Redes *Bayesianas*: o que são, para que servem, algoritmos e exemplos de aplicações.

Inteligência Artificial
Profa.: Inês Dutra
Roberto Ligeiro

Resumo

- Introdução.
- Raciocinando sobre incertezas.
- Cálculo de Probabilidades.
- Aplicando a regra de *bayes*.
- Redes *Bayesinas.*
- Inferência em redes *bayesinas*.
- Aplicações.
- Considerações finais.
- Referências.

Introdução

- Sistemas que agem racionalmente
  - Raciocínio Lógico
  - Raciocínio Probabilístico
    - Situações onde não se conhece todo o escopo do problema.
    - Redes *Bayesianas* (início da década de 90)
      - Teoria de probabilidades
      - Teoria de Grafos

Raciocinando sobre incertezas

- “A principal vantagem de raciocínio probabilístico sobre raciocínio lógico é fato de que agentes podem tomar decisões racionais mesmo quando não existe informação suficiente para se provar que uma ação funcionará”[Russel]
- Alguns fatores podem condicionar a falta de informação em uma base de conhecimento:
  - Ignorância Teórica
  - Impossibilidade

Raciocinando sobre incertezas

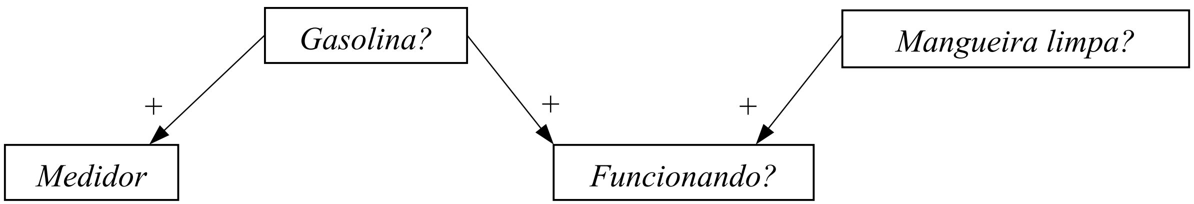
- Utilizar conectivos que manipulem níveis de certeza e não apenas valores booleanos.
  - *"Eu tenho probabilidade 0.8 fazer um bom trabalho de IA"*.
  - "*A probabilidade de um trabalho de IA ser bom é 0.7*".
  - "*A probabilidade de um bom trabalho de IA tirar A é 0.9*".
  - "*Quais são as minhas chances tirar A?*"

- Grafos podem representar relações causais entre eventos.

Raciocinando sobre incertezas

- Considere o domínio: *"Pela manha meu carro não irá funcionar. Eu posso ouvir a ignição, mas nada acontece. Podem existir várias razões para o problema. O rádio funciona, então a bateria está boa. A causa mais provável é que a gasolina tenha sido roubada durante a noite ou que a mangueira esteja entupida. Também pode ser que seja o carburador sujo, um vazamento na ignição ou algo mais sério. Para descobrir primeiro eu verifico o medidor de gasolina. Ele indica ½ tanque, então eu decido limpar a mangueira da gasolina"*.
- Estados:
  - Funcionando?{sim,não}
  - Mangueira Limpa?{sim,não}
  - Gasolina?{sim,não}
  - Medidor{vazio, ½ ,cheio}

Raciocinando sobre incertezas




```
Função RP-Agente(percepção) retorna ação
{
Estático: conjunto de sentenças probabilísticas a respeito do problema.

Calcula novas probabilidades para o estado atual baseado na evidência
   disponível incluindo a percepção atual e a ação anterior.
Calcula as probabilidades para as possíveis ações, dado a descrição das
   ações e as probabilidades atuais.
Seleciona a ação com a maior expectativa.
Retona ação.
}
```

