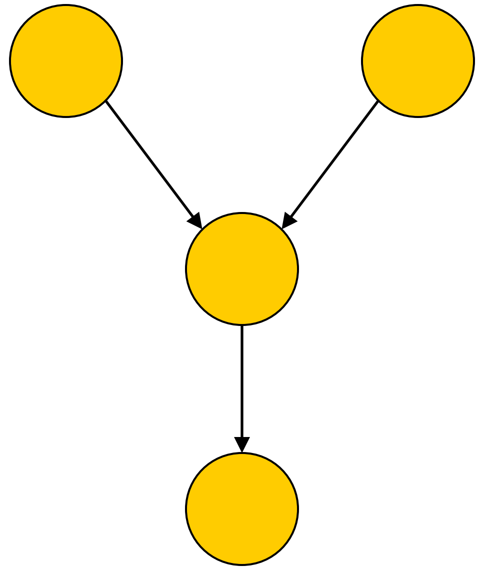


Redes Bayesianas
Material do Tutorial on Bayesian Networks

Weng-Keen Wong
School of Electrical Engineering and Computer Science
Oregon State University

Material Complementar extraído dos slides dos

Prof. Dr. Jaime Simao Sichman e da Prof. Dra. Anna Helena Reali Costa

Raciocinio Probabilistico - Livro: IA – J. Pearl e Norvig

Introduction

Suppose you are trying to determine if a patient has pneumonia. You observe the following symptoms:

- The patient has a cough
- The patient has a fever
- The patient has difficulty breathing

Introduction

You would like to determine how likely the patient has pneumonia given that the patient has a cough, a fever, and difficulty breathing

We are not 100% certain that the patient has pneumonia because of these symptoms. We are dealing with uncertainty!

Introduction

Now suppose you order a chest x-ray and the results are positive.

Your belief that that the patient has pneumonia is now much higher.

Introduction

- In the previous slides, what you observed affected your belief that the patient has pneumonia
- This is called reasoning with uncertainty
- Wouldn't it be nice if we had some methodology for reasoning with uncertainty? Why in fact, we do...

Bayesian Networks

- Bayesian networks help us reason with uncertainty
- In the opinion of many AI researchers, Bayesian networks are the most significant contribution in AI in the last 10 years
- They are used in many applications eg.:
  - Spam filtering / Text mining
  - Speech recognition
  - Robotics
  - Diagnostic systems
  - Syndromic surveillance

Outline

1. Introduction
2. Probability Primer
3. Bayesian networks
4. Bayesian networks in syndromic surveillance

Probability Primer: Random Variables

- A random variable is the basic element of probability
- Refers to an event and there is some degree of uncertainty as to the outcome of the event
- For example, the random variable *A* could be the event of getting a heads on a coin flip

Boolean Random Variables

- We deal with the simplest type of random variables – Boolean ones
- Take the values *true* or *false*
- Think of the event as occurring or not occurring
- Examples (Let A be a Boolean random variable):

  A = Getting heads on a coin flip

  A = It will rain today

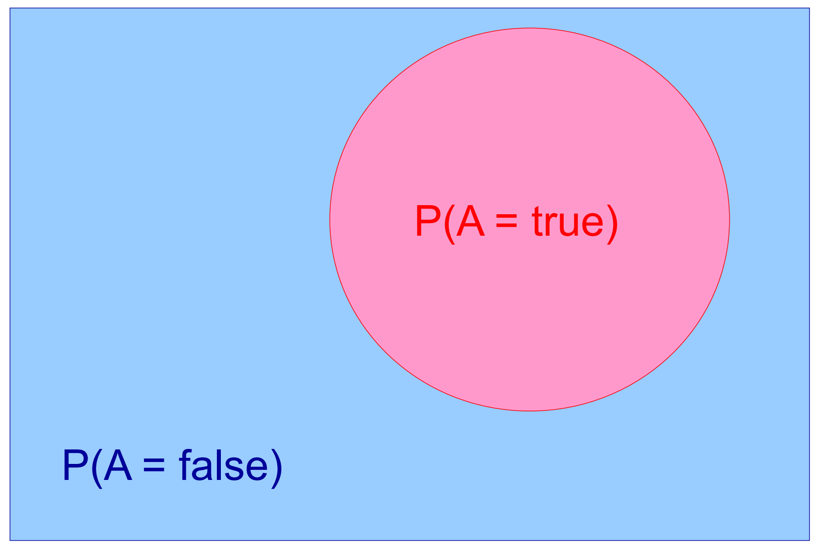
  A = There is a typo in these slides

Probabilities

We will write *P(A = true)* to mean the probability that *A = true*.

What is probability? It is the relative frequency with which an outcome would be obtained if the process were repeated a large number of times under similar conditions*

