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/2007

Independência entre as variáveis

- Os axiomas de probabilidade são fracos
  - Exemplo: Se existem *n* variáveis binárias em um domínio.
    - Existem 2^n -1diferentes números para serem associados para dar uma distribuição de probabilidade completa para este domínio.
- Para ser capaz de determinar qualquer probabilidade:
  - Temos que começar com um enorme banco de dados de probabilidades condicionais ou probabilidades dos mundos possíveis.
  - Para resolver isso podemos utilizar a idéia de **independência**:
    - O conhecimento na verdade de uma proposição não afeta a crença em outra.
      - O conhecimento na verdade da *cárie* não afeta a crença em *Tempo=chuvoso*.

Independência entre as variáveis

- Suponha um domínio com as variáveis *DorDeDente, UsarAlicate* e *Cárie* – a distribuição de probabilidade completa tem 8 valores.

- Suponha que acrescentamos mais a variável *Tempo* ao domínio.

- A distribuição de probabilidade completa se torna *P*(*DorDeDente, UsarAlicate, Cárie, Tempo*), que tem 32 entradas.
  - Pode ser representada por 4 tabelas, uma para cada valor de *Tempo.*

- Como *P*(*dordedente*, *usaralicate*, *cárie*, *Tempo*=*nublado*) e P(*dordedente*, *usaralicate*, *cárie*) estão relacionadas?

  = *P*(*Tempo*=*nublado* | *dordedente*, *usaralicate*, *cárie*)

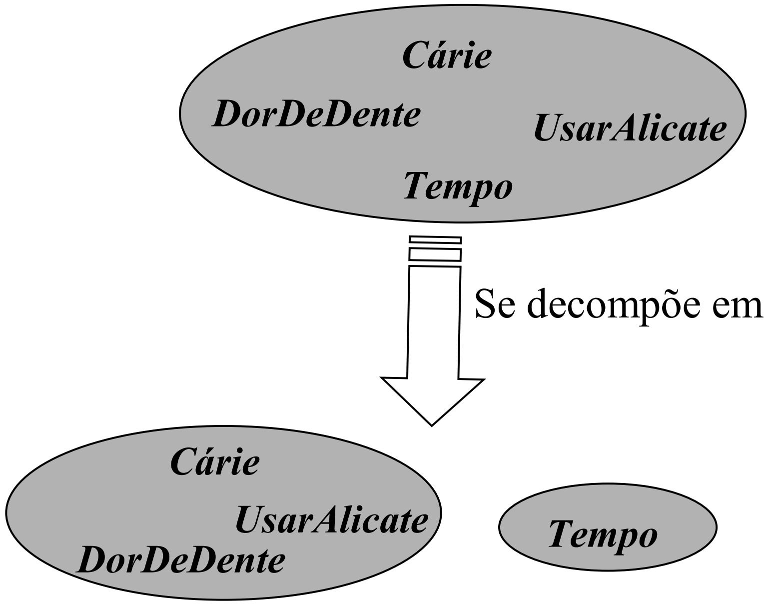
  = *P*(*Tempo*=*nublado*) *P*(*dordedente, alicate, cárie*)