Cálculo de Probabilidades

- Probabilidade incondicional
  - A probabilidade *P(a)* de um evento *a* é um número dentro do intervalo [0,1].
    - *P(a) = 1* sss *a* é certo.
    - Se *a* e *b* são mutuamente exclusivos, então: $P(a \vee b) = P(a) + P(b)$.
- Probabilidade condicional
  - Probabilidade condicional $P(a|b) = x$, pode ser interpretada como: *"Dado o evento b, a probabilidade do evento a é x"*.
    - $P(b|a) = P(a|b)P(b)/P(a)$ - Regra de *Bayes*.
- Tabela de Conjunção de probabilidades
  - $P(A|B)=P(A,B)/P(B)$
  - Tabela *n x m*, representada pela probabilidade de cada configuração (a_i, b_i)
  - Representam todo o domínio
  - Para valores booleanos teríamos 2^n entradas

Cálculo de Probabilidades

- Tabela de Conjunção de probabilidades
  - $P(X) = (a_1, \ldots, a_n); \quad a_i \geq 0; \quad \sum a_i = 1$, onde a_i é a probabilidade de X estar no estado a_i, $P(X=a_i)$.

	b_1	b_2	b_3
a_1	0.4	0.3	0.6
a_2	0.6	0.7	0.4

Tabela1. $P(X|Y)$

  - *Se* $P(Y) = <0.4,0.4,0.2>$, aplicando $P(A/B)=P(A,B)/P(B)$

	b_1	b_2	b_3
a_1	0.16	0.12	0.12
a_2	0.24	0.28	0.08

Tabela2. $P(X,Y)$

  - Com esta tabela pode-se ainda calcular $P(X)$ *e* $P(Y/X)$

Aplicando a Regra de *Bayes*

- Diagnóstico médico:
  - *"um médico sabe que a meningite causa torcicolo em 50% dos casos. Porém, o médico também conhece algumas probabilidades incondicionais que dizem que, um caso de meningite atinge 1/50000 pessoas e, a probabilidade de alguém ter torcicolo é de 1/20."*
    - Aplicando a rede de *Bayes*:
      - $P(M|T) = (P(T|M)P(M))/P(T) = (0.5 \times 1/50000)/(1/20) = 0.0002$
  - Por que não calcular estatisticamente *P(M|T)* ?
    - Surto de meningite => *P(M)* aumenta. *P(M|T)* ?

Redes Bayesianas

- Uma Rede *Bayesiana* consiste do seguinte:
  - Um conjunto de variáveis e um conjunto de arcos ligando as variáveis.
  - Cada variável possui um conjunto limitado de estados mutuamente exclusivos.
  - As variáveis e arcos formam um grafo dirigido sem ciclos (*DAG*).
  - Para cada variável A que possui como pais $B_1, \ldots, B_n$, existe uma tabela $P(A | B_1, \ldots, B_n)$.

Redes Bayesianas

- Exemplo
  - "*Você possui um novo alarme contra ladrões em casa. Este alarme 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.*"
- Estados:
  - Ladrão
  - Terremoto
  - Alarme
  - João
  - Maria

