Inteligencia Artificial

Inferencia Exacta en Redes Bayesianas

Jorge Luis Guevara Diaz

www.jorge.sistemasyservidores.com

Que veremos?

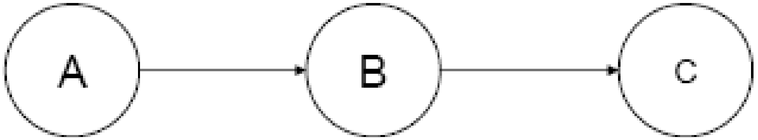
- Inferencia exacta por ennumeracion
- Inferencia exacta por eliminacion de variables
- Inferencia aproximada por simulación estocástica
- Inferencia aproximada por método de Monte Carlo y cadenas de Markov

Inferencia Básica

$P(b) = ?$

$$P(b) = \sum_a \mathrm{P}(\mathrm{a}, \mathrm{b}) = \sum_a \mathrm{P}(\mathrm{b} \mid \mathrm{a})\, \mathrm{P}(\mathrm{a})$$

Inferencia Básica

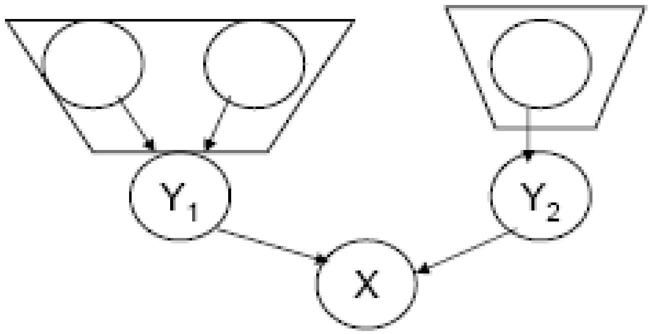


$$P(b) = \sum_a \mathrm{P(a, b)} = \sum_a \mathrm{P(b \mid a)\, P(a)}$$

$$P(c) = \sum_{a,b} P(c \mid b) P(b \mid a) P(a)$$

$$= \sum_{a,b} P(c \mid b) P(b)$$

Inferencia en árboles



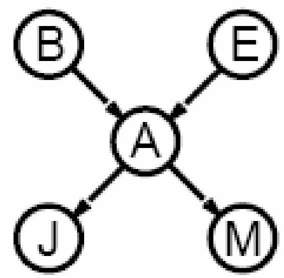
- P(X)=?

Inferencia

- Consultas simples $P(X_i | E=e)$
  - P(SinGas | Indicador=vacio, Luces=prendidas, Inicio=falso)
- Consultas conjuntivas
  - $P(X_i, X_j | E=e) = P(X_i | E=e) \, P(X_j | X_i, E=e)$

Inferencia por ennumeración

$$\begin{aligned}
&\mathbf{P}(B|j,m)\\
&= \mathbf{P}(B,j,m)/P(j,m)\\
&= \alpha\mathbf{P}(B,j,m)\\
&= \alpha \textstyle\sum_e \sum_a \mathbf{P}(B,e,a,j,m)
\end{aligned}$$



- Reescribir entradas de la distribución con completa usando entradas a la CPT

$$\begin{aligned}
&\mathbf{P}(B|j,m)\\
&= \alpha \textstyle\sum_e \sum_a \mathbf{P}(B)P(e)\mathbf{P}(a|B,e)P(j|a)P(m|a)\\
&= \alpha\mathbf{P}(B) \textstyle\sum_e P(e) \sum_a \mathbf{P}(a|B,e)P(j|a)P(m|a)
\end{aligned}$$

- Ennumeración Recursiva primero en profundidad, espacio O(n) Tiempo O(d^n)

