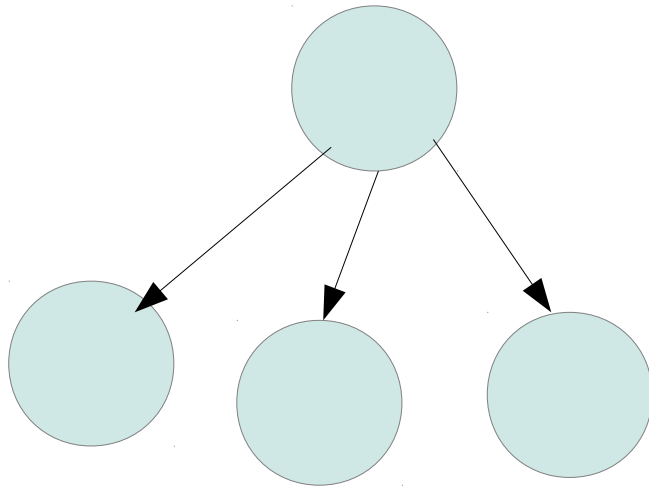


Razonamiento probabilístico y redes bayesianas



Clasificador bayesiano
Representación y semántica de las RB

Blanca A. Vargas Govea
vargasgovea@itesm.mx

Octubre 12, 2012

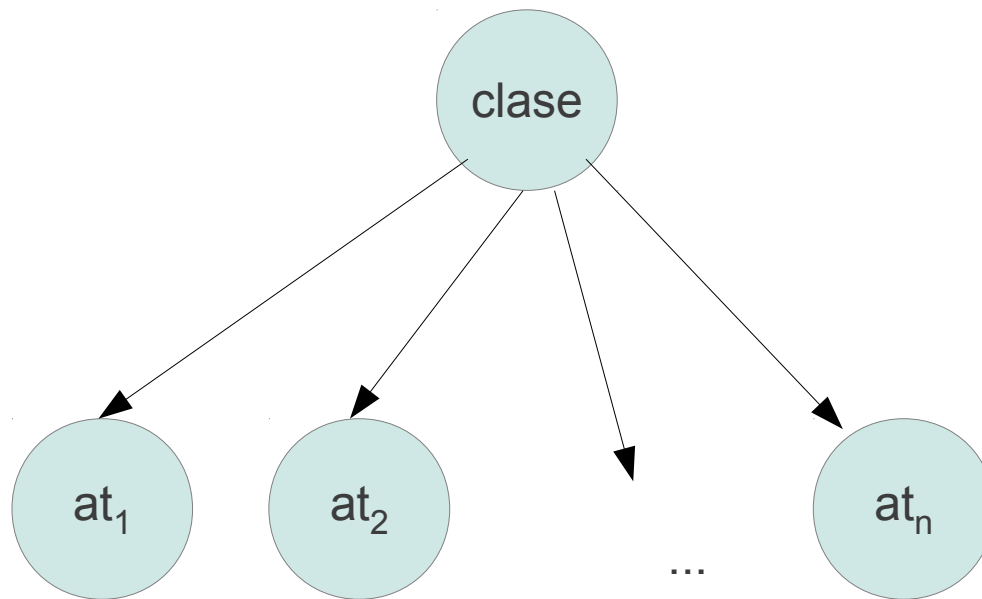
Correcciones al ejercicio de clase anterior

Objetivos

- Conocer y aplicar el clasificador bayesiano.
- Conocer cómo se representan las distribuciones de probabilidad mediante las redes bayesianas.
- Empezar a conocer la semántica de las redes bayesianas.

Clasificador bayesiano

- Asume independencia entre los atributos sobre la probabilidad de una clase determinada.



N	Ambiente	Temperatura	Jugar
1	Soleado	Alta	No
2	Soleado	Alta	No
3	Nublado	Alta	Si
4	Lluvioso	Media	Si
5	Lluvioso	Baja	Si
6	Lluvioso	Baja	No
7	Nublado	Baja	Si
8	Soleado	Media	No
9	Soleado	Baja	Si
10	Lluvioso	Media	Si
11	Soleado	Media	Si
12	Nublado	Media	Si
13	Nublado	Alta	Si
14	Lluvioso	Media	No

Ver clasificador bayesiano
ejnaiveb.xls

Redes bayesianas (RB)

- Tabla de distribución de probabilidad conjunta: no práctica.
- RB: estructura de datos que representa las dependencias entre variables.

