Conhecimento Incerto
Redes de Crença

Profa. Josiane M. P. Ferreira

Texto base:

David Poole, Alan Mackworth e Randy Goebel - "*Computational Intelligence – A logical approach*" – cap 10.

Stuart Russel e Peter Norving - "Inteligência Artificial" - cap 14.

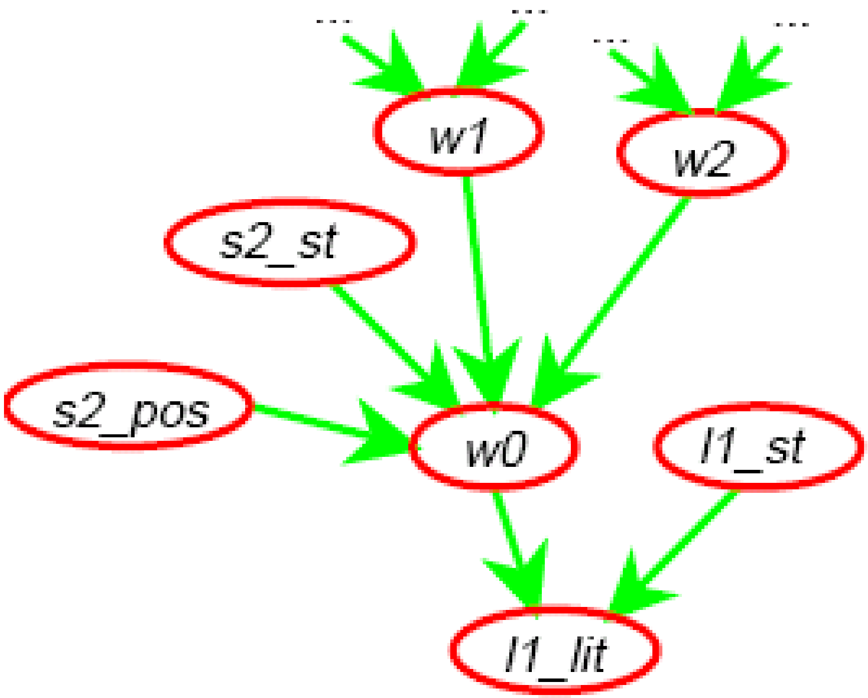
setembro/2008

Idéia das Redes de Crença

- Dado uma variável randômica *V*, existe um pequeno conjunto de variáveis que afetam diretamente o valor de V*.*
  - Qualquer outra variável é independente de *V*, dado os valores para as variáveis que afetam *V* diretamente (variáveis **pais**).
- Este princípio de localidade é explorado nas **Redes de Crença** ou **Rede Bayesianas.**
  - São representações gráficas de independência condicional.
  - São grafos onde os nós são variáveis randômicas e existe um arcos de cada um dos pais do nó para o nó.
  - A independência permite descrever os efeitos diretos no grafo e escrever quais probabilidades necessitam ser especificadas.

Idéia das Redes de Crença

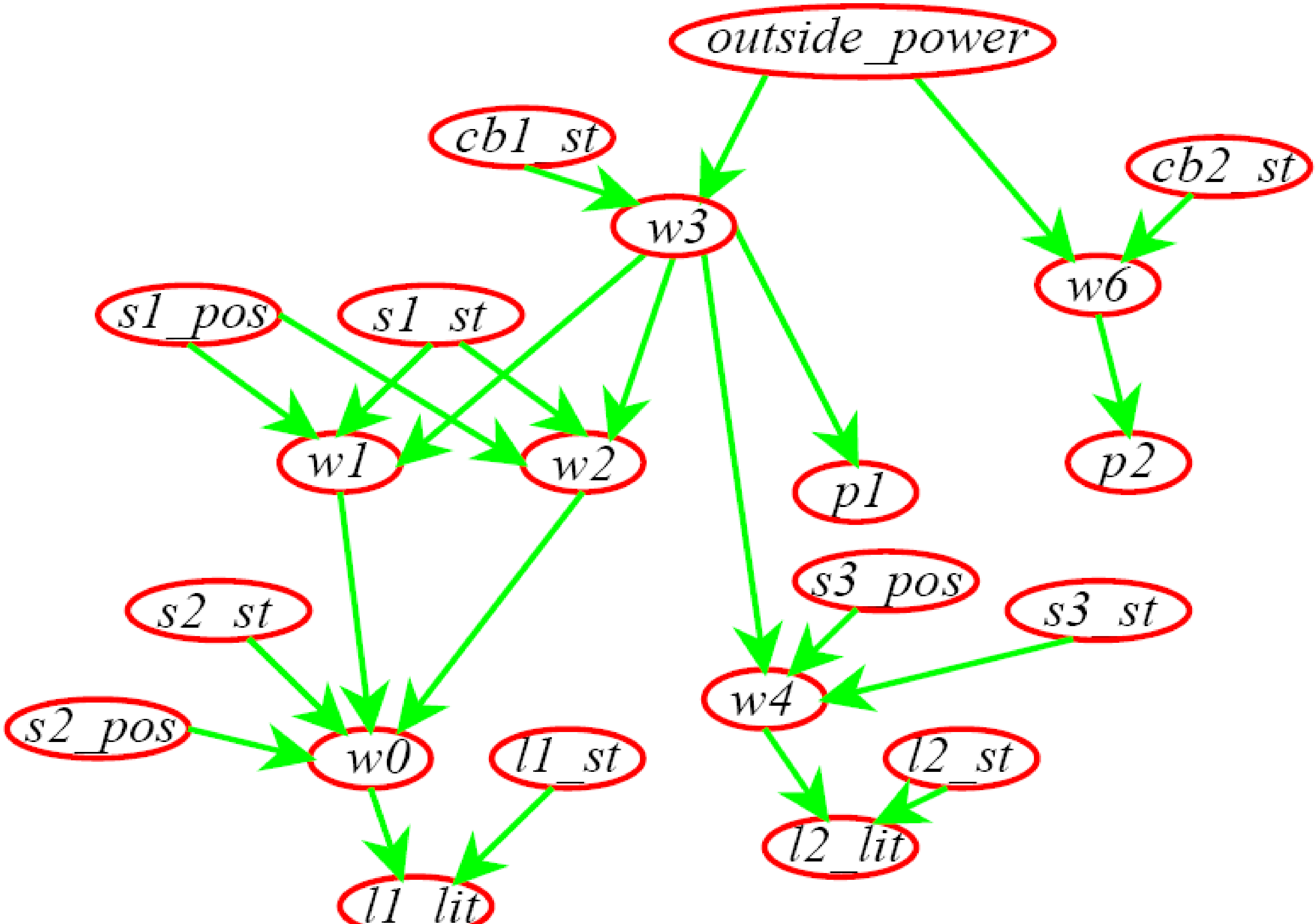
- Se *l1* está acesa (*lit_l1*) depende **somente** do estado da luz (*l1_st*) e se existe energia no fio *w0*.
- Assim, *l1_lit* é independente de outras variáveis dado *l1_st* e *w0*.
  - Em uma rede de crença, *w0* e *l1_st* **são pais** de *l1_lit*.
- Similarmente, *w0* depende **somente** se existe energia no fio *w1*, se existe energia no fio *w2*, a posição da chave *s2* (*s2_pos*), e o estado da chave *s2* (*s2_st*).