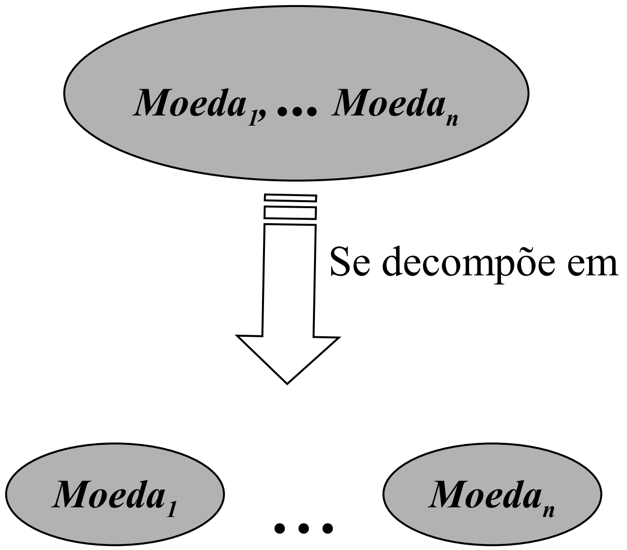
Independência entre as variáveis

- A distribuição de probabilidade completa *P*(*dordedente*, *usaralicate*, *cárie*, *Tempo*=*nublado*)
  - E a tabela de 32 entradas pode ser separada em uma de 4 entradas para *Tempo* e outra de 8 para as outras variáveis.
- Isto é chamado de **independência**: o tempo é independente dos problemas dentários de alguém.
- A independência entre as proposições *a* e *b* pode ser escrita como:
  - $P(a \mid b) = P(a)$ ou $P(b \mid a) = P(b)$ ou $P(a \wedge b) = P(a)P(b)$

Independência

- As asserções de independência em geral se baseiam no conhecimento do domínio.
- Reduzem drasticamente a quantidade de informações necessárias para especificar a distribuição de probabilidade completa.





Independência Condicional

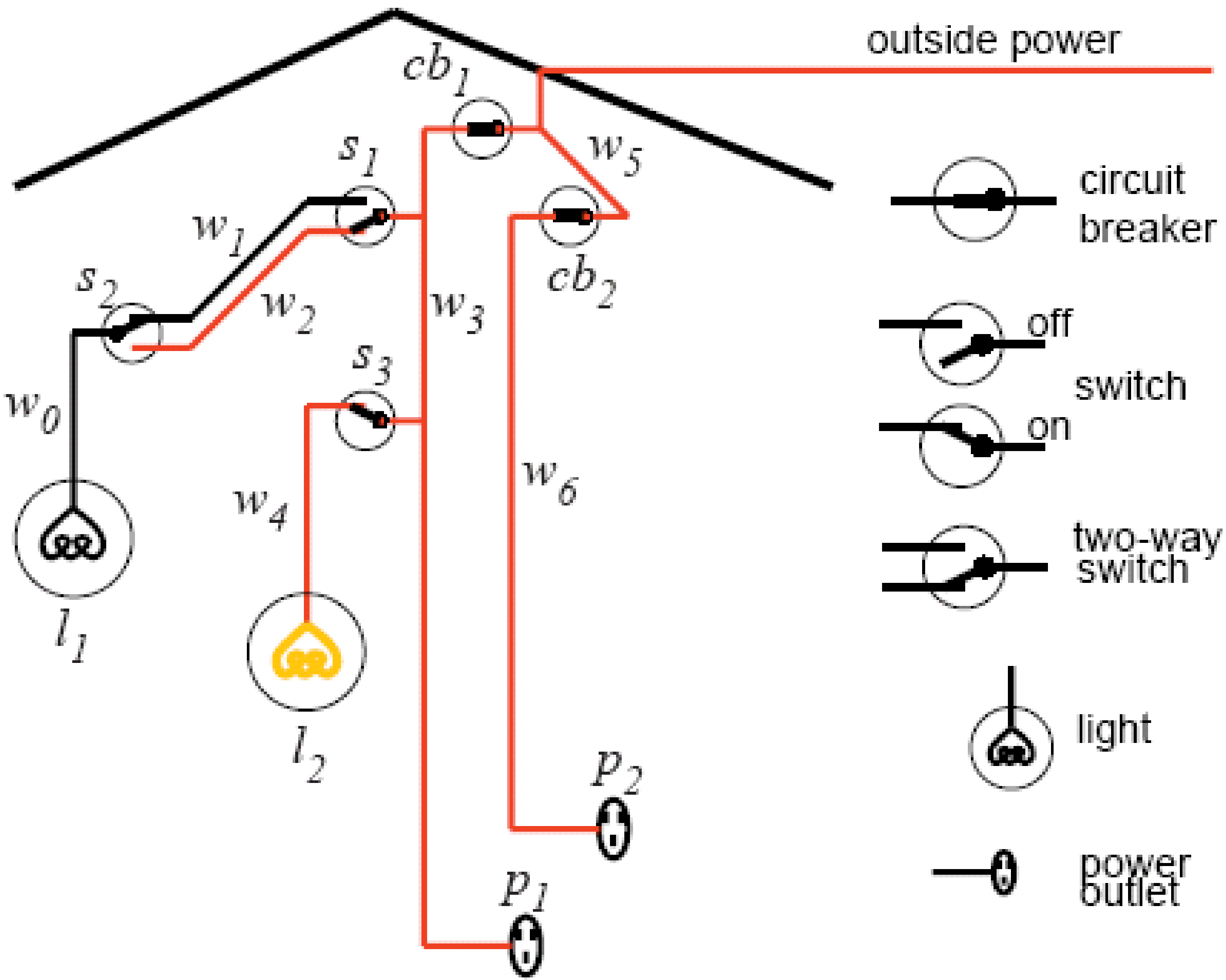
- Acontece quando as variáveis são independentes, **dada** uma (ou um conjunto de) evidência(s).