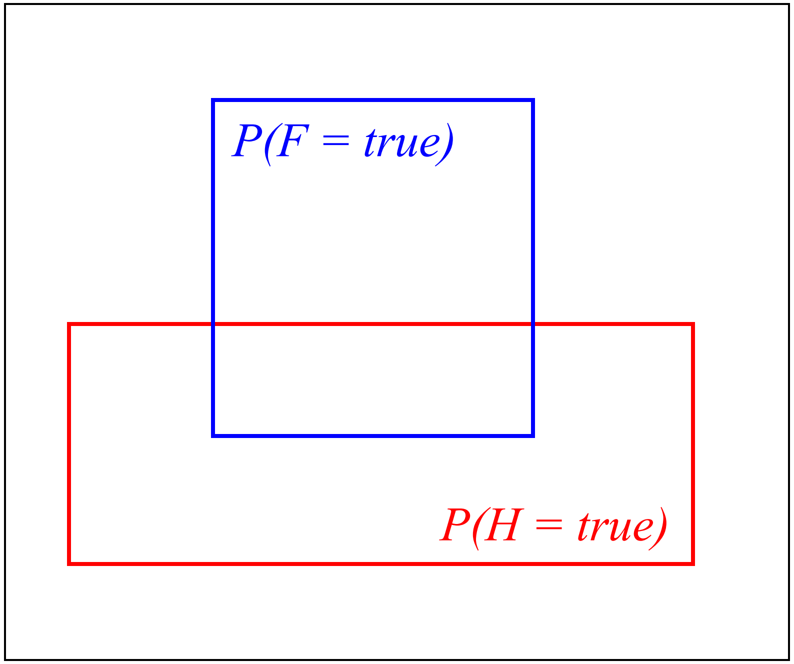
The sum of the red and blue areas is 1



*Ahem…there's also the Bayesian definition which says probability is your degree of belief in an outcome

Conditional Probability

- $P(A = true \mid B = true)$ = Out of all the outcomes in which B is true, how many also have A equal to true
- Read this as: "Probability of A conditioned on B" or "Probability of A given B"



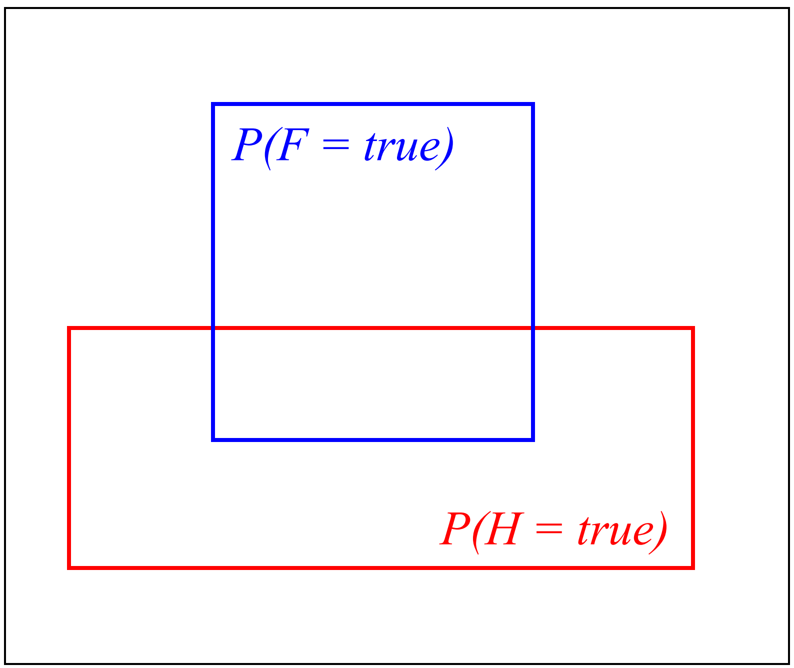
H = "Have a headache"
F = "Coming down with Flu"

$P(H = true) = 1/10$
$P(F = true) = 1/40$
$P(H = true \mid F = true) = 1/2$

"Headaches are rare and flu is rarer, but if you're coming down with flu there's a 50-50 chance you'll have a headache."

The Joint Probability Distribution

- We will write *P(A = true, B = true)* to mean "the probability of *A = true* **and** *B = true*"
- Notice that:



$$P(H{=}true \mid F{=}true)$$

$$= \frac{\text{Area of "H and F" region}}{\text{Area of "F" region}}$$

$$= \frac{P(H = true, F = true)}{P(F = true)}$$

In general, $P(X|Y)=P(X,Y)/P(Y)$

The Joint Probability Distribution

- Joint probabilities can be between any number of variables

  eg. $P(A = true, B = true, C = true)$
- For each combination of variables, we need to say how probable that combination is
- The probabilities of these combinations need to sum to 1

A	**B**	**C**	**P(A,B,C)**
false	false	false	0.1
false	false	true	0.2
false	true	false	0.05
false	true	true	0.05
true	false	false	0.3
true	false	true	0.1
true	true	false	0.05
true	true	true	0.15

Sums to 1

The Joint Probability Distribution

A	**B**	**C**	**P(A,B,C)**
false	false	false	0.1
false	false	true	0.2
false	true	false	0.05
false	true	true	0.05
true	false	false	0.3
true	false	true	0.1
true	true	false	0.05
true	true	true	0.15

- Once you have the joint probability distribution, you can calculate any probability involving A, B, and C
- Note: May need to use marginalization and Bayes rule, (both of which are not discussed in these slides)

Examples of things you can compute:

- $P(A=true)$ = sum of $P(A,B,C)$ in rows with $A=true$
- $P(A=true, B = true \mid C=true) =$

  $P(A = true, B = true, C = true) / P(C = true)$

The Problem with the Joint Distribution

- Lots of entries in the table to fill up!
- For *k* Boolean random variables, you need a table of size 2^k
- How do we use fewer numbers? Need the concept of independence

A	**B**	**C**	**P(A,B,C)**
false	false	false	0.1
false	false	true	0.2
false	true	false	0.05
false	true	true	0.05
true	false	false	0.3
true	false	true	0.1
true	true	false	0.05
true	true	true	0.15

Independence

Variables A and B are independent if any of the following hold:

- $P(A,B) = P(A)\, P(B)$
- $P(A \mid B) = P(A)$
- $P(B \mid A) = P(B)$

This says that knowing the outcome of A does not tell me anything new about the outcome of B.

Independence

How is independence useful?

