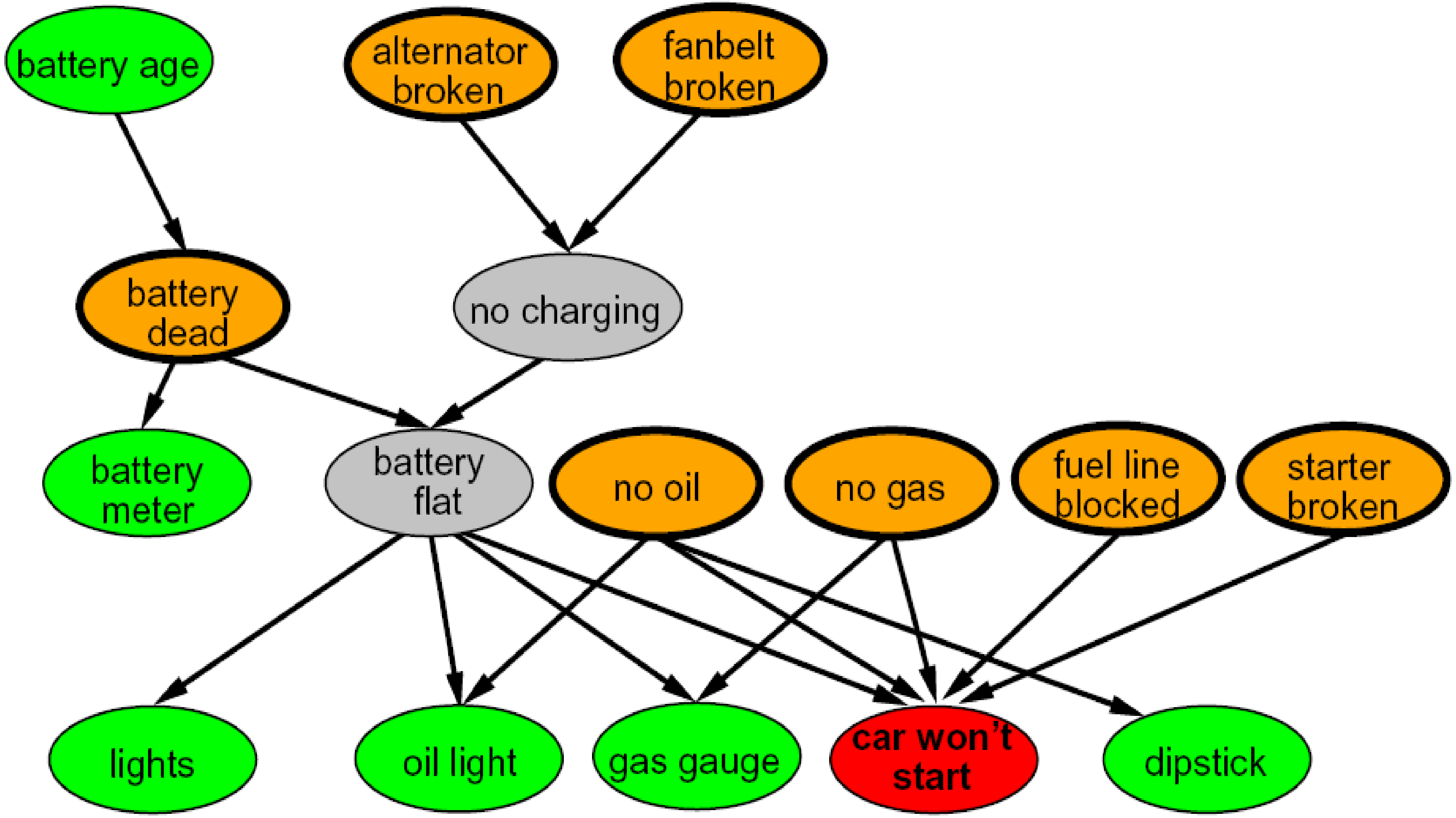
Resumo sobre probabilidade

- Probabilidade condicional $$P(x|y) = \frac{P(x,y)}{P(y)}$$
- Regra do produto $$P(x,y) = P(x|y)P(y)$$
- Regra da cadeia $$\begin{aligned} P(X_1, X_2, \ldots X_n) &= P(X_1)P(X_2|X_1)P(X_3|X_1,X_2)\ldots \\ &= \prod_{i=1}^{n} P(X_i|X_1,\ldots,X_{i-1}) \end{aligned}$$
- X, Y são **independentes** se e somente se: $\forall x,y : P(x,y) = P(x)P(y)$
  $X \perp\!\!\!\perp Y$
- X e Y são **condicionalmente independentes** dado Z se somente se:
  $X \perp\!\!\!\perp Y | Z$
  $$\forall x,y,z : P(x,y|z) = P(x|z)P(y|z)$$
  $$\forall x,y,z : P(x|z,y) = P(x|z)$$