Algoritmo de enumeración

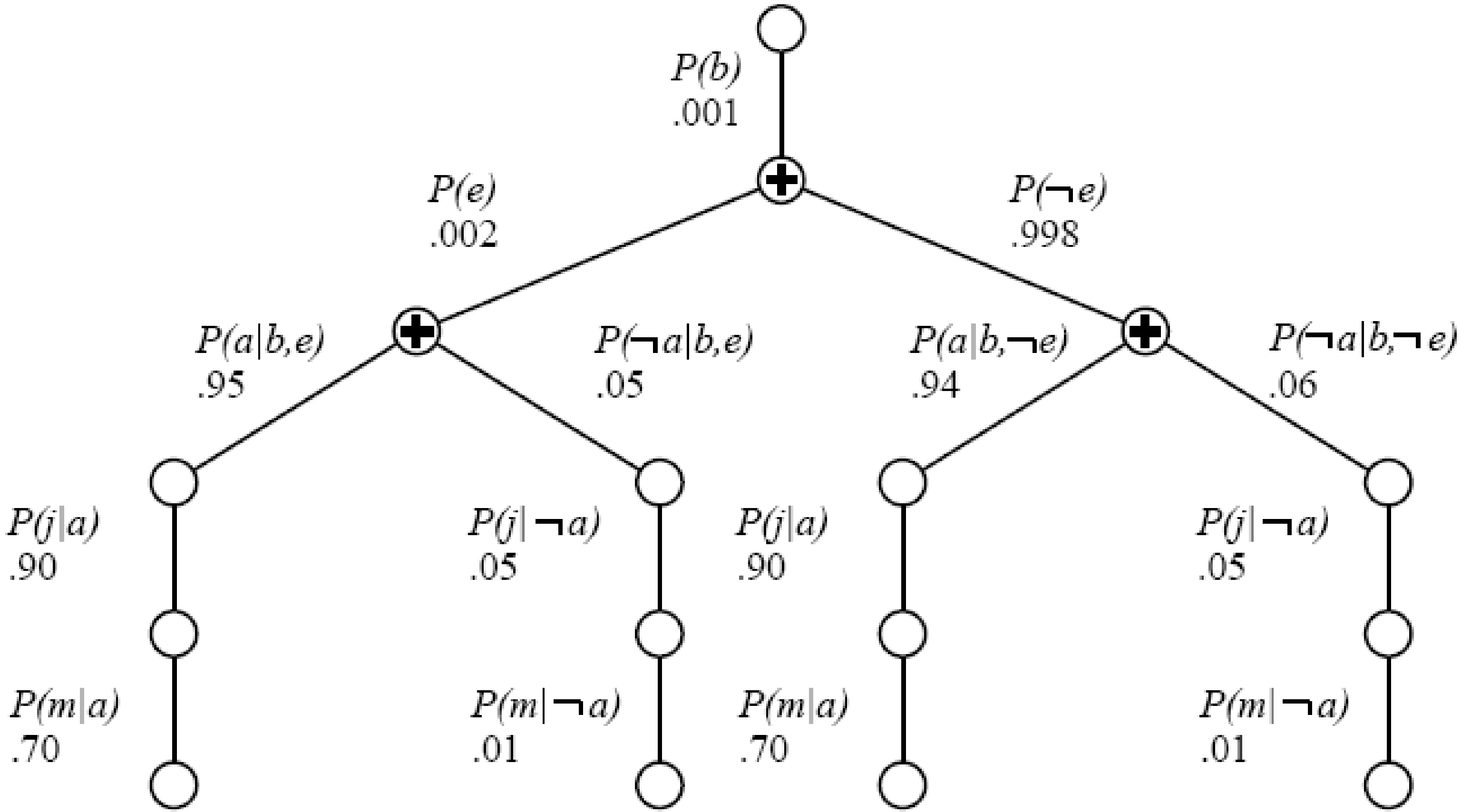
function ENUMERATION-ASK(X, $\mathbf{e}$, bn) **returns** a distribution over X
  inputs: X, the query variable
    $\mathbf{e}$, observed values for variables $\mathbf{E}$
    bn, a Bayesian network with variables $\{X\} \cup \mathbf{E} \cup \mathbf{Y}$