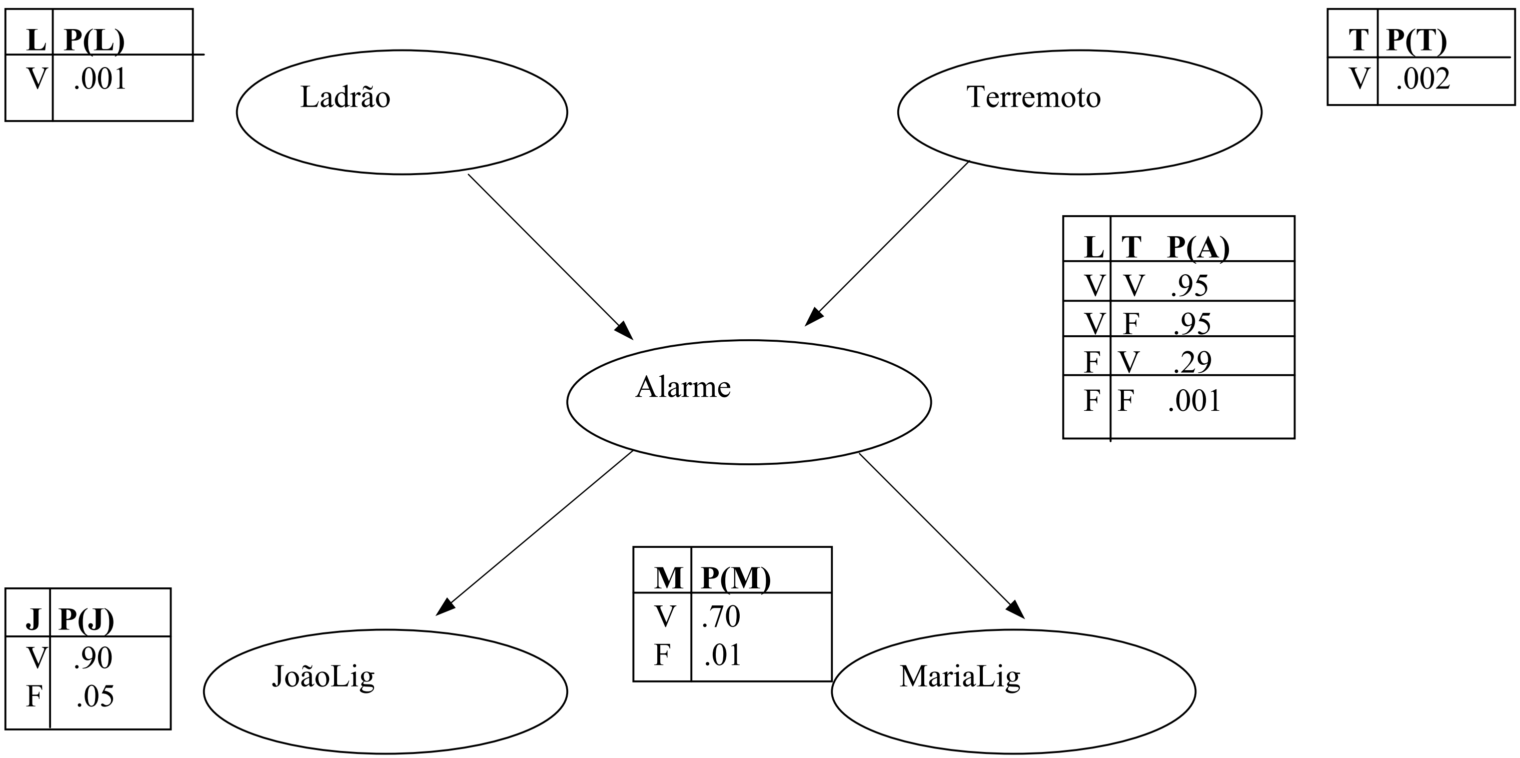
Redes Bayesianas

- Se conhecemos a probabilidade da ocorrência de um ladrão e de um terremoto, e ainda, a probabilidade de João e Maria telefonarem.
- Podemos Calcular $P(Alarme|Ladrão,Terremoto)$:

Ladrão	Terremoto	$P(Alarme\|Ladrão,Terremoto)$	
		Verdadeiro	Falso
Verdadeiro	Verdadeiro	0.95	0.050
Verdadeiro	Falso	0.95	0.050
Falso	Verdadeiro	0.29	0.71
Falso	Falso	0.001	0.999

Redes Bayesianas

- Podemos construir a seguinte rede:



L	P(L)
V	.001

T	P(T)
V	.002

L	T	P(A)
V	V	.95
V	F	.95
F	V	.29
F	F	.001

J	P(J)
V	.90
F	.05

M	P(M)
V	.70
F	.01

Redes Bayesianas

- Considere que se deseja calcular a probabilidade do alarme ter tocado, mas, nem um ladrão nem um terremoto aconteceram, e ambos, João em Maria ligaram, ou $P(J \wedge M \wedge A \wedge \neg L \wedge \neg T)$.