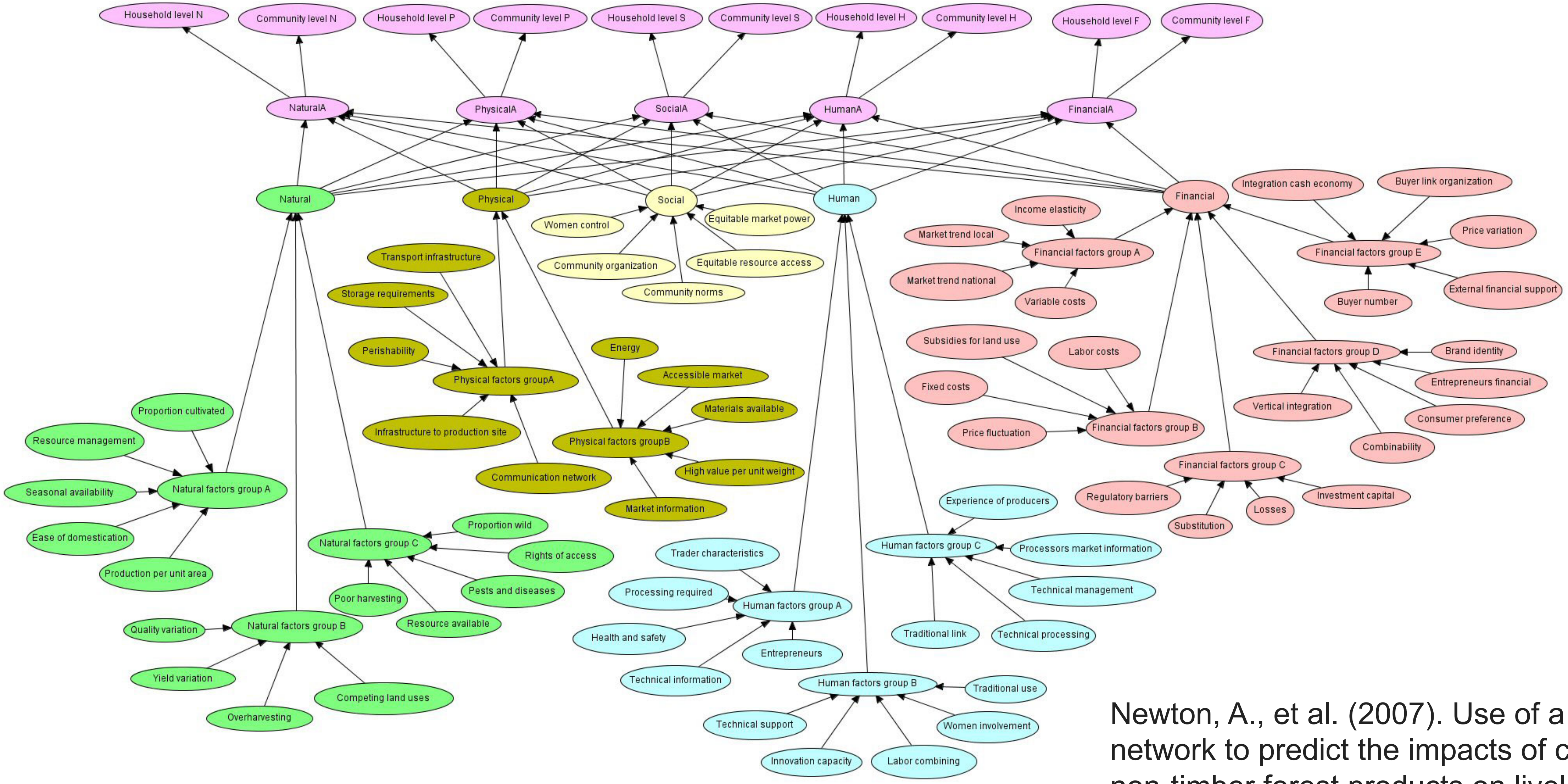
Redes (de Crença) Bayesiana (BBN)



Redes (de Crença) Bayesiana (BBN)