- Uma variável randômica *X* é independente de uma variável randômica *Y* **dada** uma variável randômica *Z* se, para todo $x_i \in dom(X)$, $y_j \in dom(Y)$ e $z_m \in dom(Z)$,

$$\begin{aligned} & P(X = x_i | Y = y_j \wedge Z = z_m) \\ = \quad & P(X = x_i | Y = y_k \wedge Z = z_m) \\ = \quad & P(X = x_i | Z = z_m). \end{aligned}$$

- Isto é, o conhecimento dos valores de *Y* não afetam a crença do agente no valor de *X*, dado o valor de *Z*.

Exemplo de domínio (Assistente de diagnóstico)



Exemplos de independência condicional

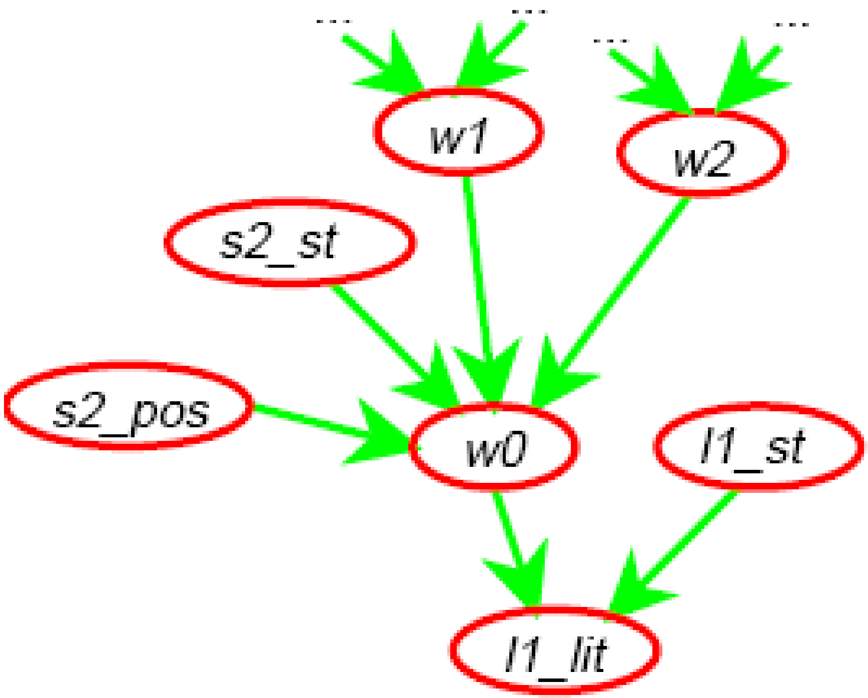
- A identidade da rainha da Inglaterra é independente de se a luz *l1* está acesa dado se existe energia de fora.
- Se existe alguém na sala é independente de se a luz *l2* está acesa dado a posição da chave *s3*.
- Se a luz *l1* está acesa é independente da posição da chave *s2* dado se existe energia no fio *w0*.
- Qualquer outra variável pode ser independente se a luz *l1* está acesa dado se existe energia no fio *w0* e o estado da luz *l1* (se ela está ok ou quebrada).