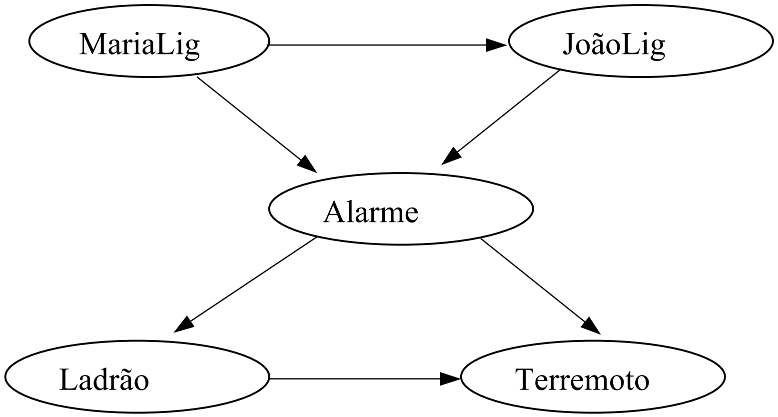
  - $P(J \wedge M \wedge A \wedge \neg L \wedge \neg T) = P(J|A)P(M|A)P(A|\neg L \wedge \neg T)P(\neg L)\ P(\neg T)$
  - $= 0.9 \times 0.7 \times 0.001 \times 0.999 \times 0.998 = 0.00062$

Redes Bayesianas

- Método para construção de redes *bayesianas*:
  - Escolha um conjunto de variáveis X_i que descrevam o domínio.
  - Escolha uma ordem para as variáveis.
  - Enquanto existir variáveis:
    - Escolha uma variável X_i e adicione um nó na rede.
    - Determine os nós *Pais*(X_i) dentre os nós que já estejam na rede e que tenham influência direta em X_i.
    - Defina a tabela de probabilidades condicionais para X_i

Redes Bayesianas

- Método para construção de redes *bayesianas*:
  - *MariaLig*: raiz.
  - *JoãoLig*: Se Maria ligou, então, provavelmente, o alarme tocou. Neste caso, MariaLig influencia *JoaoLig*. Portanto, *MariaLig* é pai de *JoaoLig*.
  - *Alarme*: Claramente, se ambos ligaram, provavelmente o alarme tocou. Portanto, *Alarme* é influenciado por *JoãoLig* e *MariaLig*.
  - *Ladrão*: Influenciado apenas por *Alarme*.
  - *Terremoto*: Se o alarme tocou, provavelmente, um terremoto pode ser acontecido. Entretanto, se existe um *Ladrão*, então as chances de um terremoto diminuem. Neste caso, *Terremoto* é influenciado por *Ladrão* e *Alarme*.



Redes Bayesianas

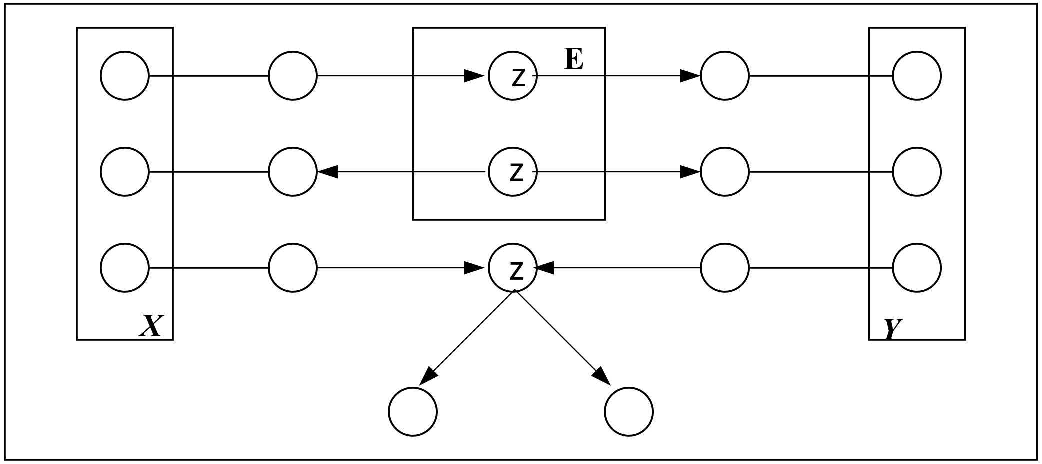
- Método para construção de redes *bayesianas*:
  - Compactação de nós
    - Se cada nó dependesse de todos os outros, teríamos uma tabela de probabilidade de 2^n entradas – para variáveis booleanas -(assim como tabela de conjunção de probabilidades).
    - Localidade estrutural
      - Padrão de relacionamento entre os nós.
      - Uma variável aleatória é influenciada por no máximo *k* outras (seu pais na rede).
      - Por isto: *P(MariaLig|JoaoLig,Alarme,Terremoto,Ladrão) = P(MariaLig|Alarme)*
      - Para uma rede com 20 nós:
        - 2^n = ~1milãho
        - Considerando *k = 5 =>* 640