O que é uma Rede de Crença

- São DAGs onde os nós são variáveis randômicas.
  - Os pais de um nó *n* são aquelas variáveis das quais *n* depende diretamente.
- São uma representação gráfica de dependência e independência.
  - Uma variável é independente dos seus **não-descendentes**, dados seus pais.
    - Exemplo: seja X uma variável randômica com os pais $Y1$, $Y2$,..., Yn:
      - $P(X=a \mid Y1=v1, Y2=v2, Yn=vn, R) = P(X=a \mid Y1=v1, Y2=v2, Yn=vn)$, se R **não** envolve um descendente de X incluindo o próprio X.
- São constituídas por:
  - Um grafo acíclico com nós rotuladas com variáveis randômicas.
  - Um domínio para cada variável randômica.
  - Um conjunto de probabilidades condicionais para cada variável dado seus pais (incluindo probabilidades a priori para nós sem pais).

Exemplo de uma Rede de Crença



Exemplo de uma Rede de Crença - continuação

- O domínio das variáveis:
  - *w0*, ... , *w6* tem o domínio {*live*, *dead*}.
  - *s1_pos*, *s2_pos*, e *s3_pos* tem o domínio {*up*, *down*}.
  - *s1_st*, tem {*ok*, *upside_down*, *short*, *intermittent*, *broken*}.
- Probabilidade condicional, incluindo os valores para:
  - $P(w1 = live \mid s1\_pos = up \wedge s1\_st = ok \wedge w3 = live)$.
  - $P(w1 = live \mid s1\_pos = up \wedge s1\_st = ok \wedge w3 = dead)$.
    - ... 2 * 5 * 2 = 20 valores.
    - O número de probabilidades que deve ser especificada para uma variável é exponencial no número de pais desta variável.
  - $P(s1\_pos = up)$ => *2 valores.*
  - $P(s1\_st = upside\_down)$ => *5 valores.*

Suposição de independência para Redes de Crença

- As variáveis *s1_pos*, *s1_st*, e *w3* são pais da variável *w1.*
- Se soubermos os valores de *s1_pos*, *s1_st*, e *w3* conhecermos os valores de *l2* e *cb2_st* não afetam a crença se existe energia em *w1.*
- Mas se soubermos os valores de *s1_pos*, *s1_st*, e *w3* ficarmos sabendo que *l1* está ligada, potencialmente muda a crença se existe energia em *w1.*
  - Neste caso a independência não é diretamente aplicável.
- Para variáveis que não tem pais, como *s1_pos*, a independência da rede de crença especifica que, para qualquer *A* que não envolva um descendente de *s1_pos*:
  - *P*(*s1_pos=up* | *A*) = *P*(*s1_pos=up*)

Utilizando Redes de Crença

- Ordene as variáveis de forma que os nós pais venham antes de seus filhos: $X_1, \ldots, X_n$.
- Utilizando a regra da cadeia:

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

- Os pais π_{Xi} de X_i são aqueles predecessores de X_i que tornam X_i independente dos outros predecessores.

$$\pi_{X_i} \subseteq X_1, \ldots, X_{i-1} \text{ and } P(X_i | \pi_{X_i}) = P(X_i | X_1, \ldots, X_{i-1})$$

- Assim,

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

Utilizando Redes de Crença

- A rede de energia pode ser empregada de várias formas
  - Condicionando sobre o estado dos interruptores e disjuntores, se existir energia na rua e a posição dos interruptores, você pode simular o funcionamento das lâmpadas.
  - Dados valores para os interruptores, a energia externa, e se as lâmpadas estão ligadas, você pode determinar a probabilidade posterior de que cada interruptor ou disjuntor estar *OK* ou *Não*.
  - Dado algumas posições e algumas saídas e alguns valores intermediários, você pode determinar a probabilidade de qualquer outra variável na rede.

Redes Crença - Exemplo

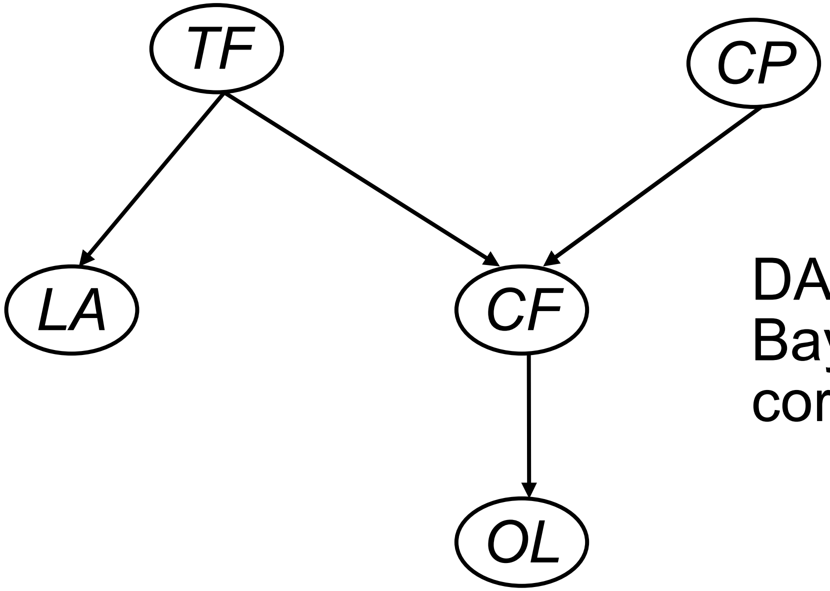
- Problema: quando eu vou p/ casa eu quero saber se alguém da minha família está em casa antes que eu entre.
- Temos:
  - Toda vez que a esposa sai de casa, freqüentemente (mas não sempre) ela liga a luz de fora. (As vezes ela também liga a luz quando está esperando um convidado).
  - Quando ninguém está em casa, o cachorro é sempre deixado pra fora.
  - Se o cachorro tiver problemas intestinais, ele também é deixado pra fora.
  - Se o cachorro estiver pra fora, eu provavelmente vou ouvi-lo latir (embora ele possa não latir, ou eu posso confundir o latido de outro cachorro com o latido do meu).