- Suppose you have n coin flips and you want to calculate the joint distribution $\boldsymbol{P}(C_1, \ldots, C_n)$
- If the coin flips are not independent, you need 2^n values in the table
- If the coin flips are independent, then

$$P(C_1, \ldots, C_n) = \prod_{i=1}^{n} P(C_i)$$

Each P(C_i) table has 2 entries and there are n of them for a total of $2n$ values

Conditional Independence

Variables A and B are conditionally independent given C if any of the following hold:

- $P(A, B \mid C) = P(A \mid C)\, P(B \mid C)$
- $P(A \mid B, C) = P(A \mid C)$
- $P(B \mid A, C) = P(B \mid C)$

Knowing C tells me everything about B. I don't gain anything by knowing A (either because A doesn't influence B or because knowing C provides all the information knowing A would give)

Outline

1. Introduction
2. Probability Primer
3. Bayesian networks
4. Bayesian networks in syndromic surveillance

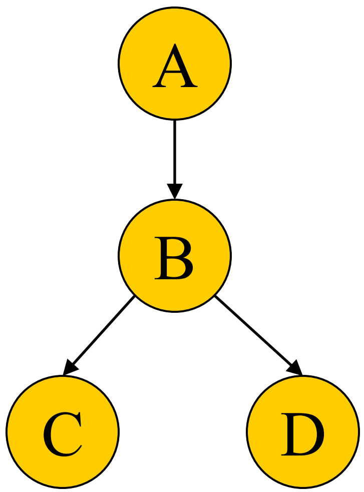
Rede Bayesiana

- Redes Bayesianas (RB)
- As **redes bayesianas** consistem em uma notação gráfica para representar independência condicional e assim especificar de forma concisa distribuições conjuntas totais
- Notação
- – Uma rede bayesiana é um grafo acíclico orientado
- – Existe um nó para cada variável aleatória
- – Um arco orientado (seta) de um nó Y a um nó X e lido como Y é pai de X;
- Pais(X)={conjunto de todos nos pais de }
- – Cada nó Xi tem uma distribuição de probabilidade condicional associada – P(Xi|Pais(Xi)). Esta distribuição quantifica o efeito das variáveis aleatórias representadas nos nós pais sobre a variável aleatória representada no nó Xi.

A Bayesian Network

A Bayesian network is made up of:

1. A Directed Acyclic Graph



2. A set of tables for each node in the graph

A	P(A)
false	0.6
true	0.4

A	B	P(B|A)
false	false	0.01
false	true	0.99
true	false	0.7
true	true	0.3

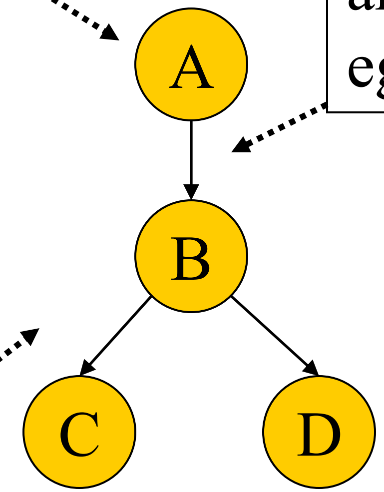
B	D	P(D|B)
false	false	0.02
false	true	0.98
true	false	0.05
true	true	0.95

B	C	P(C|B)
false	false	0.4
false	true	0.6
true	false	0.9
true	true	0.1

A Directed Acyclic Graph

Each node in the graph is a random variable

A node X is a parent of another node Y if there is an arrow from node X to node Y eg. A is a parent of B



Informally, an arrow from node X to node Y means X has a direct influence on Y

A Set of Tables for Each Node

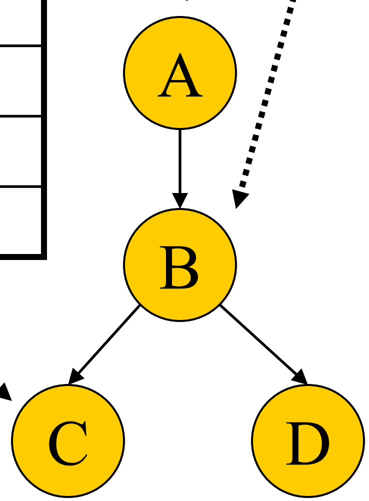
A	P(A)
false	0.6
true	0.4

A	B	P(B|A)
false	false	0.01
false	true	0.99
true	false	0.7
true	true	0.3

Each node X_i has a conditional probability distribution P(X_i | Parents(X_i)) that quantifies the effect of the parents on the node

The parameters are the probabilities in these conditional probability tables (CPTs)

B	C	P(C|B)
false	false	0.4
false	true	0.6
true	false	0.9
true	true	0.1



B	D	P(D|B)
false	false	0.02
false	true	0.98
true	false	0.05
true	true	0.95

A Set of Tables for Each Node

Conditional Probability
Distribution for C given B

B	C	P(C|B)
false	false	0.4
false	true	0.6
true	false	0.9
true	true	0.1

For a given combination of values of the parents (B in this example), the entries for P(C=true | B) and P(C=false | B) must add up to 1
eg. P(C=true | B=false) + P(C=false |B=false)=1

If you have a Boolean variable with k Boolean parents, this table has 2^{k+1} probabilities (but only 2^k need to be stored)