Inferência em redes *bayesianas*

- Independência condicional
  - Sabemos que um nó é independente de seus predecessores dado seu pai na rede.
  - Porém, para realização de inferências é necessário saber mais a respeito da relação entre os nós.
    - É necessário saber se um conjunto de nós *X* é independente de outro conjunto *Y*, dado que um conjunto de evidências *E* (*X* é *d-separado* de *Y*)
    - Se todo caminho não dirigido entre um nó em *X* e um nó em *Y* é *d-separado* por *E*, então *X* e *Y* são condicionalmente independentes dada a evidência *E*.

Inferência em redes *bayesianas*

- Independência condicional
  - Três possibilidades:

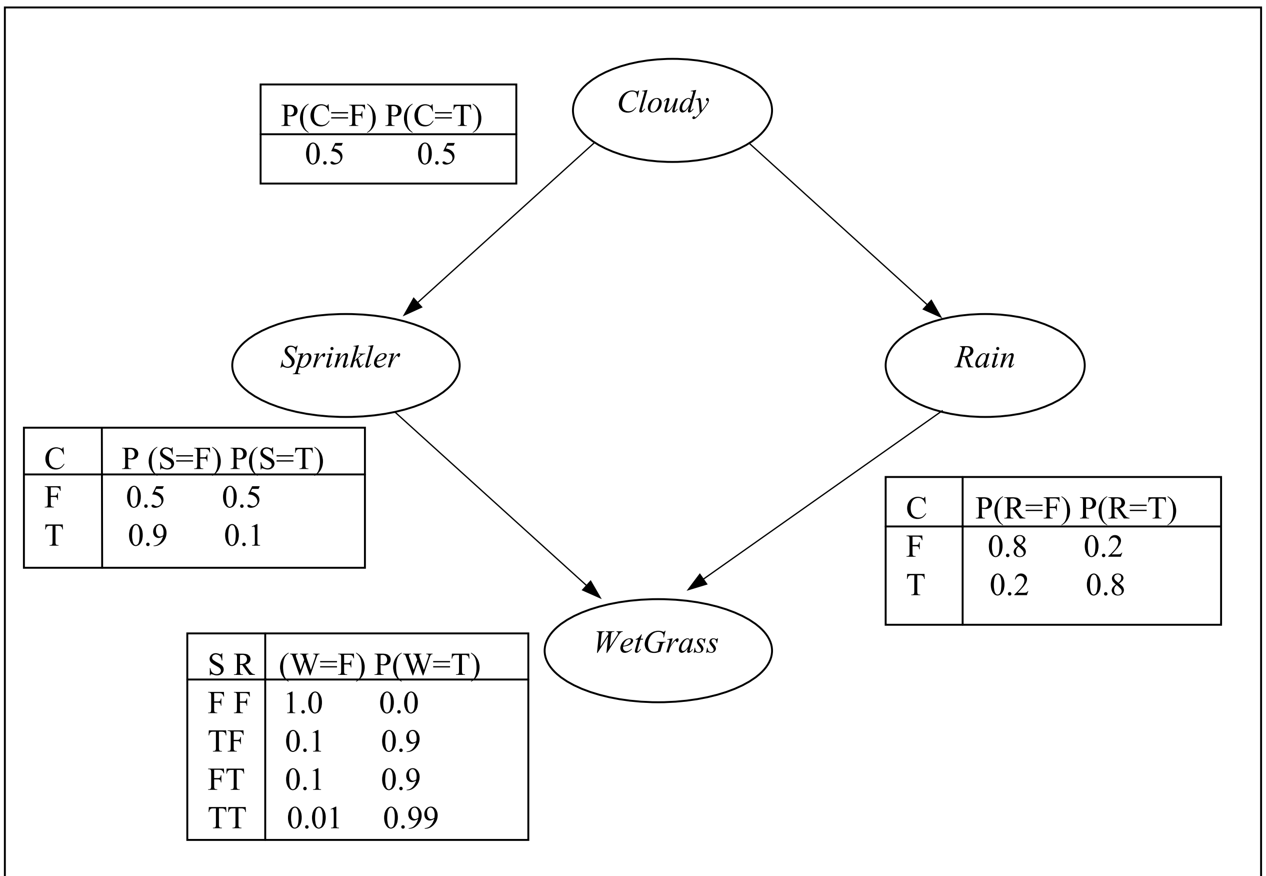


Inferência em redes *bayesianas*

- A tarefa básica de uma inferência probabilística é computar a distribuição de probabilidades posterior para um conjunto de *variáveis de consulta* (*query variables*) => *P(Query|Evidence)*.
- Para o exemplo anterior
  - *Ladrão* constitui uma boa variável de consulta.