Redes Crença - Exemplo

- Vamos definir as variáveis aleatórias booleanas:
  - *TF*: Todos estão fora de casa
  - *LA*: A luz está acesa
  - *CF*: O cachorro está fora de casa
  - *CP*: O cachorro tem problemas intestinais
  - *OL*: Eu posso ouvir o cachorro latir
- Dependências condicionais (ou influências causais diretas):
  - *OL* é influenciado diretamente somente por *CF*. Portanto, *OL* é condicionalmente independente de *LA*, de *TF* e de *CP* dado *CF*.
  - *CF* é influenciado diretamente somente por *TF* e *CP*. Portanto, *CF* é condicionalmente independente de *LA* dado *TF* e *CP*.

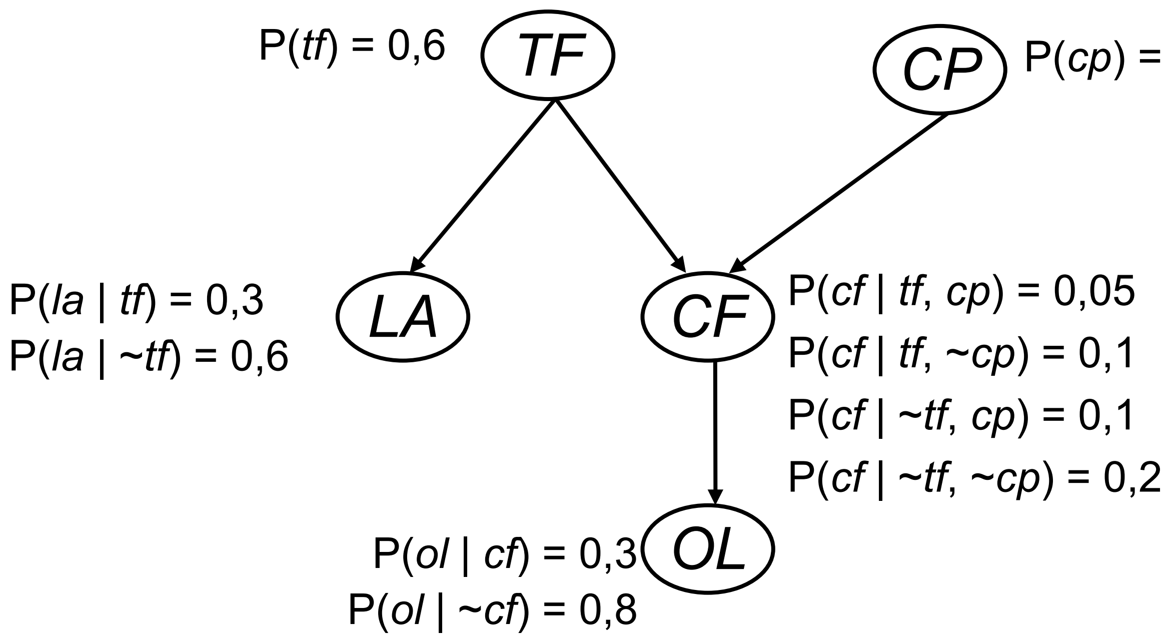
Redes de Crença - Exemplo

- Influências causais diretas:
  - *LA* é influenciado diretamente somente por *TF*. Portanto, *LA* é condicionalmente independente de *CF*, de *OL* e de *CP* dado *TF*.
  - *TF* e *CP* são independentes.



DAG da Rede Bayesiana correspondente

Redes de Crença – informações quantitativas



Total de **10** probb.

Enquanto a distribuição completa de probb. seria uma tabela com **32** probb.

Esta redução é devido ao fato de muitas v.a. serem condicionalmente independentes

Independência em Redes de Crença

- Duas variáveis que não são conectadas diretamente por um arco ainda podem afetar uma a outra.
  - Exemplo: *CP* e *OL* não são ligadas diretamente por um arco, mas *CP* afeta *OL* indiretamente.
- Relações condicionais de independência representadas podem ser facilmente visualizadas na rede:
  - Cada nó *V* é condicionalmente independente de todos os nós que não são descendentes de *V*, dado os pais de *V*.
  - Exemplo: *OL* é condicionalmente independente de *CP*, *TF* e *LA* dado *CF*. Assim, $P(OL \mid CP, CF, TF, LA) = P(OL \mid CF)$.

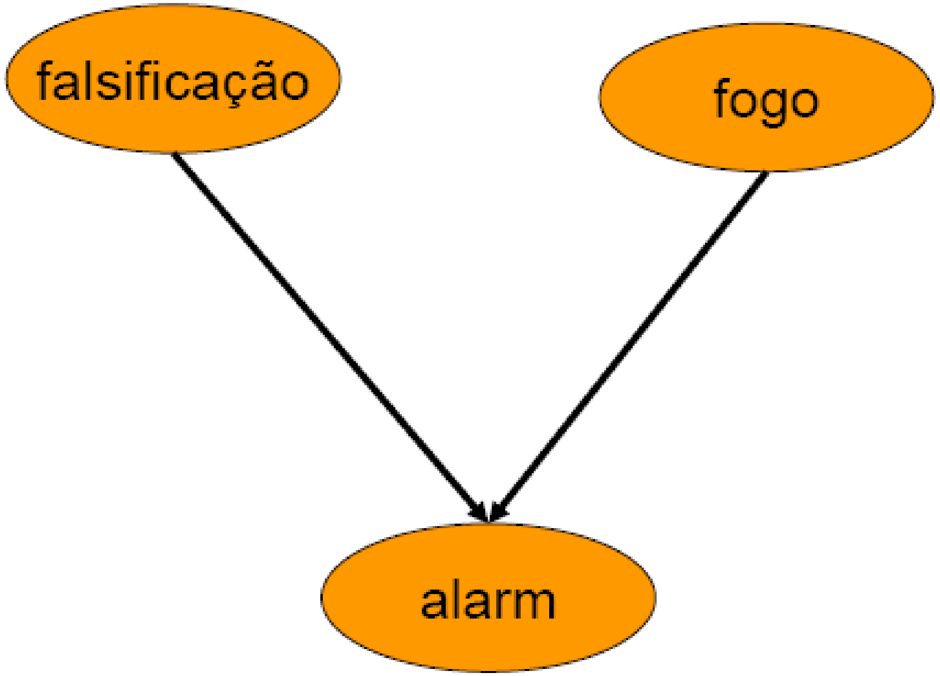
Que variáveis são afetadas pela observação?