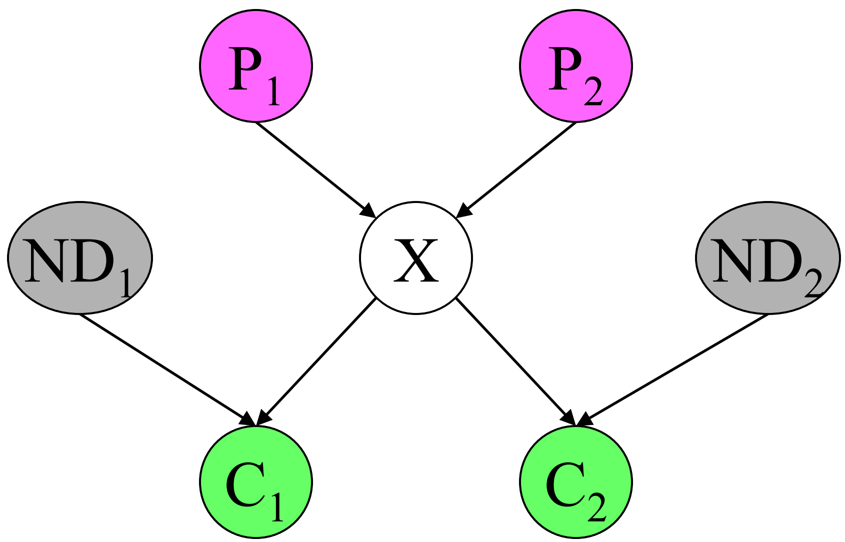
Bayesian Networks

Two important properties:

1. Encodes the conditional independence relationships between the variables in the graph structure
2. Is a compact representation of the joint probability distribution over the variables

Conditional Independence

The Markov condition: given its parents (P_1, P_2), a node (X) is conditionally independent of its non-descendants (ND_1, ND_2)



The Joint Probability Distribution

Due to the Markov condition, we can compute the joint probability distribution over all the variables $X_1, \ldots, X_n$ in the Bayesian net using the formula:

$$P(X_1 = x_1, \ldots, X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i \mid Parents(X_i))$$

Where Parents(X_i) means the values of the Parents of the node X_i with respect to the graph

Using a Bayesian Network Example

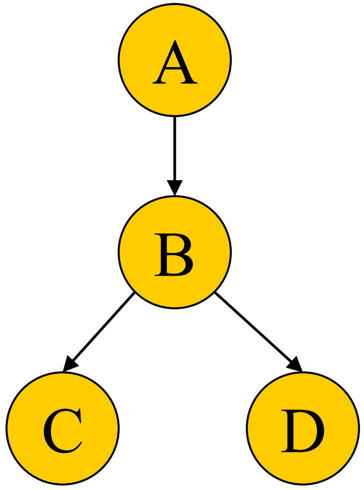
Using the network in the example, suppose you want to calculate:

$P(A = true, B = true, C = true, D = true)$

$= P(A = true) * P(B = true \mid A = true) *$

$P(C = true \mid B = true)\ P(D = true \mid B = true)$

$= (0.4)*(0.3)*(0.1)*(0.95)$



Using a Bayesian Network Example

Using the network in the example, suppose you want to calculate:

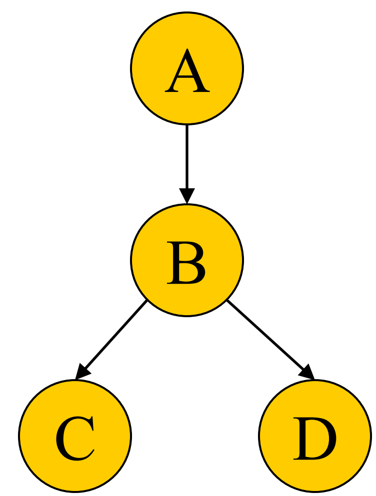
$P(A = \text{true}, B = \text{true}, C = \text{true}, D = \text{true})$

$= P(A = \text{true}) * P(B = \text{true} \mid A = \text{true}) *$
$\quad P(C = \text{true} \mid B = \text{true})\, P(D = \text{true} \mid B = \text{true})$

This is from the graph structure

$= (0.4)*(0.3)*(0.1)*(0.95)$

These numbers are from the conditional probability tables



Inference

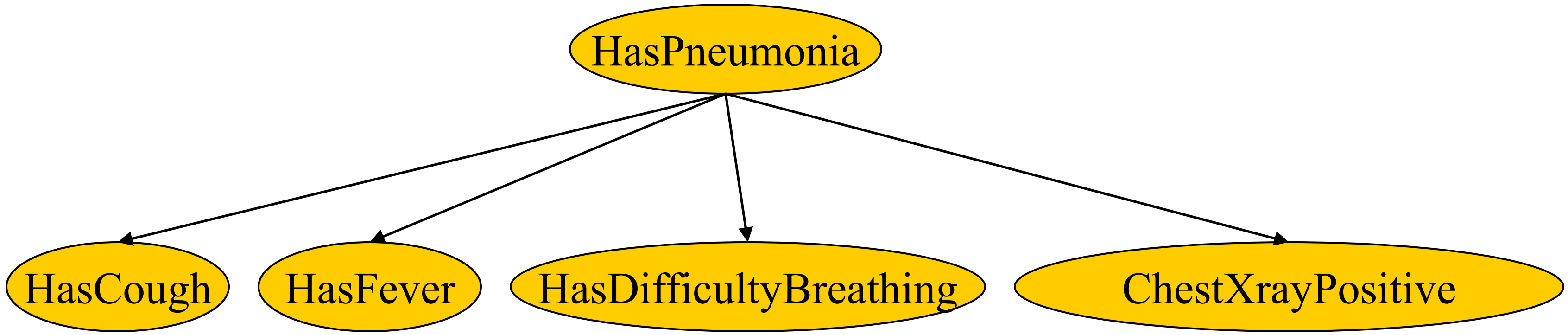
- Using a Bayesian network to compute probabilities is called inference
- In general, inference involves queries of the form:

$P(X \mid E)$

E = The evidence variable(s)

X = The query variable(s)

Inference



- An example of a query would be:
  P(*HasPneumonia* = *true* | *HasFever* = *true*, *HasCough* = *true*)
- Note: Even though *HasDifficultyBreathing* and *ChestXrayPositive* are in the Bayesian network, they are not given values in the query (ie. they do not appear either as query variables or evidence variables)
- They are treated as unobserved variables

The Bad News

- Exact inference is feasible in small to medium-sized networks
- Exact inference in large networks takes a very long time
- We resort to approximate inference techniques which are much faster and give pretty good results

How is the Bayesian network created?