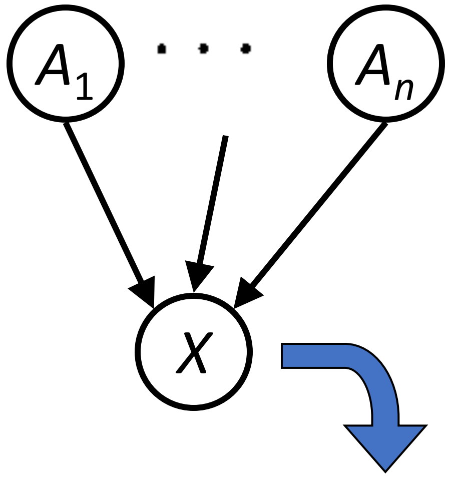
BBN usada para modelar e prever o impacto da comercialização dos **Non-Timber Forest Product** (NTFP) (qualquer recurso biológico encontrado em florestas exceto madeira) em meios de subsistência.



Newton, A., et al. (2007). Use of a Bayesian belief network to predict the impacts of commercializing non-timber forest products on livelihoods. *Ecology and Society*, *11*(2).

Semântica

- Conjunto de nós, um por variável
- Grafo acíclico direcionado
- Distribuição condicional para cada nó
  - Descrição de um processo causal ruidoso



$$P(X|A_1 \ldots A_n)$$

$$P(x_1, x_2, \ldots x_n) = \prod_{i=1}^{n} P(x_i|\mathit{parents}(X_i))$$

Rede bayesiana =

Topologia (grafo) + Probabilidades condicionais locais

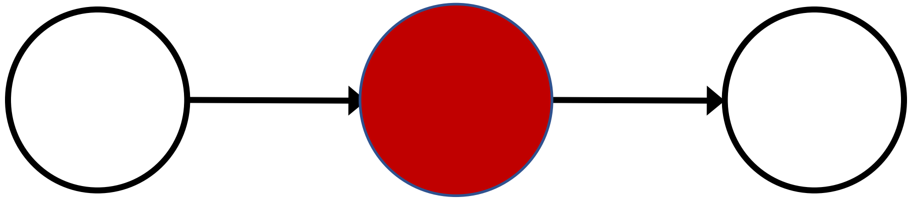
Redes Bayesianas

- Como responder a consultas sobre a distribuição?

D-separation

Método para responder às consultas
Estudo das propriedades de independência para triplas de variáveis

Cadeias causais



X: Pressão baixa   Y: Chuva   Z: Tráfego

$$P(x, y, z) = P(x)P(y|x)P(z|y)$$

- X independente de Z dado Y?

$$P(z|x, y) = \frac{P(x, y, z)}{P(x, y)}$$

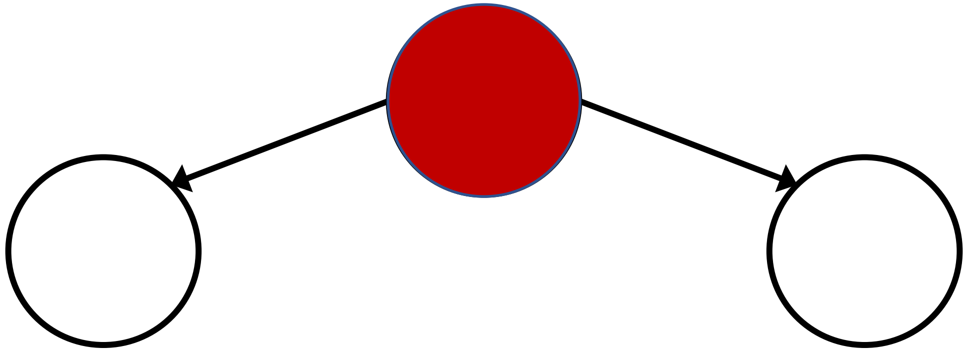
$$= \frac{P(x)P(y|x)P(z|y)}{P(x)P(y|x)}$$

$$= P(z|y)$$

Sim!

- Evidência na cadeia **bloqueia** a influência

Causa comum



$$P(x, y, z) = P(y)P(x|y)P(z|y)$$

- X e Z independente dado Y?