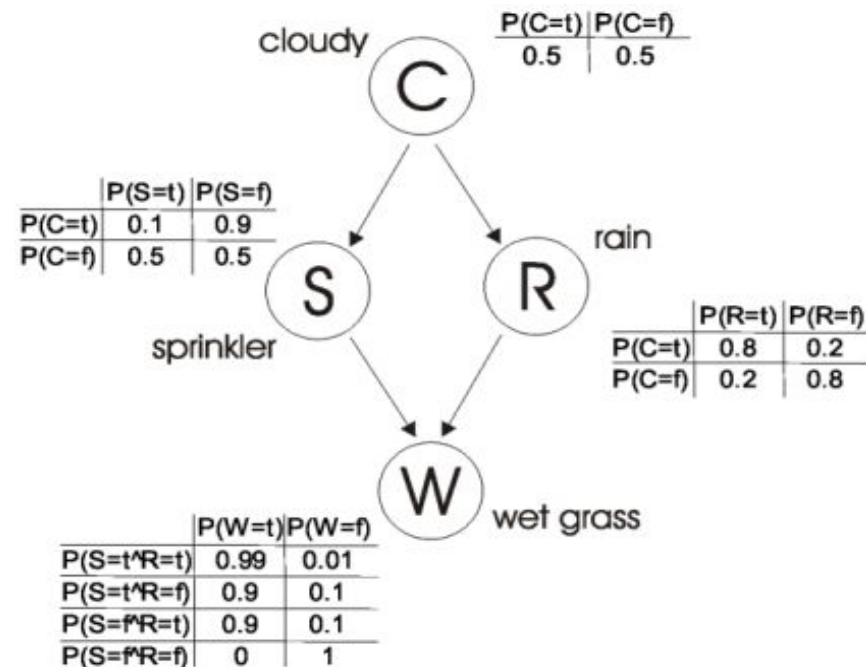
Red de creencia

Red causal

Red probabilista

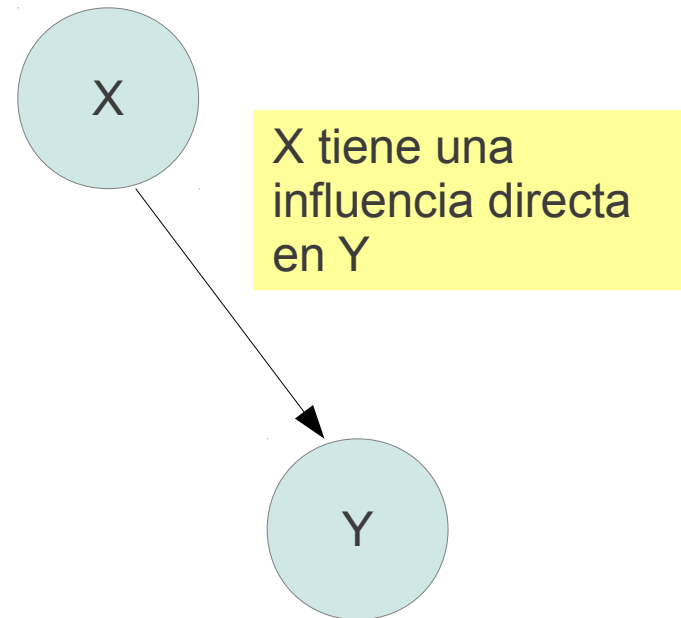
Redes bayesianas

Una RB es un **grafo dirigido** en el que cada nodo es anotado con información de probabilidad.