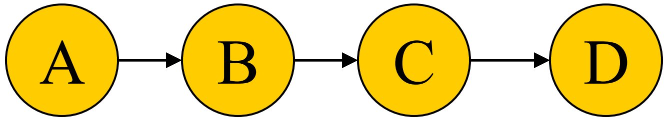
1. Get an expert to design it
   - Expert must determine the structure of the Bayesian network
     - This is best done by modeling direct causes of a variable as its parents
   - Expert must determine the values of the CPT entries
     - These values could come from the expert's informed opinion
     - Or an external source eg. census information
     - Or they are estimated from data
     - Or a combination of the above

2. Learn it from data
   - This is a much better option but it usually requires a large amount of data
   - This is where Bayesian statistics comes in!

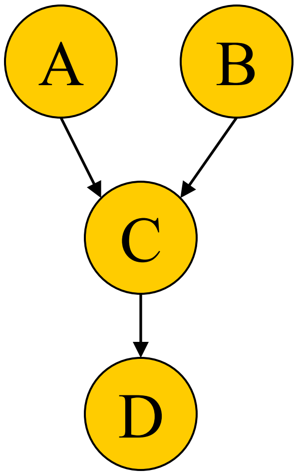
Learning Bayesian Networks from Data

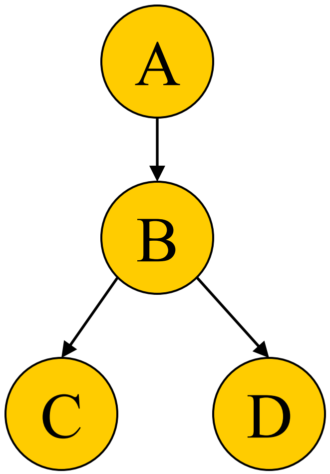
Given a data set, can you learn what a Bayesian network with variables A, B, C and D would look like?

A	B	C	D
true	false	false	true
true	false	true	false
true	false	false	true
false	true	false	false
false	true	false	true
false	true	false	false
false	true	false	false
:	:	:	:



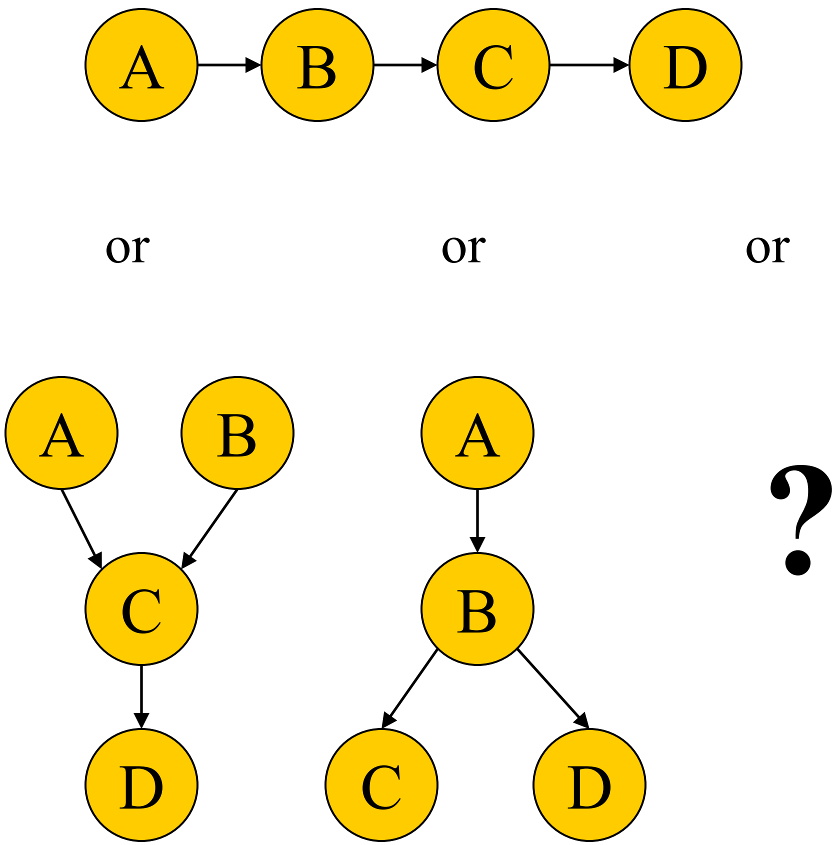
or or or





?

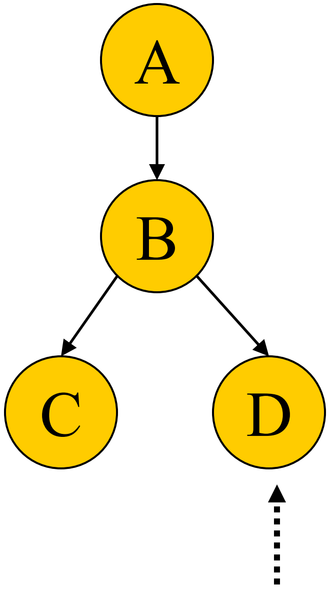
Learning Bayesian Networks from Data



- Each possible structure contains information about the conditional independence relationships between A, B, C and D
- We would like a structure that contains conditional independence relationships that are supported by the data
- Note that we also need to learn the values in the CPTs from data

Learning Bayesian Networks from Data

How does Bayesian statistics help?