- Se você observa a variável Y, as variáveis cuja probabilidade posterior são diferentes de seus valores a priori são:
  - Os antecedentes de Y e
  - Seus descendentes.

- Intuitivamente (se você tiver uma rede de crença causal):
  - Você realiza abdução para as possíveis causas e
  - Predição a partir das causas.

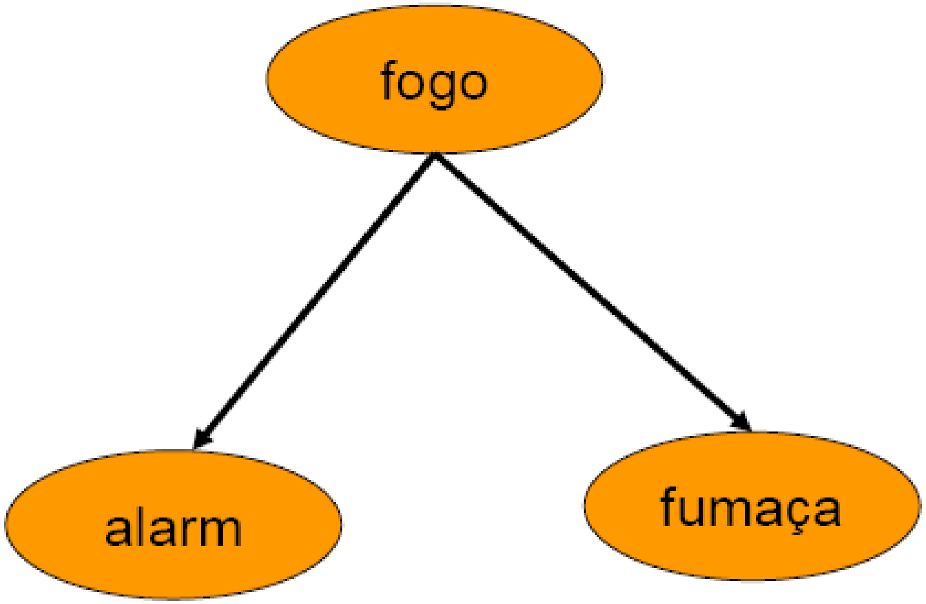
Independência em Redes de Crença – descendentes comuns

- *falsificação* e *fogo* são independentes
- *falsificação* e *fogo* são dependentes dado *alarme*
- Intuitivamente, *falsificação* pode **minimizar a importância** de *fogo*



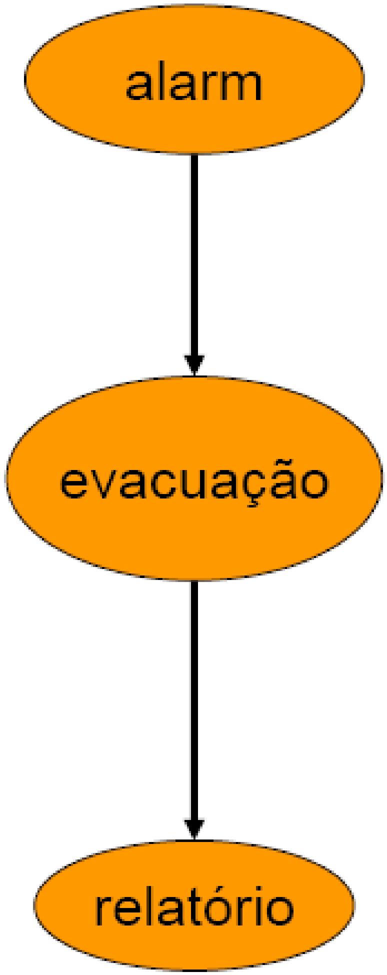
Independência em Redes de Crença – ancestrais comuns

- *alarme* e *fumaça* são dependentes
- *alarme* e *fumaça* são independentes dado *fogo*
- Intuitivamente, *fogo* pode **explicar** *alarme* e *fumaça*; aprendendo um pode afetar o outro pela troca de sua crença em *fogo*.



Independência em Redes de Crença – cadeia

- *alarme* e *relatório* são dependente
- *alarme* e *relatório* são independentes dado *evacuação*
- Intuitivamente, a única forma de que o *alarme* afeta *relatório é por afetar evacuação*.