$$P(z|x, y) = \frac{P(x, y, z)}{P(x, y)}$$

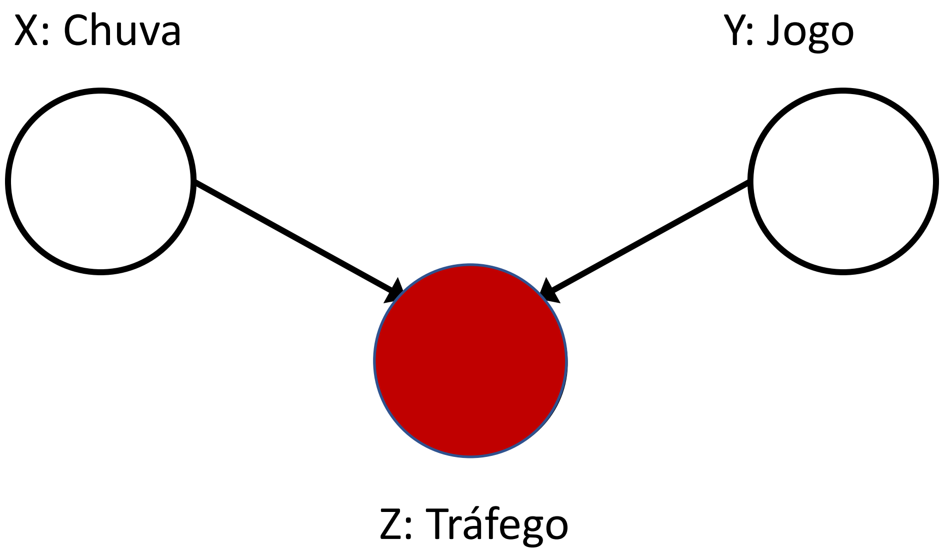
$$= \frac{P(y)P(x|y)P(z|y)}{P(y)P(x|y)}$$

$$= P(z|y)$$

Sim!

- Observar a causa **bloqueia** influência entre efeitos.

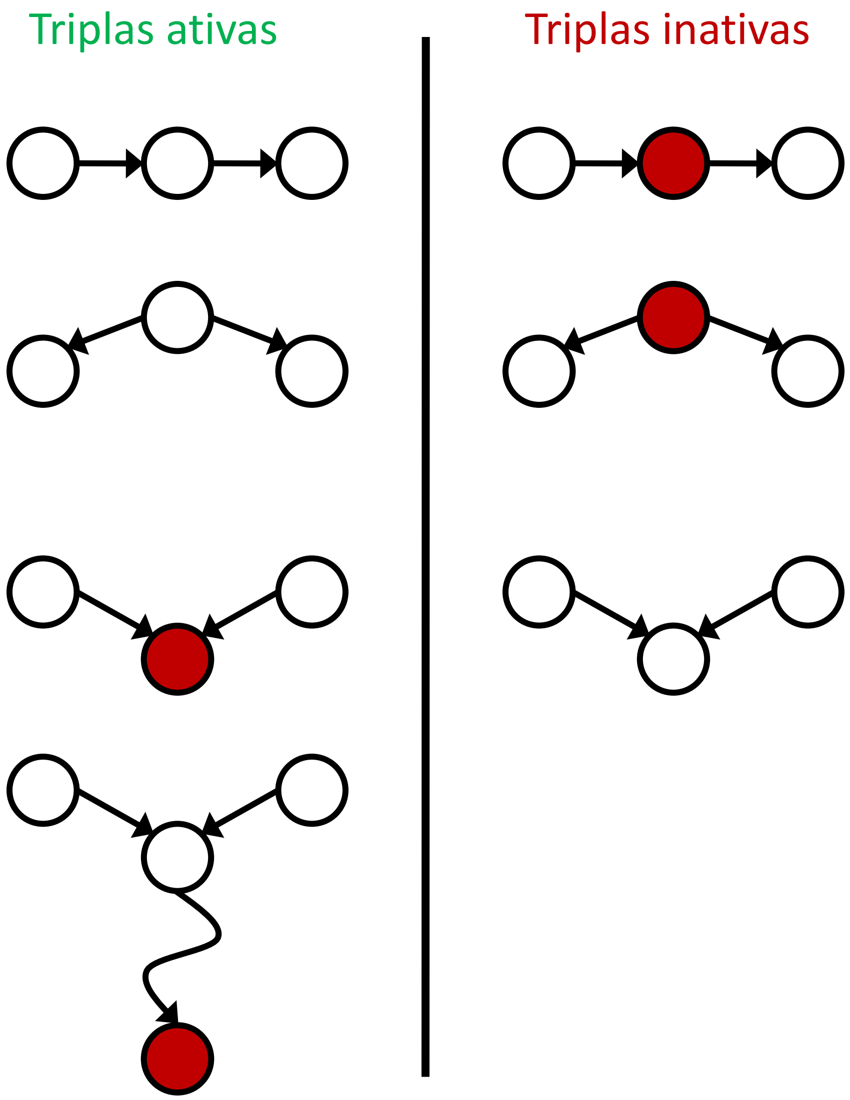
Efeito comum (estruturas v)