  $\mathbf{Q}(X) \leftarrow$ a distribution over X, initially empty
  for each value x_i of X **do**
    extend $\mathbf{e}$ with value x_i for X
    $\mathbf{Q}(x_i) \leftarrow$ ENUMERATE-ALL(VARS[bn], $\mathbf{e}$)
  return NORMALIZE($\mathbf{Q}(X)$)

---

function ENUMERATE-ALL($vars$, $\mathbf{e}$) **returns** a real number
  if EMPTY?($vars$) **then return** 1.0
  $Y \leftarrow$ FIRST($vars$)
  if Y has value y in $\mathbf{e}$
    then return $P(y \mid Pa(Y)) \times$ ENUMERATE-ALL(REST($vars$), $\mathbf{e}$)
    else return $\Sigma_y\ P(y \mid Pa(Y)) \times$ ENUMERATE-ALL(REST($vars$), $\mathbf{e}_y$)
      where $\mathbf{e}_y$ is $\mathbf{e}$ extended with $Y = y$

Arbol de ennumeración



Inferencia por ennumeración

- El algoritmo de ennumeración es ineficiente
  - Se repite el cálculo de P(j\a) P(m\a) para cada posible valor de e
  - Una mejora es el algoritmo de eliminación de variables

Inferencia por eliminación de variables

- Operaciones básicas:

  Pointwise product (PP) y Summing out (SO)

  - Pointwise product (PP)
    - Sean los factores f_1 y f_2

$$f_1(x_1, \ldots, x_j, y_1, \ldots, y_k) \times f_2(y_1, \ldots, y_k, z_1, \ldots, z_l)$$
$$= f(x_1, \ldots, x_j, y_1, \ldots, y_k, z_1, \ldots, z_l)$$

Ejemplo:

$$f_1(a,b) \times f_2(b,c) = f(a,b,c)$$

Inferencia por eliminación de variables

- Ejemplo:
  - Si

A	B	$\mathbf{f}_1(A,B)$
T	T	.3
T	F	.7
F	T	.9
F	F	.1

B	C	$\mathbf{f}_2(B,C)$
T	T	.2
T	F	.8
F	T	.6
F	F	.4

A	B	C	$\mathbf{f}_3(A,B,C)$
T	T	T	$.3 \times .2$
T	T	F	$.3 \times .8$
T	F	T	$.7 \times .6$
T	F	F	$.7 \times .4$
F	T	T	$.9 \times .2$
F	T	F	$.9 \times .8$
F	F	T	$.1 \times .6$
F	F	F	$.1 \times .4$

Inferencia por eliminación de variables

- Summing out (SO)
  - Mover cualquier valor constante fuera de la sumatoria y realizar la sumatoria interna del producto de factores

$$\sum_x f_1 \times \cdots \times f_k = f_1 \times \cdots \times f_i \sum_x f_{i+1} \times \cdots \times f_k = f_1 \times \cdots \times f_i \times f_{\bar{X}}$$