  - *JoãoLig, MariaLig* seriam boas variáveis de evidência.

Inferência em redes *bayesianas*

- Exemplo:



- Dado que a grama está molhada, qual a probabilidade de ter chovido?

- Dado que a grama está molhada, qual a probabilidade de o regador ter sido ligado?

Inferência em redes *bayesianas*

- Neste caso, *S* e *R* tornam-se dependentes dado que o nó *W*, a evidência, é filho de ambos, ou seja, não *S* e *R* não são d-separados. Assim, para o calculo de *P(S|W)* deve-se considerar *P(R)* e vice-versa. A equação seria (1 = T e 0 = V):

$$P(S=1|W=1) = P(S=1,W=1)/P(W=1)$$
$$= \Sigma_{c,r}\,(P(C=c_i,\ S=1,R= r_i,\ W=1)/P(W=1)) = 0.2781/0.6471$$

$$P(R=1|W=1) = P(R=1,W=1)/P(W=1)$$
$$= \Sigma_{c,s}\,(P(C = c_i,\ S = s_i,R =1,\ W=1)/P(W=1)) = 0.4581/0.6471$$

Onde *(W = 1)*:

$$P(W=1) = \Sigma_{c,r,s}\,P(C = c_i,\ S = s_i,R = r_i,\ W=1) = 0.6471$$

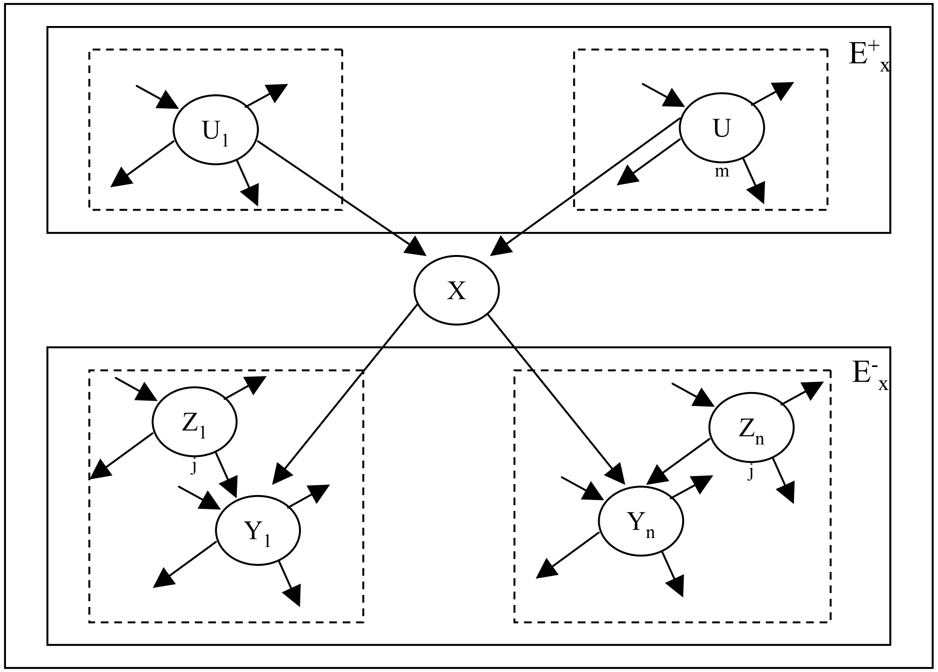
Inferência em redes *bayesianas*

- Inferências podem ser realizadas sobre redes *Bayesianas* para:

  - Diagnósticos: Dos efeitos para as causas. Dado *JoaoLig, P(Ladrão|JoaoLig)*
  - Causas: De causas para efeitos. Dado *Ladrão, P(JoaoLig|Ladrão)*

Inferência em redes *bayesianas*

- Algoritmo para inferências
  - Pode ser *NP-Hard*, dependendo de como o problema foi modelado.
  - Para uma classe de redes possui sempre tempo linear: redes simplesmente conexas.



Inferência em redes *bayesianas*

- Estratégia geral:
  - 1. Expressar *P(X/E)* em relação a E^-_x, E^+_x.
  - 2. Computar a contribuição de E^+_x através de seus efeitos em *Pais(X)*, e então transportar tais efeitos para *X*.
  - 3. Computar a contribuição de E^-_x através de seus efeitos em *Filhos(X)*, e então transportar tais efeitos para *X*. Note que computar os efeitos nos filhos de *X* é uma recursão do problema de computar os efeitos em *X*.

- Chamadas recursivas a partir de *X* por todos seus arcos.
  - Termina em nós de evidência, raízes e folhas da árvore.
  - Cada chamada recursiva exclui o nó que a chamou, desta forma, a árvore é coberta apenas uma vez.

Inferência em redes *bayesianas*

- Inferência em Redes *Bayesianas* Multiconectadas
  - *Clustering* – Transforma probabilisticamente (não topologicamente) a rede em uma rede simplesmente conexa.
  - *Conditioning* – Faz uma transformação na rede instanciando variáveis em valores definidos, e então, e então produz uma rede simplesmente conexa para cada variável instanciada.
  - *Stochastic simulation* – Usa a rede para gerar um grande número de modelos concretos de um domínio. A partir destes modelos o algoritmo calcula uma aproximação de uma inferência.

Aplicações

- *Pathfinder*, Heckerman 1990. Stanford – Sistema para diagnósticos de problemas nas glândulas linfáticas.
- *Map Learning*, Ken Basye 1990. Brown University – Este projeto combina problemas de diagnostico e teoria de decisão. Um robô deve percorrer um "labirinto", procurando aprender os caminhos percorridos e, ao mesmo tempo, explorar caminhos desconhecidos.
- *AutoClass*, NASA´s Ames Research Center, 1998 - Sistema de exploração e aquisição de conhecimento espacial. Este projeto está desenvolvendo uma rede *bayesiana* que permita a interpolação automática de dados espaciais oriundos de diferentes observatórios e planetários espalhados pelo mundo.
- *Lumiere*, Microsoft Research, 1998 – O projeto pretende criar um sistema que possa automaticamente e inteligentemente interagir com outros sistemas, antecipando os objetivos e necessidades dos usuários.

Considerações Finais

- Redes *bayesianas* constituem uma forma natural para representação de informações condicionalmente independentes.
- Boa solução a problemas onde conclusões não podem ser obtidas apenas do domínio do problema.
- Inferências sobre redes *bayesinas*.
  - Podem ser executadas em tempo linear.
  - *NP-hard* para maioria dos casos.
  - Aplicação de técnicas.

Referências

- Charniak, Eugene. "Bayesians Networks without Tears". *IA Magazine*, 1991.
- Darwiche, Adan & Huang, Cecil. "Inference in Belief Networks: A procedural guide". International Journal of Approximate Reasoning, 1994.
- Jensen, V. Finn. "Bayesian Networks and Decision Graphs". Springer-Verlag,2001.
- Murphyk, P. Kevin. "A Brief Introduction to Graphical Models and Bayesian Networks". http://www.cs.berkeley.edu/~murphyk/Bayes/bayes.html
- Niedermayer, Daryle. "An Introduction to Bayesian Networks and their Contemporary Applications".1998.
- Russel, J. Stuart & Norvig, Peter. "Artifical Intelligence: A modern Approach".Prentice Hall.