- X e Y independentes?
  - *Sim*: Chuva e Jogo não são relacionados
- X e Y independentes dado Z?
  - *Não*: observar tráfego coloca Chuva e Jogo em competição pela **explicação** do raciocínio
- Observar um efeito **ativa** a influência entre causas possíveis

Caminhos ativos / inativos

- X e Y são condicionalmente independentes dadas as variáveis de evidência {Z}?
  - Sim, se X e Y são **"d-separados"** por Z
  - Dentre todos os caminhos (indiretos) de X à Y, se não houverem caminhos ativos = independência!
- Caminhos ativos:
  - Cadeia causal: $A \rightarrow B \rightarrow C$ com B não observado (qualquer direção)
  - Causa comum: $A \leftarrow B \rightarrow C$ com B não observado
  - Efeito comum: (estrutura v): $A \rightarrow B \leftarrow C$ com B ou um de seus descendentes ser observado

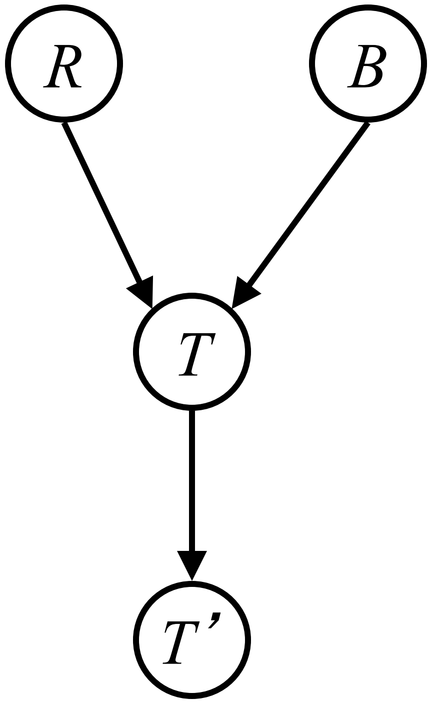


Exemplos

$R \perp\!\!\!\perp B$ *Sim*

$R \perp\!\!\!\perp B | T$ *Não*

$R \perp\!\!\!\perp B | T'$ *Não*



Exemplos

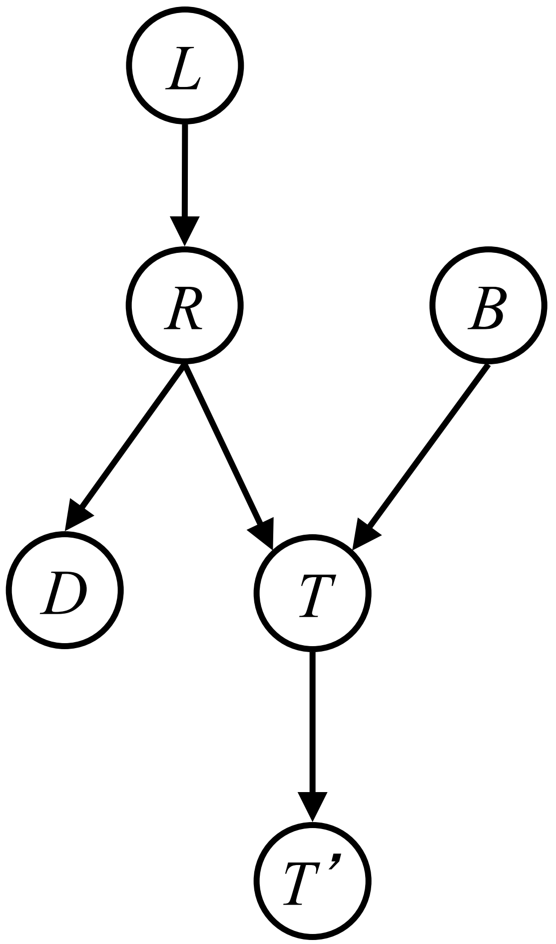
$L \perp\!\!\!\perp T' | T$ Sim