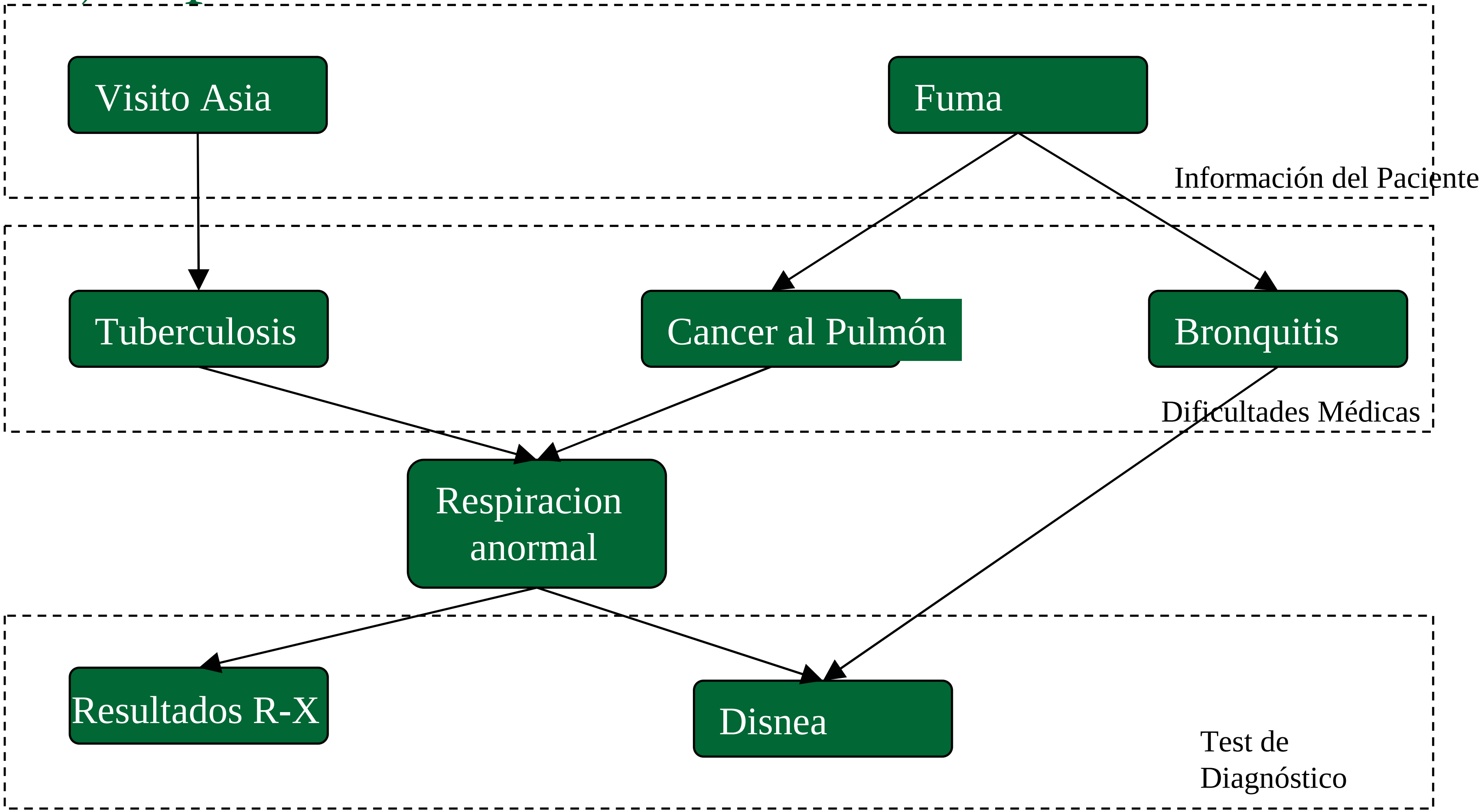
  - Donde $f_1 \ldots f_i$ no dependen de x

Inferencia por eliminación de variables

- Almacena los resultados (programación dinámica) intermedios, para evitar cálculos repetidos
- Redistribuye la sumatoria de las probabilidades