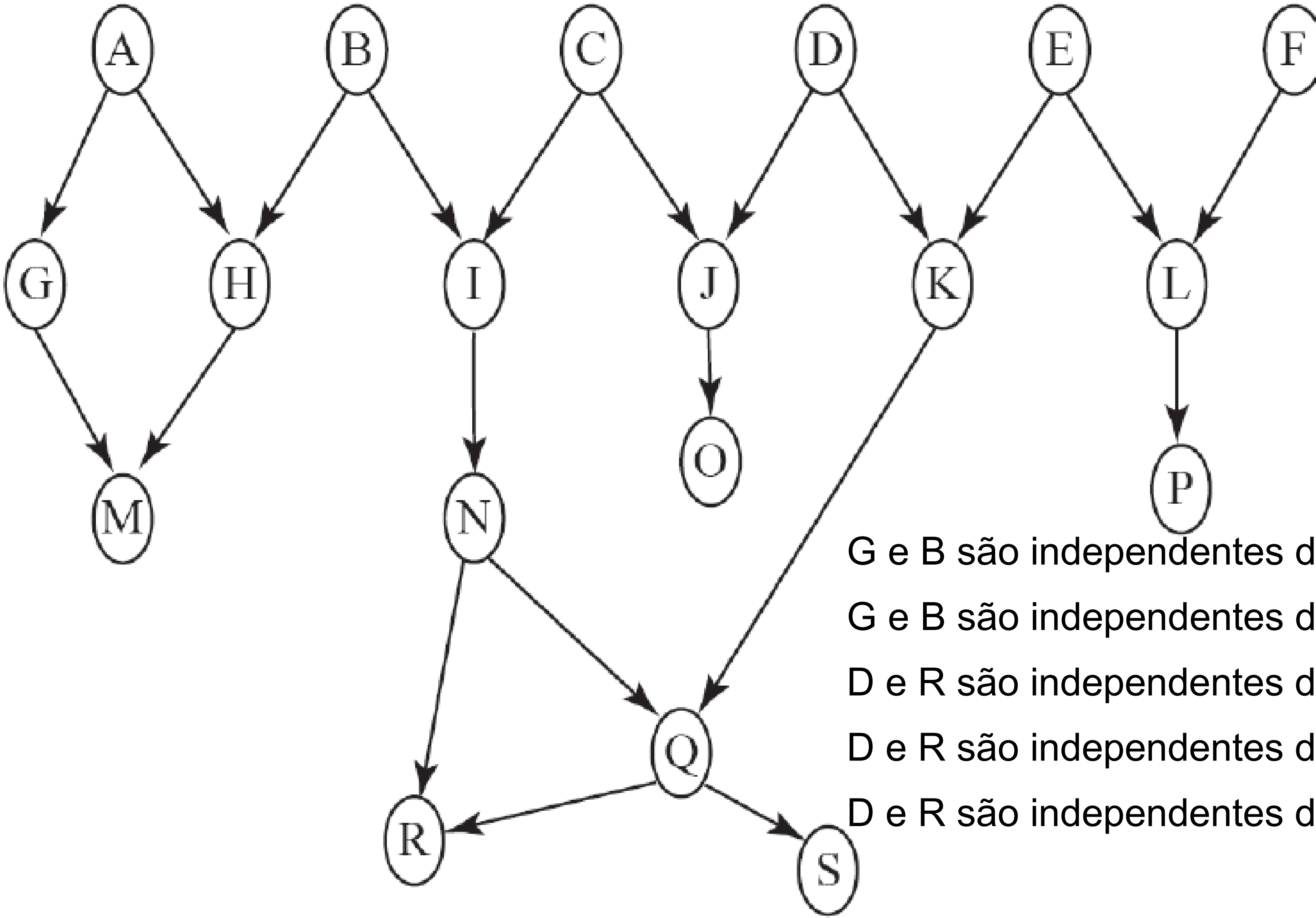


D-Separação

- Dois nós V_i e V_j são condicionalmente independentes dado um conjuntos de evidência E se:
  - Para todo caminho não direcionado da rede entre V_i e V_j, existe algum nó V_b, no caminho com alguma das três seguintes propriedades:
    - V_b está em E, e ambos os arcos no caminho saem de V_b.
    - V_b está em E, e um arco entra em V_b e um arco sai de V_b.
    - Nem V_b nem qualquer descendente de V_b está em E,

    e ambos os arcos no caminho entram em V_b.
  - Neste caso dizemos que E d-separa V_i e V_j.

D-Separação – Exemplo



G e B são independentes dado H?

G e B são independentes dado A?

D e R são independentes dado N?

D e R são independentes dado N e Q?

D e R são independentes dado Q e K?

Calculando probabilidades conjuntas em Redes de Crença

- Objetivo: calcular P(cp*, ~tf, cf, ~la, ol*).
- Eu ouça o cachorro latir, ele está fora com problemas intestinais a luz esteja apagada e ninguém tenha saído.
- Vamos enumerar (1) para a regra da cadeia e ({x},{y}) para independência entre as variáveis do conjunto x dado o conjunto y.
- *P*(cp*, ~tf, cf, ~la, ol*) = *P*(*~tf, cp, ~la , cf, ol*) (pais antes dos filhos) (1)
- = *P*(*~tf*) * *P*(*cp* | *~tf*) * *P*(*~la*| *~tf, cp*) * *P*(*cf* | *~tf, cp, ~la*) * *P*(*ol*|*~tf, cp, ~la, cf*)
- = *P*(*~tf*) * *P*(*cp*) * *P*(*~la*| *~tf*) * *P*(*cf* | *~tf, cp*) * *P*(*ol*|*cf*)
- Onde todos os valores numéricos estão disponíveis diretamente na rede (uma vez que P(~A|B) = 1 – P(A|B)).
- = (1- 0,6) * 0,3 * (1-0,6) * 0,1 * 0,3 = **0,00144**

Inferência por enumeração em uma Rede de Crença

- Qualquer probabilidade condicional pode ser calculada pelo somatório de termos da distribuição conjunta total.

$$P(X \mid e) = \frac{P(X \wedge e)}{P(e)} = \alpha\, P(X \wedge e) = \alpha \sum_{y=v} P(X \wedge e \wedge y)$$

  - onde *X* é a variável de consulta, ***e*** é o conjunto de valores observados e ***y*** são todas as variáveis ocultas.
- A rede de crença fornece uma representação completa da distribuição conjunta total.
- Então, podemos responder uma consulta com o uso de uma rede de crença, calculando-se somas de produtos de probabilidades condicionais da rede.

Inferência por enumeração em uma Rede de Crença - Exemplo

- Suponha que eu queira calcular a probabilidade de eu ouvir o cachorro latir dado que a luz está acesa e o cachorro está com problemas intestinais: *P*(*ol* | *la, cp*).

$$P(ol \mid la \wedge cp) = \frac{P(ol \wedge la \wedge cp)}{P(la \wedge cp)} = \alpha\, P(ol \wedge la \wedge cp)$$

$$= \alpha \sum_{TF} \sum_{CF} P(TF \wedge cp \wedge la \wedge CF \wedge ol)$$

$$= \alpha \sum_{TF} \sum_{CF} P(TF) * P(cp) * P(la \mid TF) * P(CF \mid TF \wedge cp) * P(ol \mid CF)$$

$$= \alpha \{P(cp) * [\sum_{TF} P(TF) * P(la \mid TF) * (\sum_{CF} P(CF \mid TF \wedge cp) * P(ol \mid CF))]\}$$

$$= \alpha \{P(cp) * [\sum_{TF} P(TF) * P(la \mid TF) * (P(cf \mid TF \wedge cp) * P(ol \mid cf) + P(\sim cf \mid TF \wedge cp) * P(ol \mid \sim cf))]\}$$

Inferência por enumeração em uma Rede de Crença - Exemplo

$= \alpha \{P(cp) * [(P(tf) * P(la \mid tf)) * (P(cf \mid tf \wedge cp) * P(ol \mid cf) + P(\sim cf \mid tf \wedge cp) * P(ol \mid \sim cf)) +$

$(P(\sim tf) * P(la \mid \sim tf)) * (P(cf \mid \sim tf \wedge cp) * P(ol \mid cf) + P(\sim cf \mid \sim tf \wedge cp) * P(ol \mid \sim cf))]\}$