B	D	P(D|B)
false	false	0.02
false	true	0.98
true	false	0.05
true	true	0.95

1. I might have a prior belief about what the structure should look like.

2. I might have a prior belief about what the values in the CPTs should be.

These beliefs get updated as I see more data

Learning Bayesian Networks from Data

- We won't have enough time to describe how we actually learn Bayesian networks from data
- If you are interested, here are some references:
  - Gregory F. Cooper and Edward Herskovits. A Bayesian Method for the Induction of Probabilistic Networks from Data. Machine Learning, 9:309-347, 1992.
  - David Heckerman. A Tutorial on Learning Bayesian Networks. Technical Report MSR-TR-95-06, Microsoft Research. 1995. (Available online)

Outline

1. Introduction
2. Probability Primer
3. Bayesian networks
4. Bayesian networks in syndromic surveillance

Raciocínio Probabilístico
Mat. Complementar

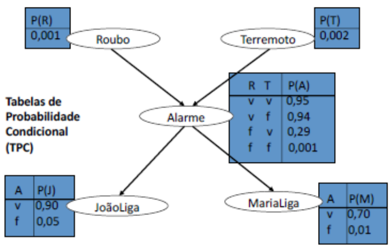
- Material Extraído dos Slides dos
- Prof. Dr. Jaime Simao Sichman e da Prof. Dra. Anna Helena Reali Costa
- Raciocinio Probabilistico
- Livro: IA – S. Russel e P. Norvig (Cap. 14)

Raciocínio Probabilístico

RB – exemplo 2

- Outro Exemplo: (by J. Pearl)

Tabelas de Probabilidade Condicional (TPC)



33

RB – Tabelas Prob. Condicional

- Comentários
  - Uma TPC para variáveis aleatórias booleans X_i com **k** pais booleanos tem 2^k linhas para as combinações de valores dos pais
  - Cada linha requer um número **p** para X_i=verdadeiro (o número para X_i=falso é 1 – **p**)
  - Se cada variável tem no máximo **k** pais, a rede completa requer **O(n*2^k)** números
  - Isto é, cresce linearmente com **n**, versus **O(2^n)** para a distribuição conjunta total
  - No exemplo do alarme:

    1+1+4+2+2=10 números

    (vs., 2^5 =32 para a distribuição conjunta total)

34

RB – Semântica

- Duas maneiras de ver
  - **Semântica Numérica**
    - RB como uma representação da distribuição probabilidades conjunta total
  - **Semântica Topológica**
    - RB como uma codificação de uma coleção de declarações sobre independência condicional
- As duas visões são equivalentes, mas:
  - A primeira é útil na compreensão de como construir redes
  - A segunda é útil no projeto de procedimentos de inferência aproximados

35

RB – Semântica Numérica

Representação da Distribuição Conjunta Total

$$p(x_1,\dots,x_n) = \prod_{i=1}^{n} p(x_i \mid pais(X_i))$$

$$onde:$$

$$p(x_1,\dots,x_n) \equiv p(X_1 = x_1 \wedge \dots \wedge X_n = x_n)$$

$$pais(X_i) \Rightarrow \text{valores específicos das variáveis em } Pais(X_i)$$

- Exemplo
  - Probabilidade de que o alarme tenha soado, mas não tenha ocorrido nenhum roubo nem terremoto, e que tanto João quanto Maria tenham ligado???

p(j ∧m ∧a ∧¬r ∧¬t)
= p(j|a) * p(m|a) * p(a| ¬r ∧¬t) * p(¬r) * p(¬t)
= 0,9 * 0,7 * 0,001 * (1 – 0,001) * (1 – 0,002)
≈ 0,000628

36

Raciocínio Probabilístico

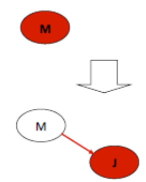
RB – Método de Construção

- Requisito
  - Garantir semântica numérica
- Método
  1. Escolha uma ordem para as variáveis $X_1,\dots,X_n$
  2. Para i=1 até n faça
     1. adicione X_i a rede
     2. Defina $Pais(X_i) \subseteq \{X_1,\dots,X_{i-1}\}$ de tal maneira que: $\mathbf{P}(X_i|Pais(X_i)) = \mathbf{P}(X_i \mid X_1,\dots,X_{i-1})$
- Primeiro, um exemplo de utilização ...
- Depois, por que o método funciona ...

37

RB – Método de Construção Exemplo do Alarme

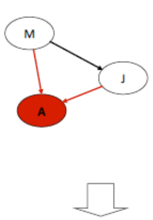
- Variáveis
  - A(larme), J(oãoLiga) , M(ariaLiga), R(oubo), T(erremoto)
- Vamos supor que escolhemos a ordem
  - M,J, A, R, T
- Processo
  - Adiciona-se M:
    - não há pais
  - Adiciona-se J:
    - P(J | M) = P(J) ?Não



38

RB – Método de Construção Exemplo do Alarme

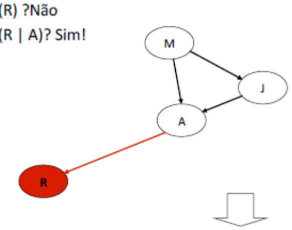
- Processo (cont.)
  - Adiciona-se A:
    - P(A | J,M) = P(A) ?Não
    - P(A | J,M) = P(A|J)? Não
    - P(A | J,M) = P(A|M)? Não



39

RB – Método de Construção Exemplo do Alarme

- Processo (cont.)
  - Adiciona-se R:
    - P(R | A,J,M) = P(R) ?Não
    - P(R | A,J,M) = P(R | A)? Sim!