$L \perp\!\!\!\perp B$ Sim

$L \perp\!\!\!\perp B | T$ Não

$L \perp\!\!\!\perp B | T'$ Não

$L \perp\!\!\!\perp B | T, R$ Sim

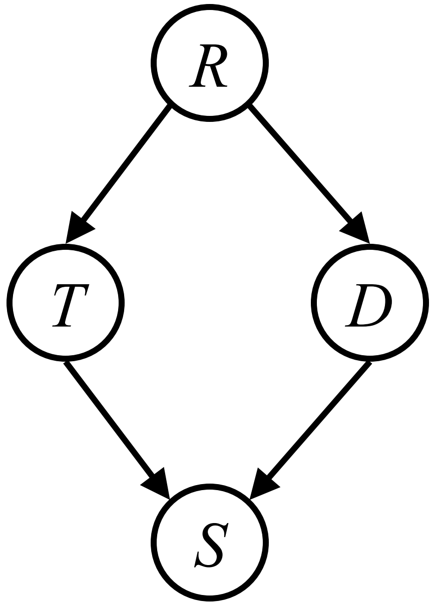


Exemplo

$T \perp\!\!\!\perp D$ Não

$T \perp\!\!\!\perp D | R$ Sim

$T \perp\!\!\!\perp D | R, S$ Não



Inferência

- Calcular algo a partir de uma distribuição de probabilidade conjunta
  - Probabilidade posterior $P(Q|E_1 = e_1, \ldots E_k = e_k)$
  - Explicação mais comum $\text{argmax}_q \, P(Q = q|E_1 = e_1 \ldots)$

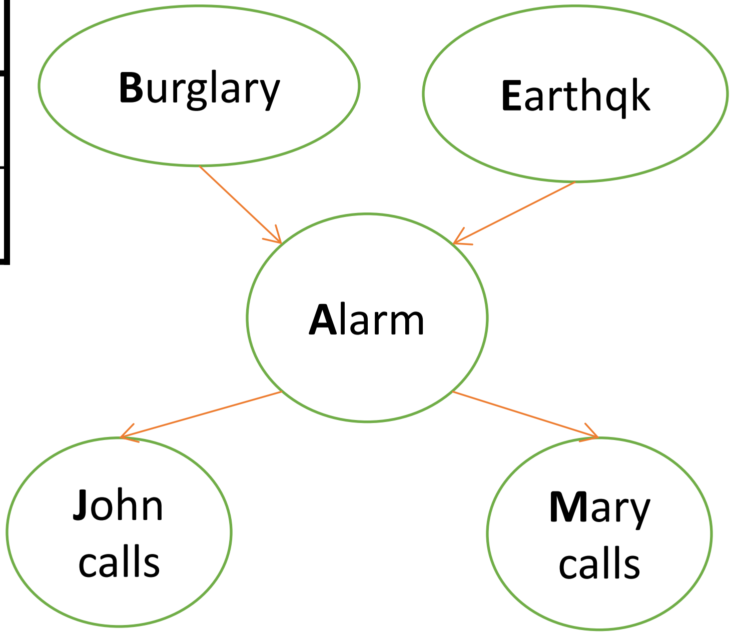
Exemplo: alarme

"Você possui um novo alarme contra ladrões em casa que é muito confiável na detecção de ladrões. Entretanto, ele também pode disparar caso ocorra um terremoto.

Você tem dois vizinhos, João e Maria, os quais prometeram telefonar-lhe no trabalho caso o alarme dispare. João sempre liga quando ouve o alarme. Entretanto, algumas vezes confunde o alarme com o telefone e também liga nestes casos. Maria, por outro lado, gosta de ouvir música alta e às vezes não escuta o alarme."