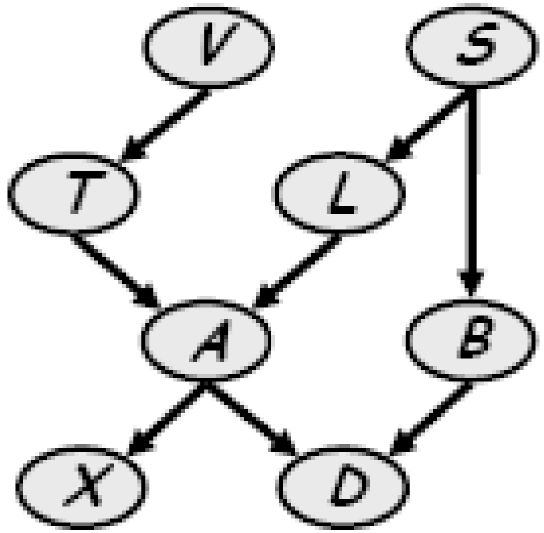
$$
\begin{aligned}
&\mathbf{P}(B|j,m) \\
&\quad = \alpha \underbrace{\mathbf{P}(B)}_{B} \textstyle\sum_e \underbrace{P(e)}_{E} \textstyle\sum_a \underbrace{\mathbf{P}(a|B,e)}_{A} \underbrace{P(j|a)}_{J} \underbrace{P(m|a)}_{M} \\
&\quad = \alpha \mathbf{P}(B) \textstyle\sum_e P(e) \textstyle\sum_a \mathbf{P}(a|B,e) P(j|a) f_M(a) \\
&\quad = \alpha \mathbf{P}(B) \textstyle\sum_e P(e) \textstyle\sum_a \mathbf{P}(a|B,e) f_J(a) f_M(a) \\
&\quad = \alpha \mathbf{P}(B) \textstyle\sum_e P(e) \textstyle\sum_a f_A(a,b,e) f_J(a) f_M(a) \\
&\quad = \alpha \mathbf{P}(B) \textstyle\sum_e P(e) f_{\bar{A}JM}(b,e) \text{ (sum out } A) \\
&\quad = \alpha \mathbf{P}(B) f_{\bar{E}\bar{A}JM}(b) \text{ (sum out } E) \\
&\quad = \alpha f_B(b) \times f_{\bar{E}\bar{A}JM}(b)
\end{aligned}
$$

Ejemplo:



Visito Asia
Fuma
Información del Paciente
Tuberculosis
Cancer al Pulmón
Bronquitis
Dificultades Médicas
Respiracion anormal
Resultados R-X
Disnea
Test de Diagnóstico

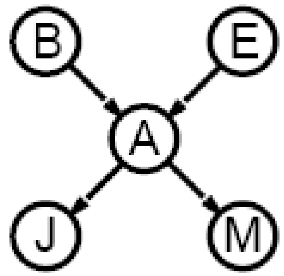
Ejemplo:



- Calcular P(d)
- Se necesita eliminar v,s,x,t,l,a,b

Variables Irrelevantes

- P(J\b)



$$P(J|b) = \alpha P(b) \sum_e P(e) \sum_a P(a|b,e) P(J|a) \sum_m P(m|a)$$

- M es irrelevante
- Teorema 1
  - Y es irrelevante a menos que pertenezca a los antecesores de X y E
  - Ejemplo:
  - X={J}, E={B}, Antecesores(X,E)={A, E}
  - M es irrelevante