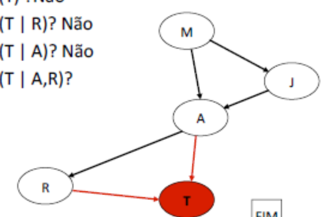


40

Raciocínio Probabilístico

RB – Método de Construção
Exemplo do Alarme

- Processo (cont.)
  - Adiciona-se T:
    - P(T |R, A,J,M) = P(T) ?Não
    - P(T |R, A,J,M) = P(T | R)? Não
    - P(T |R, A,J,M) = P(T | A)? Não
    - P(T |R, A,J,M) = P(T | A,R)?
      - SIM!



FIM

41

RB – Método de Construção
Exemplo do Alarme

- Considerações
  - Como escolher a ordem "correta" das variáveis?
    - A ordem mais apropriada consiste em adicionar primeiro variáveis de "causa"; variáveis de efeito devem ser deixadas por último
    - Ex.:
      - Variáveis de causa = R, T, A
      - Variáveis de efeito = A, M, J
  - Se tentarmos construir na ordem inversa (efeito para causa, como foi feito no exemplo) acabaremos sendo forçados a especificar dependências adicionais entre causas que de outra forma seriam independentes (e, com freqüência, também entre efeitos que ocorrem separadamente)
  - Uma ordenação causal leva a uma quantidade menor de números, e os números freqüentemente são mais fáceis de avaliar

42

Raciocínio Probabilístico

RB – Método de Construção
Por que funciona?

$$p(x_1,\ldots,x_n) = \boxed{\prod_{i=1}^{n} p(x_i \mid pais(X_i))}$$

Requisito:
Semântica Numérica

pela Regra do Produto

$\Downarrow$

$$\begin{aligned} p(x_1,\ldots,x_n) &= p(x_n \mid x_{n-1},\ldots,x_1) * p(x_{n-1},\ldots,x_1) \\ &= p(x_n \mid x_{n-1},\ldots,x_1) * p(x_{n-1} \mid x_{n-2},\ldots,x_1) * p(x_{n-2},\ldots,x_1) \\ &\ldots \\ &= p(x_n \mid x_{n-1},\ldots,x_1) * \ldots * p(x_2 \mid x_1) * p(x_1) \\ &= \boxed{\prod_{i=1}^{n} p(x_i \mid x_{i-1},\ldots,x_1)} \end{aligned}$$

43

RB – Método de Construção
Por que funciona?

$$p(x_1,\ldots,x_n) = \boxed{\prod_{i=1}^{n} p(x_i \mid pais(X_i))}$$

$\Updownarrow$ *comparando*

$$= \boxed{\prod_{i=1}^{n} p(x_i \mid x_{i-1},\ldots,x_1)}$$

$\Downarrow$

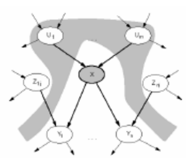
$$\left.\begin{aligned} &\mathbf{P}(X_i / X_{i-1},\ldots,X_1) = \mathbf{P}(X_i \mid Pais(X_i)) \\ &desde\ que\ Pais(X_i) \subseteq \{X_{i-1},\ldots,X_1\} \end{aligned}\right\}$$

Passo 2.2 do método de Construção

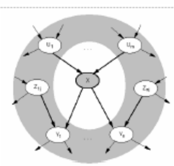
44

RB – Semântica Topológica

- Duas formulações equivalentes:
  - Um nó é condicionalmente independente de seus não-descendentes, dados seus pais



  - Ou, um nó é condicionalmente independente de todos os outros, dados seus pais, filhos e pais de seus filhos (isto é, dada a sua Cobertura de Markov)



45

RB – Semântica

- TEOREMA
  - A semântica numérica e a semântica topológica são equivalentes

46

Raciocínio Probabilístico

RB – Inferência

- Idéia Geral, terminologia e notação
  - Calcular distribuição

$$\mathbf{P}(X|\mathbf{e})$$

Onde:
X – variável de consulta
e – evidência observada

$$\mathbf{e} \equiv (E_1=e_1 \wedge \ldots \wedge E_n=e_n)$$

$\mathbf{E}=\{E_1,\ldots,E_n\}$ – variáveis de evidência
$\mathbf{Y}=\{Y_i;\ Y_i \neq X \ \wedge \ Y_i \notin \mathbf{E}\}$ – variáveis ocultas

$$\mathbf{y} \equiv (Y_1=y_1 \wedge \ldots \wedge Y_m=y_m)$$

$\mathbf{X} = \{X\} \cup \mathbf{E} \cup \mathbf{Y}$ – variáveis da RB

47

RB – Inferência

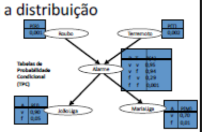
- Exemplo
  - **P**(Roubo | JoãoLiga=sim ∧ MariaLiga=sim) = ?
  - X ≡ Roubo
  - **e** ≡ JoãoLiga=sim ∧ MariaLiga=sim
  - **E** = {JoãoLiga,MariaLiga} - var de evidência
  - **Y** = {Alarme,Terremoto}- var ocultas
  - No que segue
    - V ≡ **Variável**
    - v ≡ **valor de V; por default V=sim, no caso booleano**

48

Raciocínio Probabilístico

RB – Inferência por Enumeração

- Usa o fato de que uma RB representa a distribuição conjunta total



$$\mathbf{P}(R \mid j,m) = \frac{\mathbf{P}(R,j,m)}{p(j,m)}$$

$$= \alpha * \mathbf{P}(R,j,m) \quad \text{onde} \quad \alpha = \frac{1}{p(j,m)}$$

$$= \alpha * \sum_t \sum_a \mathbf{P}(R,t,a,j,m)$$

usando semântica numérica

R=sim, ...