Exemplo: alarme

B	P(B)
+b	0.001
-b	0.999



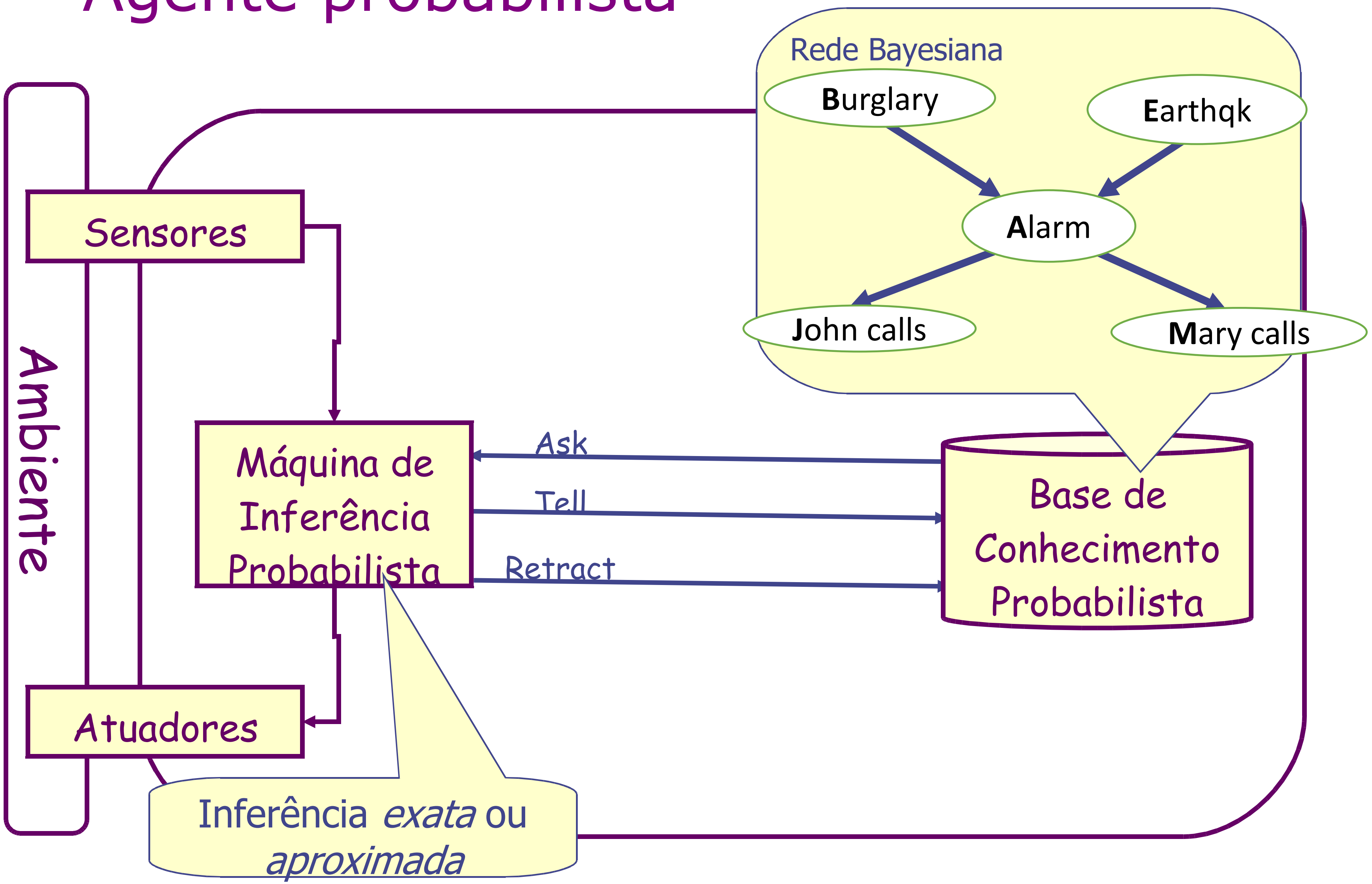
E	P(E)
+e	0.002
-e	0.998

B	E	A	P(A|B,E)
+b	+e	+a	0.95
+b	+e	-a	0.05
+b	-e	+a	0.94
+b	-e	-a	0.06
-b	+e	+a	0.29
-b	+e	-a	0.71
-b	-e	+a	0.001
-b	-e	-a	0.999