Variables Irrelevantes

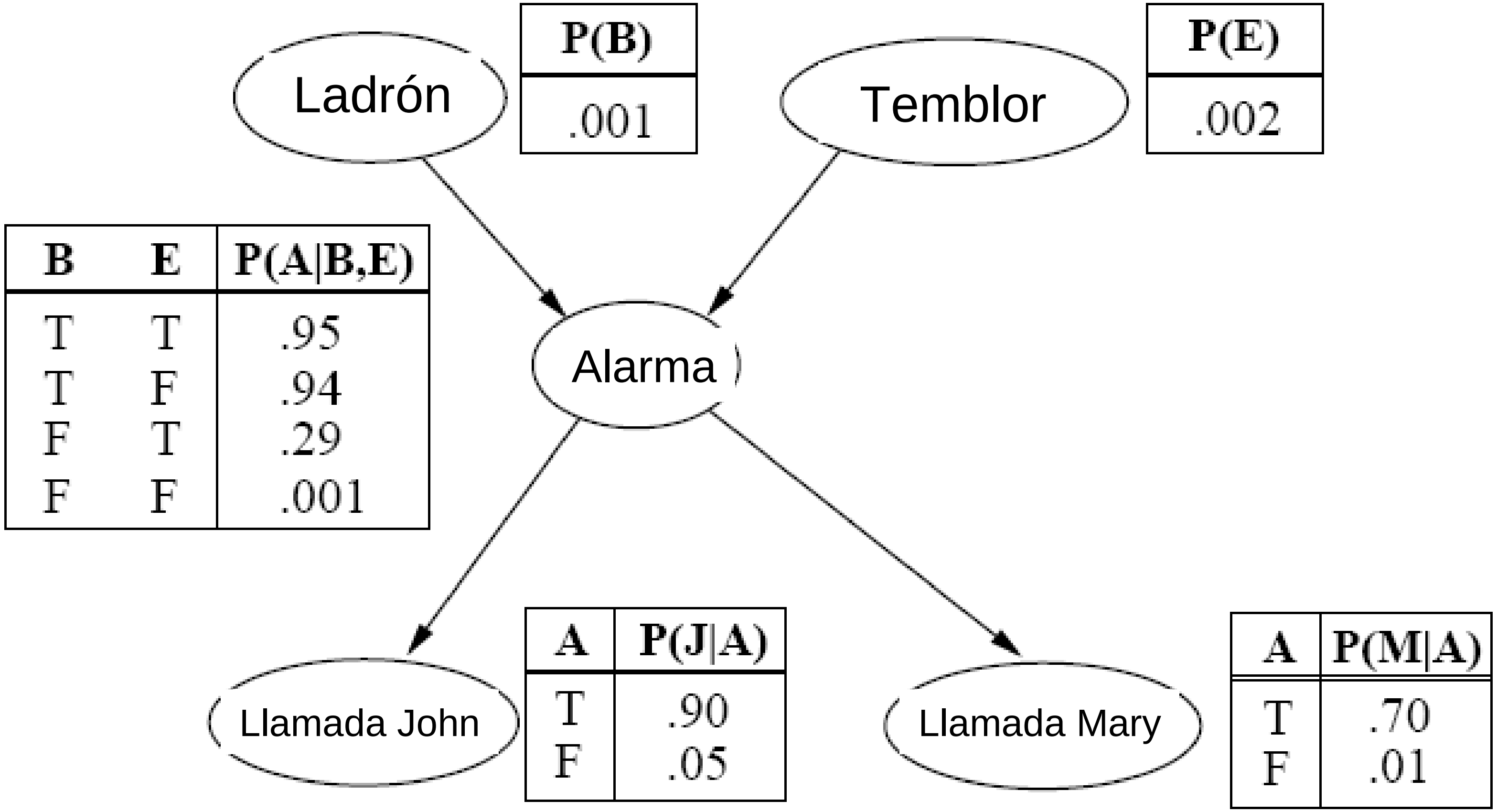
- Definicion:
  - Grafo Moral de una red Bayesiana: emparejar los padres y borrar las flechas
- Definicion:
  - A es m-separad de B por C, si y solo si está separado por C en el grado moral
- Teorema 2:
  - Y es irrelevante, si es m-separado de X por E

  - Ejemplo:
  - Para P(J\A=true) E, y B son irrelevantes

Ejercicio

- Calcular P(b / j,m) con el algoritmo de eliminacion de variables y validar la respuesta calculandola de manera convencional

- Cuantas operaciones aritméticas realiza el algoritmo de eliminacion de variables frente al de ennumeracion?