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

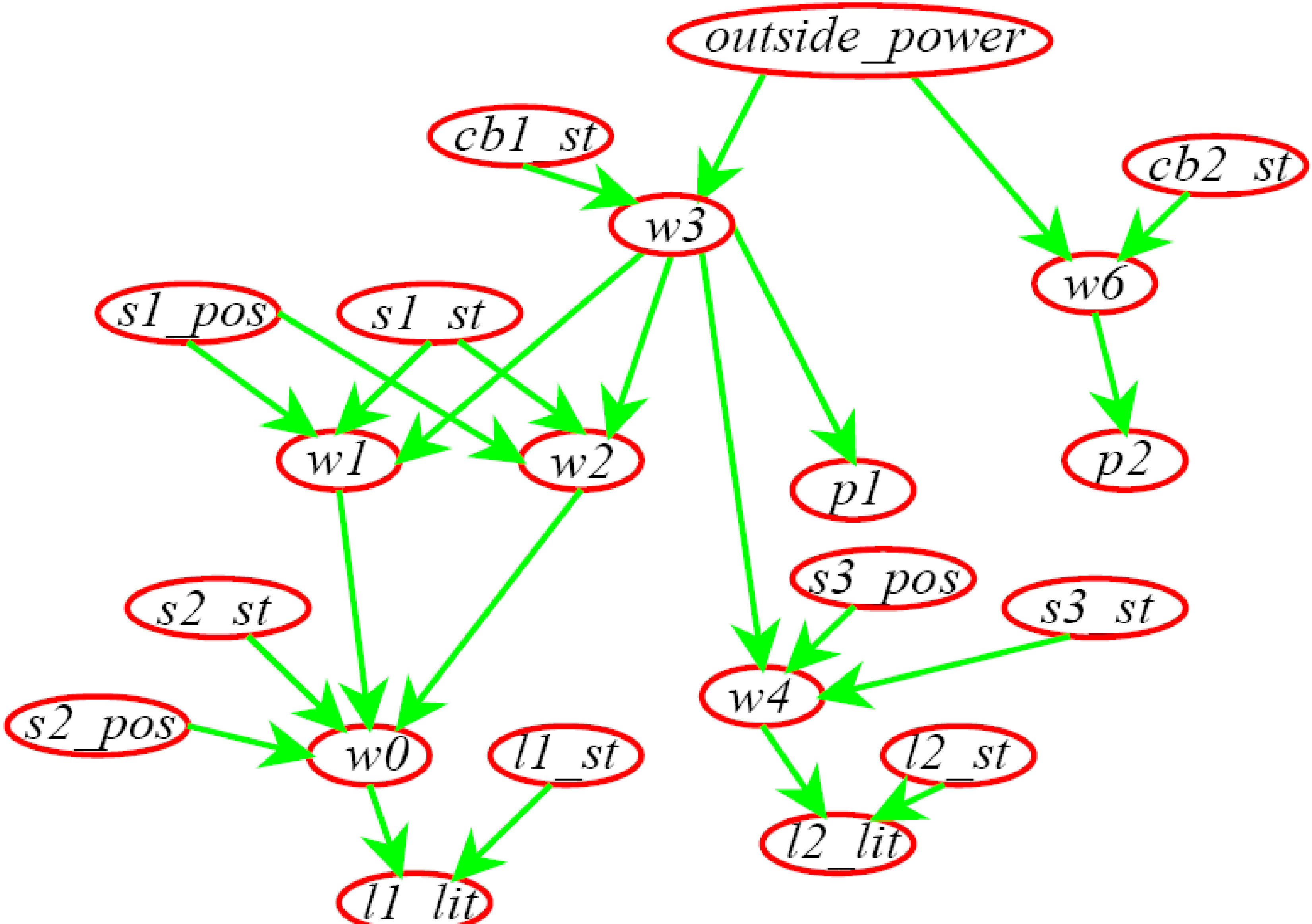
- Se *l1* está acesa (*lit_l1*) depende **somente** do estado da luz (*l1_st*) e se existe energia no fio *w0*.
- Assim, *l1_lit* é independente de outras variáveis dado *l1_st* e *w0*.
  - Em uma rede de crença, *w0* e *l1_st* **são pais** de *l1_lit*.
- Similarmente, *w0* depende **somente** se existe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