A	J	P(J|A)
+a	+j	0.9
+a	-j	0.1
-a	+j	0.05
-a	-j	0.95

A	M	P(M|A)
+a	+m	0.7
+a	-m	0.3
-a	+m	0.01
-a	-m	0.99

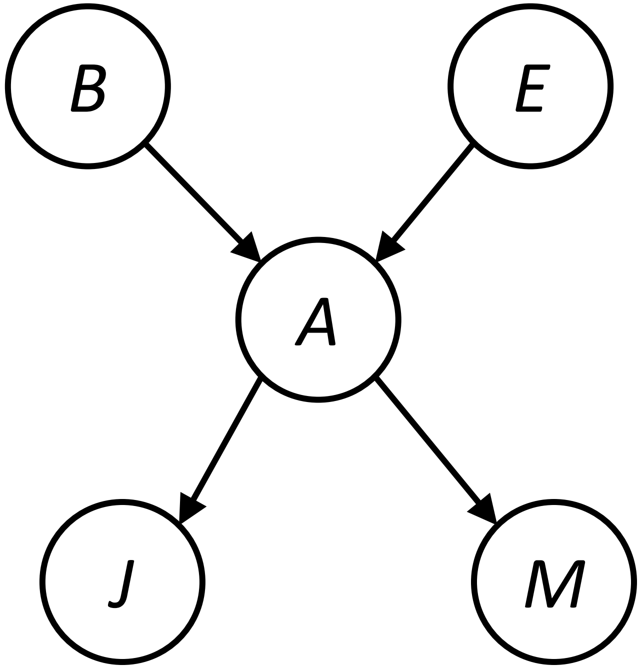
Agente probabilista



Video of Demo BN Applet

Exemplo: alarme

B	P(B)
+b	0.001
-b	0.999



E	P(E)
+e	0.002
-e	0.998

A	J	P(J|A)
+a	+j	0.9
+a	-j	0.1
-a	+j	0.05
-a	-j	0.95

A	M	P(M|A)
+a	+m	0.7
+a	-m	0.3
-a	+m	0.01
-a	-m	0.99

B	E	A	P(A|B,E)
+b	+e	+a	0.95
+b	+e	-a	0.05
+b	-e	+a	0.94
+b	-e	-a	0.06
-b	+e	+a	0.29
-b	+e	-a	0.71
-b	-e	+a	0.001
-b	-e	-a	0.999

$$P(+b, -e, +a, -j, +m) =$$

$$P(+b)P(-e)P(+a|+b,-e)P(-j|+a)P(+m|+a) =$$

$$0.001 \times 0.998 \times 0.94 \times 0.1 \times 0.7$$

Inferência por enumeração

- Caso geral:
  - Variáveis de evidência: $E_1 \ldots E_k = e_1 \ldots e_k$
  - Variável de consulta: Q
  - Variáveis escondidas: $H_1 \ldots H_r$

$$\left.\begin{array}{l} E_1 \ldots E_k = e_1 \ldots e_k \\ Q \\ H_1 \ldots H_r \end{array}\right\} \begin{array}{l} X_1, X_2, \ldots X_n \\ \textit{Todas as variáveis} \end{array}$$

- Queremos:

$$P(Q|e_1 \ldots e_k)$$

- Passo 1: selecionar as entradas consistentes com as evidências
- Passo 2: Somatório em H para computar prob conjunta entre Q e E

$$P(Q, e_1 \ldots e_k) = \sum_{h_1 \ldots h_r} P(Q, \underbrace{h_1 \ldots h_r, e_1 \ldots e_k}_{X_1, X_2, \ldots X_n})$$

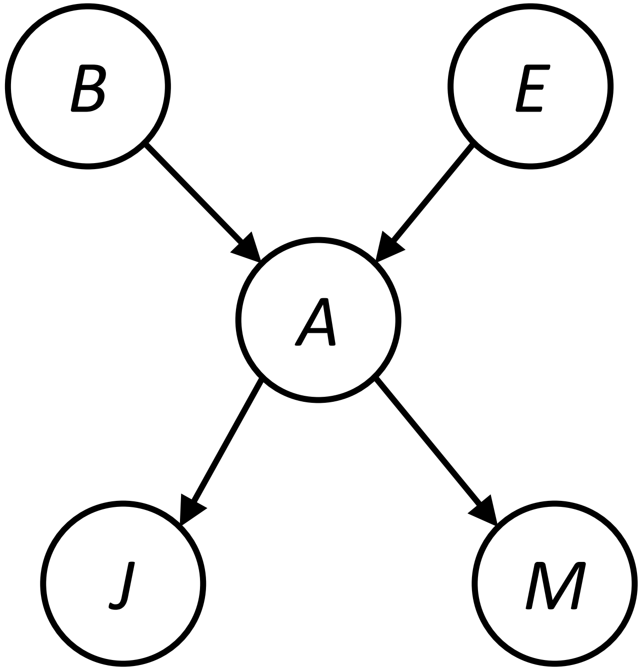
- Passo 3: Normalizar

$$\times \frac{1}{Z}$$

$$Z = \sum_q P(Q, e_1 \cdots e_k)$$

$$\boxed{P(Q|e_1 \cdots e_k) = \frac{1}{Z} P(Q, e_1 \cdots e_k)}$$

Inferência por enumeração



$$P(B \mid +j, +m) \propto_B P(B, +j, +m)$$

$$= \sum_{e,a} P(B, e, a, +j, +m)$$

$$= \sum_{e,a} P(B)P(e)P(a|B,e)P(+j|a)P(+m|a)$$

$$= P(B)P(+e)P(+a|B,+e)P(+j|+a)P(+m|+a) + P(B)P(+e)P(-a|B,+e)P(+j|-a)P(+m|-a)$$
$$P(B)P(-e)P(+a|B,-e)P(+j|+a)P(+m|+a) + P(B)P(-e)P(-a|B,-e)P(+j|-a)P(+m|-a)$$

$= \propto < 0.000593,\ 0.001492 >$

Inferência por enumeração

$$P(M \mid +b, +a) = ?$$