P(B)
.001

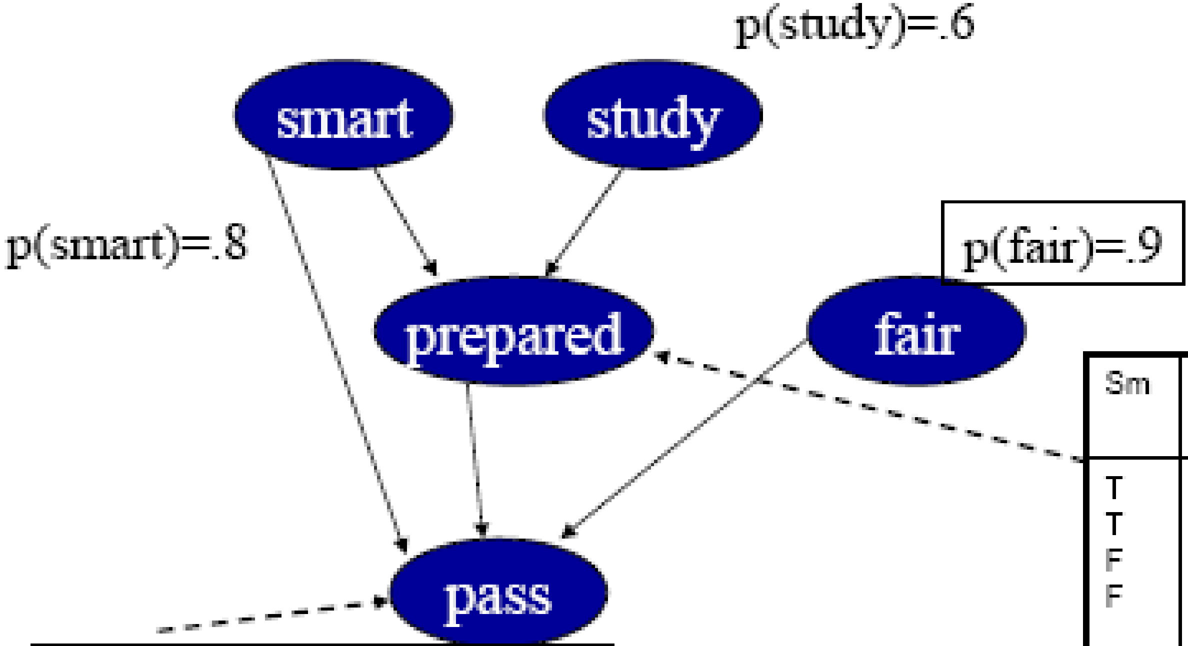
P(E)
.002

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

Ejercicio



Sm	St	P(Pr|...)
T	T	.9
T	F	.5
F	T	.7
F	F	.1

Sm	Pr	F	P(Pa|...)
T	T	T	.9
T	T	F	.1
T	F	T	.7
T	F	F	.1
F	T	T	.7
F	T	F	.1
F	F	T	.2
F	F	F	.1

Query: What is the probability that a student is smart, given that he/she passes the exam?

Complejidad de la inferencia Exacta

- Redes Singularmente Conectadas (Polytrees)
  - Dos nodos están conectados a lo mucho por un camino no dirigido
  - Complejidad de tiempo y espacio para el algoritmo de eliminación de variables es $O(d^k n)$

- Redes Multiplemente Conectadas
  - NP-hard, prueba por reduccion de 3-SAT

Algoritmos de Inferencia Exacta:

- Eliminacion de Variables
- Inferencia Simbólica (D'Ambrosio)
- Algoritmo de Paso de Mensajes (Pearl)
- Algoritmos de Clustering y Juntura de Árboles (Lauritzen, Spiegelhalter)

Referencias Bibliográficas

- Capitulo 14 Artificial Intelligence: A Modern Approach, Russell and Norvig
- http://www.cs.iastate.edu/~cs572/studyguide.html
- http://www.aaai.org/AITopics/pmwiki/pmwiki.php/AITopics/Uncertainty