Redes 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.
- 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
  - $P(s1\_pos = up)$ => 2 valores
  - $P(s1\_st = upside\_down)$ => 5 valores

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.

Utilizando redes de crença

- Ordene as variáveis de forma que os nós filhos venham antes de seus pais: X_1, ... , 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 sucessores de *Xi* que tornam X_i independente dos outros sucessores.
- Assim,
- $P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i | \pi_{X_i})$

Redes Bayesianas - Exemplo

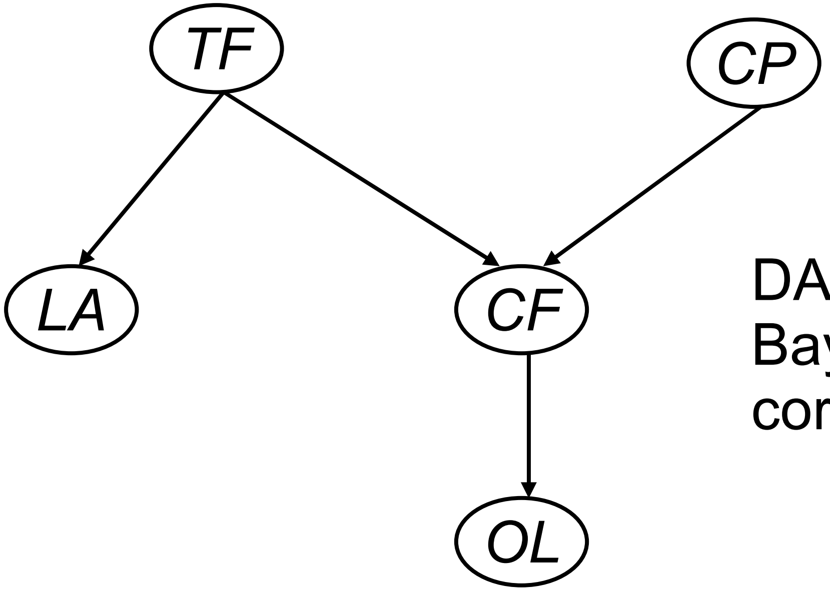
- Problema: quando eu vou p/ casa eu quero saber se alguém da minha família está em casa antes que eu entre.
- Temos:
  - Toda vez que a esposa sai de casa, freqüentemente (mas não sempre) ela liga a luz de fora. (As vezes ela também liga a luz quando está esperando um convidado).
  - Quando ninguém 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 Bayesianas - Exemplo

- Vamos definir as variáveis booleanas aleatórias:
  - *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*.