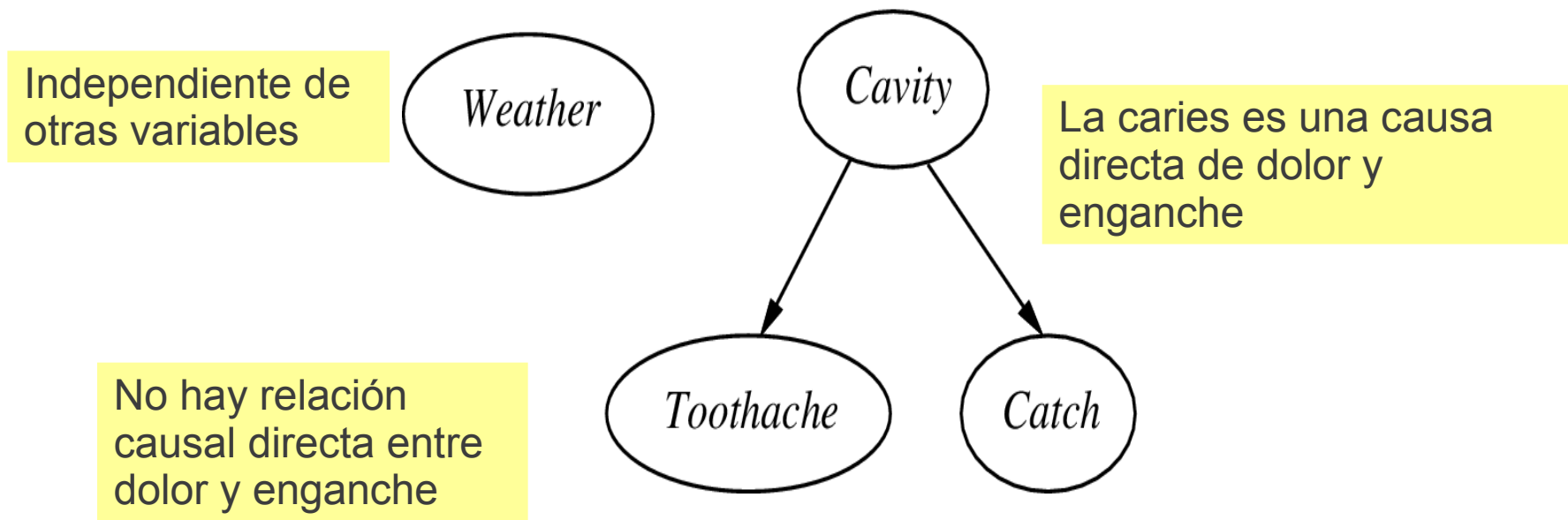
Especificación de una RB

- **Nodos:** conjunto de variables aleatorias. Se etiquetan con mayúsculas y sus variables con minúsculas. Por ejemplo, si el nodo A representa un interruptor donde $a_1 = \text{on}$, $a_2 = \text{off}$.
- **Arcos** dirigidos conectando pares de nodos.



Especificación de una RB

- **Distribución de probabilidad** condicional que cuantifica el efecto de los padres en el nodo. Especifica la distribución conjunta completa para todas las variables.
- El grafo no tiene ciclos. Grafo **acíclico** dirigido (GAD).