$= \alpha \{0,3 * [(0,6 * 0,3) * (0,05 * 0,3 + 0,95 * 0,8) + (0,4 * 0,6) * (0,1 * 0,3 + 0,9 * 0,8)]\}$

$= \alpha\ 0,09585$

$P(\sim ol \mid la \wedge cp) =$

$= \alpha \{P(cp) * [(P(tf) * P(la \mid tf)) * (P(cf \mid tf \wedge cp) * P(\sim ol \mid cf) + P(\sim cf \mid tf \wedge cp) * P(\sim ol \mid \sim cf)) +$

$(P(\sim tf) * P(la \mid \sim tf)) * (P(cf \mid \sim tf \wedge cp) * P(\sim ol \mid cf) + P(\sim cf \mid \sim tf \wedge cp) * P(\sim ol \mid \sim cf))]\}$

$= \alpha \{0,3 * [(0,6 * 0,3) * (0,05 * 0,7 + 0,95 * 0,2) + (0,4 * 0,6) * (0,1 * 0,7 + 0,9 * 0,2)]\}$

$= \alpha\ 0,03015$

$\alpha = 0,09585 + 0,03015 = 0,126$

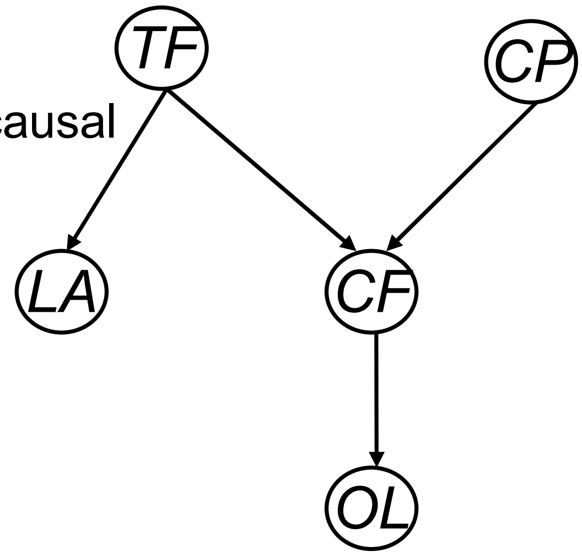
$P(ol \mid la \wedge cp) = 0,09585 / 0,126 = 0,761$

$P(\sim ol \mid la \wedge cp) = 0,03015 / 0,126 = 0,239$

Outros tipos de inferência

- Raciocínio causal:
  - Utilizando a probabilidade condicional e a regra da cadeia:
  - $P(cf \mid cp) = P(cf, tf \mid cp) + P(cf, \sim tf \mid cp)$
  - $P(cf \mid cp) = P(cf \mid tf, cp) * P(tf \mid cp) + P(cf \mid \sim tf, cp) * P(\sim tf \mid cp)$
  - $P(cf \mid cp) = P(cf \mid tf, cp) * P(tf) + P(cf \mid \sim tf, cp) * P(\sim tf)$

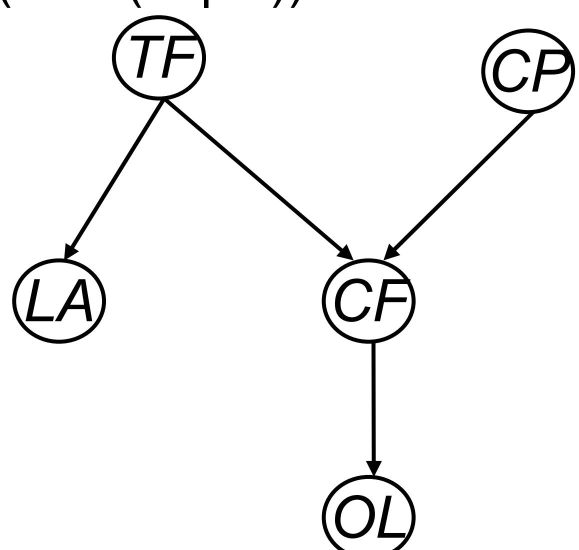
- Raciocínio de diagnóstico:
  - Utilizando a regra de Bayes:
  - $P(cp \mid cf) = (P(cf \mid cp) * P(cp)) / P(cf)$
    - $P(cf \mid cp)$ => pode ser calculado por raciocínio causal
    - $P(cp)$ => na rede
    - $P(cf)$ => por marginalização ou enumeração



Outros tipos de inferência

- Seja *y* uma variável booleana, e o único pai de *x*. Se a evidência *e* não contém qualquer descendente de *x*:
  - $P(x \mid e) = P(x \mid y, e) * P(y \mid e) + P(x \mid \sim y, e) * P(\sim y \mid e)$
  - $P(x \mid e) = P(x \mid y) * P(y \mid e) + P(x \mid \sim y) * P(\sim y \mid e)$
  - $P(x \mid e) = P(x \mid y) * P(y \mid e) + P(x \mid \sim y) * (1 - P(y \mid e))$

- Exemplo:
  - $P(ol \mid la) = \mathbf{P(ol \mid cf)} * P(cf \mid la) + \mathbf{P(ol \mid \sim cf)} * (1 - P(cf \mid la))$
  - ...
  - $P(cf \mid la)$ => *cf* tem mais de um pai



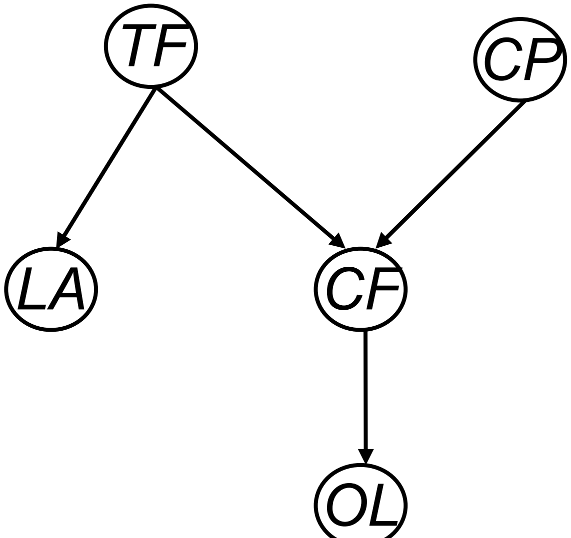
Outros tipos de inferência

- Seja $y_1, y_2, \ldots y_n$ variáveis pais de de *x*. Onde y_i tem o domínio D_i. Se a evidência *e* não contém qualquer descendente de *x*:

$$P(x|e) = \sum_{d \in D} P(x|y = d \wedge e) * P(y = d|e)$$

$$P(x|e) = \sum_{d \in D} P(x|y = d) * P(y = d|e)$$

- Exemplo:
  - *P*(*cf* | *la*) = **P(cf | tf, cp)** * *P*(*tf*, *cp* | *la*) +
    P(cf | tf, ~cp) * *P*(*tf*, ~*cp* | *la*) +
    P(cf | ~tf, cp) * *P*(~*tf*, *cp* | *la*) +
    P(cf | ~tf, ~cp) * *P*(~*tf*, ~*cp* | *la*)