$\Downarrow$

$$p(r \mid j,m)$$

$$= \alpha * \sum_t \sum_a p(r) * p(t) * p(a \mid r,t) * p(j \mid a) * p(m \mid a)$$

$$= \alpha * p(r) * \sum_t p(t) * \sum_a p(a \mid r,t) * p(j \mid a) * p(m \mid a)$$

49

RB – Inferência por Enumeração

- Método Geral:

$$\mathbf{P}(X \mid \mathbf{e})$$

$$= \alpha * \mathbf{P}(X,\mathbf{e}) \quad \text{onde} \quad \alpha = \frac{1}{p(\mathbf{e})}$$

$$= \alpha * \sum_{\mathbf{y}} \mathbf{P}(X,\mathbf{e},\mathbf{y})$$

Aplica semântica numérica para cada valor de X

onde

$\mathbf{e} \equiv E_1 = e_1 \wedge \ldots \wedge E_n = e_n$ Evidência observada

$\mathbf{y} \equiv Y_1 = y_1 \wedge \ldots \wedge Y_m = y_m$ Combinação possível de valores das Variáveis ocultas

50

RB – Inferência por Enumeração – Algoritmo

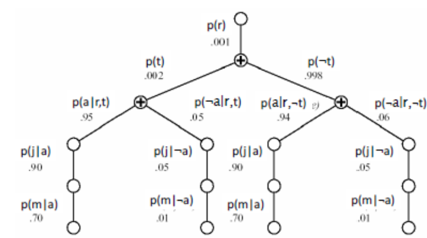
Função ASK-ENUMERAÇÃO(X,**e**,rb) **retorna** distribuição sobre X

Q(X) ← uma distribuição sobre X, vazia
para cada valor x_i de X **faça**
estender **e** com valor x_i para X
Q(x_i) ← ENUMERAR-TODOS(VARS[rb], **e**)
retornar NORMALIZAR(Q(X))

Função ENUMERAR-TODOS(vars, **e**) **retorna** um número real
se VAZIO?(vars)
então retornar 1,0
Y ← PRIMEIRO(vars)
se Y tem valor y em **e então**
retornar p(y | pais(Y)) * ENUMERAR-TODOS(RESTO(vars),**e**)
senão
retornar $\sum_y$ * p(y | pais(Y)) * ENUMERAR-TODOS(RESTO(vars),$\mathbf{e}_y$)
onde $\mathbf{e}_y$ é **e** estendido com Y=y

51

RB – Inferência por Enumeração – Computação



52

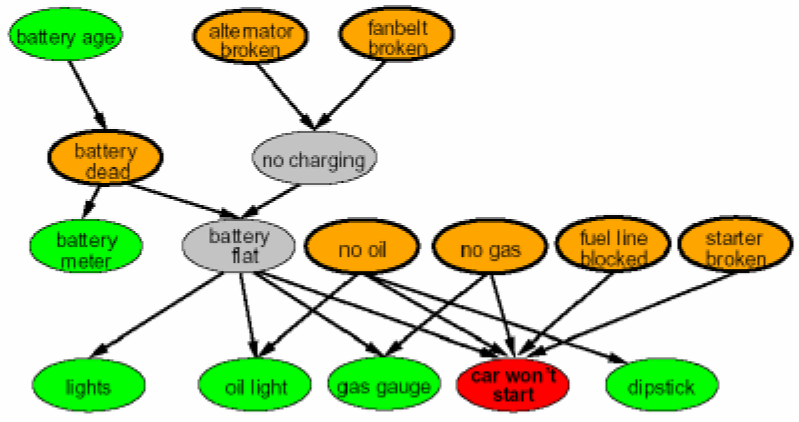
RB – Inferência

- Considerações
  - Algoritmo ASK-ENUMERAÇÃO
    - Complexidade de Espaço $O(n)$
    - Complexidade de Tempo $O(2^n)$
  - O algoritmo pode ser melhorado eliminando-se cálculos repetidos
    - Programação dinâmica
  - Pode-se mostrar que para redes **unicamente conexas** ou **poliárvores** (redes onde só existe no máximo um caminho orientado entre dois nós)
    - A complexidade de tempo e de espaço da inferência é $O(n)$
  - No caso de **redes multiconexas**
    - Inferência problema **NP-hard**
  - Na prática, no entanto, existem algoritmos eficientes para inferência aproximada

Considera

- As redes bayesianas fornecem u
  - Relacionamentos de independên
  - Uma distribuição conjunta total; das entradas correspondentes na RB com freqüência é exponencial total
- A inferência para redes bayesian um tempo linear no tamanho da intratável.
- Pontos não discutidos
  - Como construir as TPC dos nodos
  - Variáveis aleatórias contínuas
  - Métodos de inferência aproximad

RB – Apêndice
Diagnóstico de Carro



Evidência inicial : carro não dá partida! (em vermelho)
Verde – variáveis de teste (outras evidências)
Laranja – variáveis "broken, so fix it"
Cinza – variáveis escondidadas

Acknowledgements

- These slides were partly based on a tutorial by Andrew Moore
- Greg Cooper, John Levander, John Dowling, Denver Dash, Bill Hogan, Mike Wagner, and the rest of the RODS lab

References

Bayesian networks:

- “Bayesian networks without tears” by Eugene Charniak
- “Artificial Intelligence: A Modern Approach” by Stuart Russell and Peter Norvig
- “Learning Bayesian Networks” by Richard Neopolitan
- “Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference” by Judea Pearl

Other references:

- My webpage
  http://www.eecs.oregonstate.edu/~wong