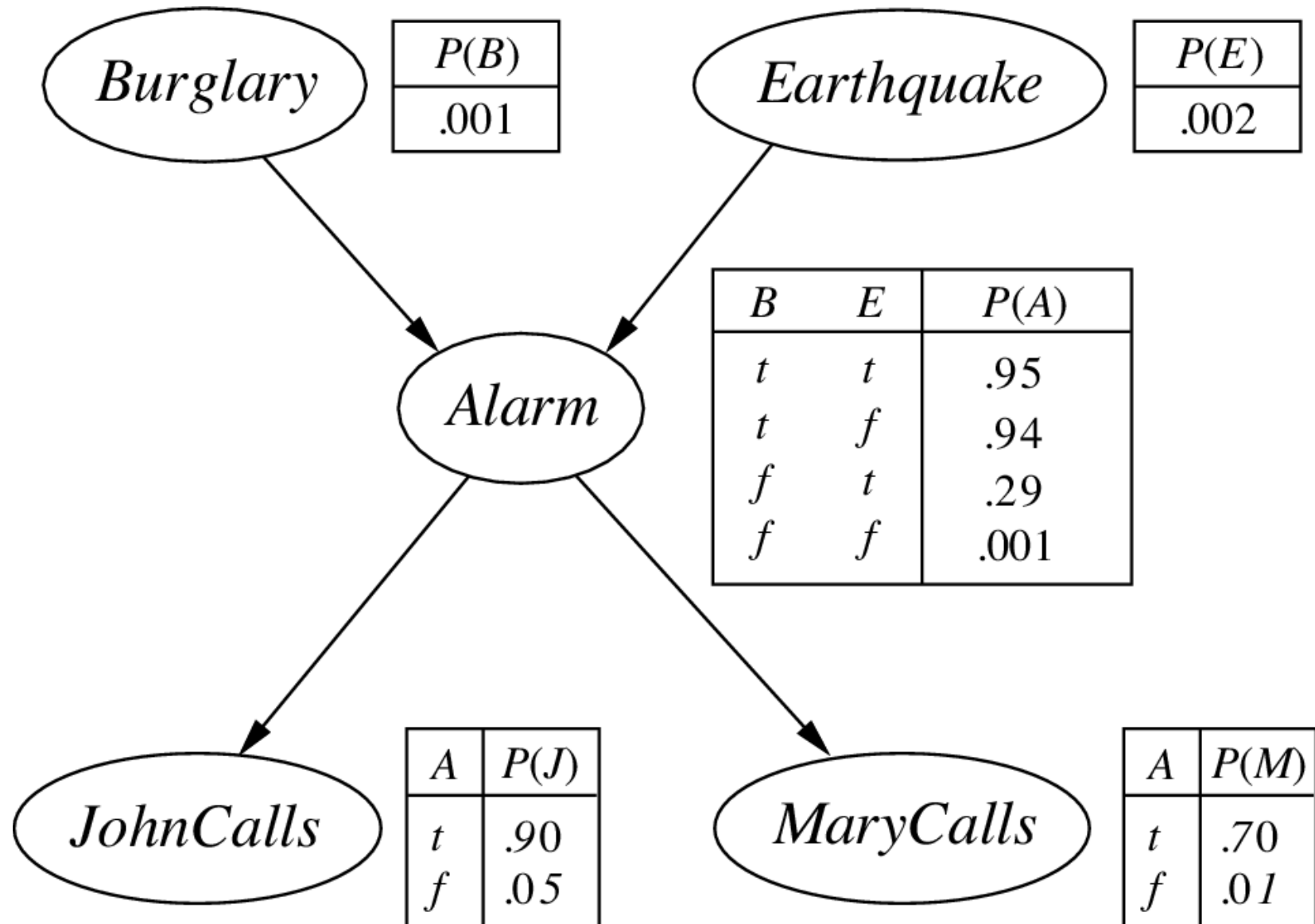
Ejemplo

- Instalaste una nueva alarma.
- Es confiable a detectar ladrones pero también responde a temblores.
- Tienes 2 vecinos: Mary y John que te llamarán al escuchar la alarma.
- John a veces confunde la alarma con el teléfono.
- Mary escucha música a alto volumen y a veces no oye la alarma.

¿Cómo se modelaría la RB si se deseara estimar la probabilidad de un robo?



Ejemplo



La topología: ¿qué significa?

- El robo y los temblores afectan directamente la probabilidad de activación de la alarma.
- El que John o Mary llamen depende solamente de la alarma.
- La red representa nuestras suposiciones sobre John y Mary: no perciben a los ladrones directamente, no notan los temblores, no consultan antes de llamar.

La topología: ¿qué significa?

- La red no tiene nodos que representen el que Mary escucha la música a alto volumen o que el teléfono confunda a John.
- Estos factores se resumen en la incertidumbre asociada con los arcos de Alarma a JohnLlama.
- La probabilidad resume un conjunto infinito de circunstancias en las que la alarma podría fallar.

Tabla de probabilidad condicional (CPT)

- Cada fila contiene la probabilidad condicional de las variables dados sus padres.
- Un nodo sin padres tiene una sola fila. Representa las probabilidades previas de cada posible valor de las variables.
- Cada renglón suma 1.
- Si la variable es booleana, el otro valor es $1-p$ por lo que se omite (como en este caso).

Semántica de las RB

Hay dos formas en las que se pueden entender

1. Pueden verse como una representación de la distribución de probabilidad conjunta.
2. También como una codificación de una colección de enunciados de independencia condicional.

Aunque son equivalentes, la primera nos sirve para entender cómo **construir** las redes mientras que la segunda nos ayuda a diseñar métodos de **inferencia**.

Representando la distribución conjunta completa

- Una RB proporciona una descripción completa del dominio.
- Cada entrada puede calcularse a partir de la información en la red.
- Una entrada es la probabilidad de una conjunción de asignación de valores a cada variable $P(X_1 = x_1 \cap \dots \cap X_n = x_n)$

que se abrevia como: $P(x_1, \dots, x_n)$

Representando la distribución conjunta completa

- Cada entrada en la distribución conjunta se representa por el producto de los elementos apropiados de las tablas (CPT).

$$P(x_1, \dots, x_n) = \prod_{i=1}^n P(x_i | \text{padres}(X_i))$$

Representando la distribución conjunta completa

- Podemos calcular la probabilidad de que la alarma haya sonado pero no hay ni robo ni temblor y que tanto Mary como John llamen:

$$\begin{aligned} P(j \cap m \cap a \cap \sim b \cap \sim e) &= \\ P(j|a)P(m|a)P(a|\sim b \cap \sim e)P(\sim b)P(\sim e) &= \\ 0.90 \times 0.70 \times 0.001 \times 0.999 \times 0.998 &= 0.00062 \end{aligned}$$

Ejercicio – clasificador bayesiano

Ejemplo aplicado al robo de autos [Meisner. 2003]

Hacer un clasificador bayesiano que permita, dados ciertos atributos de auto como color, tipo y origen, conocer la probabilidad de que sea robado.

Los datos se encuentran en: `dautos.xls`

Tarea

- El avance del proyecto. Incluye reporte y lo que llevan desarrollado – Martes 12
- Los puntos a cubrir se encuentran en el archivo octubre16.txt