Fatores

- Um **fator** é a representação de uma função de uma tupla de variáveis aleatórias em um número.
- Escreveremos fator *f* sobre as variáveis $X_1, \ldots, X_j$ como $f(X_1, \ldots, X_j)$.
- Podemos atribuir algumas ou todas as variáveis de um fator:
  - $f(X_1 = v_1, X_2, \ldots, X_j)$, onde $v_1 \in dom(X_1)$, é um fator sobre $X_2, \ldots, X_j$.
  - A equação anterior também é escrita como $f(X_1, X_2, \ldots, X_j)_{X1 = v1}$.
  - $f(X_1 = v_1, X_2 = v_2, \ldots, X_j = v_j)$ é um número que é o valor de *f* onde cada X_i tem valor v_i.

Exemplo de Fatores

$r(X, Y, Z)$:

X	Y	Z	val
t	t	t	0.1
t	t	f	0.9
t	f	t	0.2
t	f	f	0.8
f	t	t	0.4
f	t	f	0.6
f	f	t	0.3
f	f	f	0.7

$r(X{=}t, Y, Z)$:

Y	Z	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

$r(X{=}t, Y, Z{=}f)$:

Y	val
t	0.9
f	0.8

$$r(X{=}t, Y{=}f, Z{=}f) = 0.8$$

Exemplo de Fatores – Representação matricial

$$r(X,Y,Z) = \left\{ \begin{array}{l} \begin{bmatrix} x,y,z & x,y,\sim z \\ x,\sim y,z & x,\sim y,\sim z \end{bmatrix} = \begin{bmatrix} 0.1 & 0.9 \\ 0.2 & 0.8 \end{bmatrix} \\ \begin{bmatrix} \sim x,y,z & \sim x,y,\sim z \\ \sim x,\sim y,z & \sim x,\sim y,\sim z \end{bmatrix} = \begin{bmatrix} 0.4 & 0.6 \\ 0.3 & 0.7 \end{bmatrix} \end{array} \right\}$$

$$r(x,Y,Z) = \left\{ \begin{bmatrix} x,y,z & x,y,\sim z \\ x,\sim y,z & x,\sim y,\sim z \end{bmatrix} = \begin{bmatrix} 0.1 & 0.9 \\ 0.2 & 0.8 \end{bmatrix} \right\}$$

$$r(x,Y,\sim z) = \left\{ \begin{bmatrix} x,y,\sim z \\ x,\sim y,\sim z \end{bmatrix} = \begin{bmatrix} 0.9 \\ 0.8 \end{bmatrix} \right\}$$

$$r(x,\sim y,\sim z) = \left\{ \begin{bmatrix} x,\sim y,\sim z \end{bmatrix} = \begin{bmatrix} 0.8 \end{bmatrix} \right\}$$

Multiplicando Fatores

- O produto do fator $f_1(X, Y)$ e $f_2(Y, Z)$, onde Y são as variáveis em comum, é o fator $(f_1 \times f_2)(X, Y, Z)$ definido por:
  - $(f_1 \times f_2)(X, Y, Z) = f_1(X, Y)\, f_2(Y, Z)$

f_1:

A	B	val
t	t	0.1
t	f	0.9
f	t	0.2
f	f	0.8

f_2:

B	C	val
t	t	0.3
t	f	0.7
f	t	0.6
f	f	0.4

$f_1 \times f_2$:

A	B	C	val
t	t	t	0.03
t	t	f	0.07
t	f	t	0.54
t	f	f	0.36
f	t	t	0.06
f	t	f	0.14
f	f	t	0.48
f	f	f	0.32

Multiplicando Fatores – representação matricial

$$f_1(A,B) = \left\{ \begin{bmatrix} a,b & a,\sim b \\ \sim a,b & \sim a,\sim b \end{bmatrix} = \begin{bmatrix} 0.1 & 0.9 \\ 0.2 & 0.8 \end{bmatrix} \right.$$

$$f_2(B,C) = \left\{ \begin{bmatrix} b,c & \sim b,c \\ b,\sim c & \sim b,\sim c \end{bmatrix} = \begin{bmatrix} 0.3 & 0.6 \\ 0.7 & 0.4 \end{bmatrix} \right.$$

$$f_1 \times f_2 = f_B(A,B,C) = \left\{ \begin{array}{l} \begin{bmatrix} a,b,c & \sim a,b,c \\ a,b,\sim c & \sim a,b,\sim c \end{bmatrix} = \begin{bmatrix} 0.03 & 0.6 \\ 0.07 & 0.14 \end{bmatrix} \\ \begin{bmatrix} a,\sim b,c & \sim a,\sim b,c \\ a,\sim b,\sim c & \sim a,\sim b,\sim c \end{bmatrix} = \begin{bmatrix} 0.54 & 0.48 \\ 0.36 & 0.32 \end{bmatrix} \end{array} \right.$$

Sumarizando variáveis

- Podemos ***sumarizar*** uma variável, digamos X_1 com domínio $\{v_1, \ldots, v_k\}$, de um fator $f(X_1, \ldots, X_j)$, resultando em um fator sobre $X_2, \ldots, X_j$ definido por:

$$\left(\sum_{X_1} f\right)(X_2, \ldots, X_j)$$

$$f(X_1 = v_1, X_2, \ldots, X_j) + \ldots + f(X_1 = v_k, X_2, \ldots, X_j)$$

Sumarizando uma variável - Exemplo

f_3:

A	B	C	val
t	t	t	0.03
t	t	f	0.07
t	f	t	0.54
t	f	f	0.36
f	t	t	0.06
f	t	f	0.14
f	f	t	0.48
f	f	f	0.32

$\sum_B f_3$:

A	C	val
t	t	0.57
t	f	0.43
f	t	0.54
f	f	0.46

Sumarizando uma variável – representação matricial

$$f_B(A,B,C) = \begin{cases} \begin{bmatrix} a,b,c & \sim a,b,c \\ a,b,\sim c & \sim a,b,\sim c \end{bmatrix} = \begin{bmatrix} 0.03 & 0.6 \\ 0.07 & 0.14 \end{bmatrix} \\ \begin{bmatrix} a,\sim b,c & \sim a,\sim b,c \\ a,\sim b,\sim c & \sim a,\sim b,\sim c \end{bmatrix} = \begin{bmatrix} 0.54 & 0.48 \\ 0.36 & 0.32 \end{bmatrix} \end{cases}$$

$$f_{\bar{B}}(A,B,C) = \left\{ \begin{bmatrix} a,b,c & a,\sim b,c \\ \sim a,b,\sim c & \sim a,\sim b,\sim c \end{bmatrix} = \begin{bmatrix} 0.57 & 0.54 \\ 0.43 & 0.46 \end{bmatrix} \right.$$

Probabilidades posteriores como conjunções

- Se queremos computar a probabilidade a posterior de *Z* dado a evidência $Y_1 = v_1 \wedge \ldots \wedge Y_j = v_j$:

$$
\begin{aligned}
& P(Z \mid Y_1 = v_1, \ldots, Y_j = v_j) \\
= \quad & \frac{P(Z, Y_1 = v_1, \ldots, Y_j = v_j)}{P(Y_1 = v_1, \ldots, Y_j = v_j)} \\
= \quad & \frac{P(Z, Y_1 = v_1, \ldots, Y_j = v_j)}{\sum_Z P(Z, Y_1 = v_1, \ldots, Y_j = v_j).}
\end{aligned}
$$

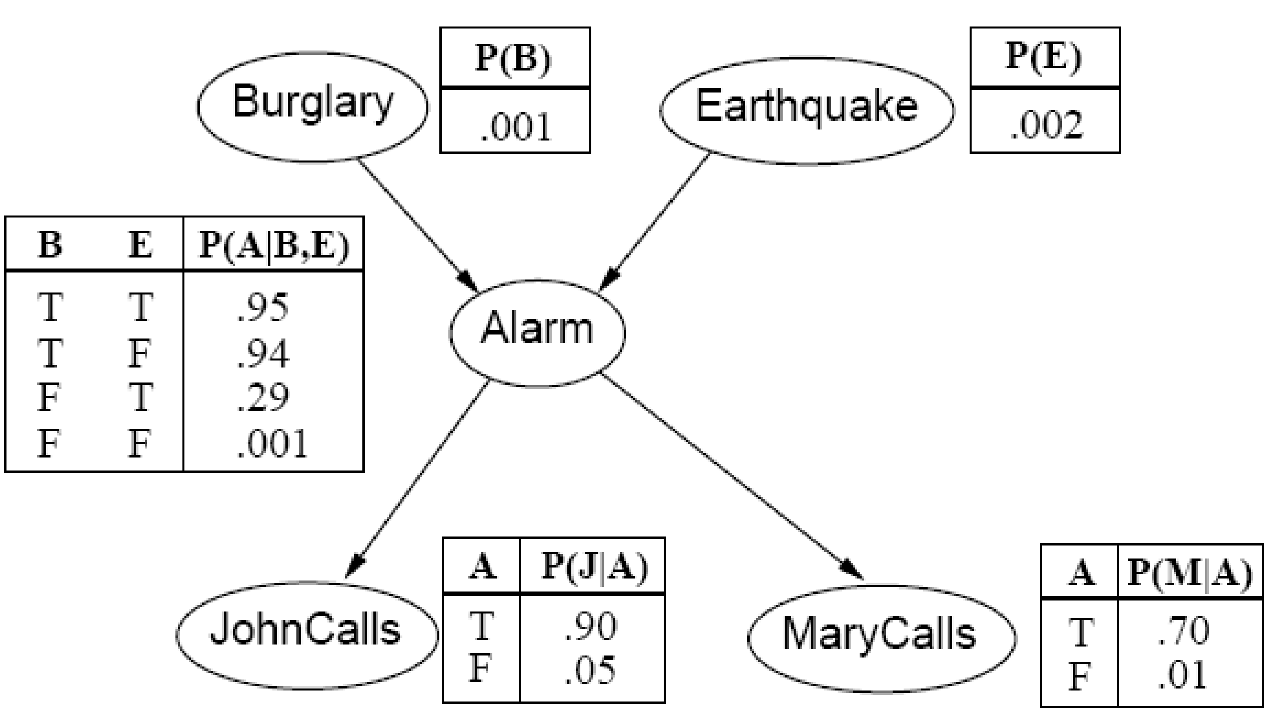
- Assim, a computação se reduz a probabilidade de $P(Z, Y_1 = v_1, \ldots, Y_j = v_j)$.
- Normalizamos ao final.

Exemplo de probabilidade conjunta usando fatores

Construindo Redes de Crença

- Para representar um domínio em uma rede de crença, precisamos considerar:
  - Quais são as variáveis relevantes?
  - Quais valores estas variáveis deveriam ter?
  - Qual é o relacionamento entre elas?
    - Isto deveria ser expressado em termos de influência local?
  - Como os valores de uma variável dependem das variáveis que localmente a influenciam (seus pais)?
    - Isto é expressado em termos de tabelas de probabilidade condicional.

Como as variáveis devem ser colocadas na rede



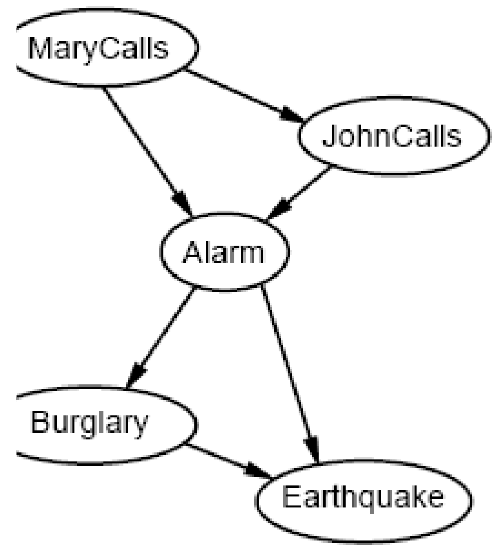
P(B)
.001

P(E)
.002

B	E	P(A|B,E)
T	T	.95
T	F	.94
F	T	.29
F	F	.001

A	P(J|A)
T	.90
F	.05

A	P(M|A)
T	.70
F	.01



- Qualquer uma das redes pode representar a mesma distribuição conjunta total
  - Alguma pode deixar de representar alguma independência condicional, e, desta forma, especificar muitos números desnecessários.