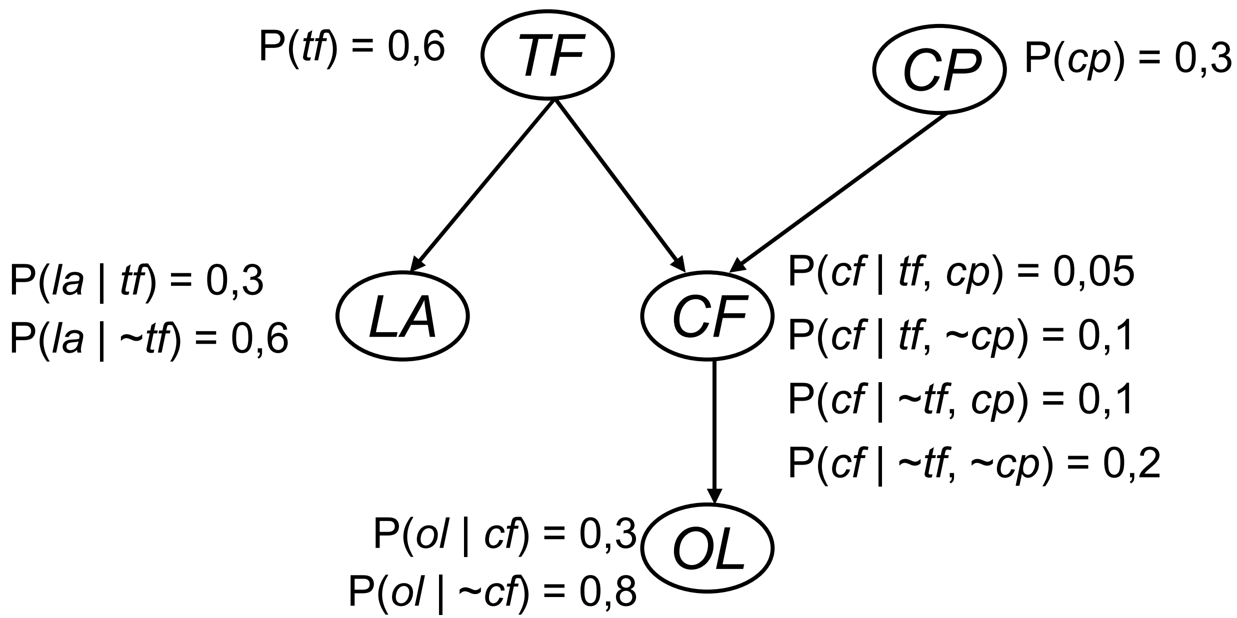
Redes Bayesianas - Exemplo

- 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 Bayesianas – 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 de Bayes:
  - 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* | *CP*, *CF*, *TF*, *LA*) = P(*OL* | *CF*).

Calculando probabilidades conjuntas

- 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(*ol* , *~la* , *cf, ~tf, cp*) (filhos antes dos pais) (1)
- = P(*ol* | *~la*, *cf, ~tf, cp*) * P*(~la, cf, ~tf, cp*) ({OL, LA, TF, CP},{CF})
- = P(*ol* | *cf*) * **P*(~la, cf, ~tf, cp*)** (1)
- = P(*ol* | *cf*) * P*(~la* | *cf, ~tf, cp*) * P(cf*, ~tf, cp*) ({LA, CF, CP},{TF})
- = P(*ol* | *cf*) * P*(~la* | *~tf*) * **P(cf*, ~tf, cp*)** (1)

Calculando probabilidades conjuntas

= P(ol | *cf*) * P(~*la* | ~*tf*) * P(*cf* | ~*tf*, *cp*) * P(~*tf*, *cp*) (1)

= P(*OL* | *CF*) * P(~*LA* | ~*TF*) * P(*CF* | ~*TF*, *CP*) * **P(~*TF* | *CP*)** * P(*CP*) ({TF, CP})

= P(*OL* | *CF*) * P(~*LA* | ~*TF*) * P(*CF* | ~*TF*, *CP*) * P(~*TF*) * P(*CP*)

= (0,3) (1-0,6) (0,1) (1-0,6) (0,3) = **0,00144**

- Onde todos os valores numéricos estão disponíveis diretamente na rede (uma vez que P(~*A*|*B*) = 1 – P(*A*|*B*)).