

Curso de aprendizaje automatizado

PCIC, UNAM

Tarea 4: Gráficos probabilísticos

Fecha límite: 20 de mayo

Formato: Archivo ZIP con ejercicios.

Forma de entrega: Enviar correo a gibranfp@unam.mx, bere.mcic@gmail.com y richardt.pcic@gmail.com con asunto *Tarea de gráficos probabilísticos* y archivo adjunto.

Descripción

Realiza los ejercicios que se describen a continuación.

Ejercicio 1

En una planta nuclear hay una alarma que se activa cuando un indicador de temperatura excede un umbral. El indicador mide la temperatura del núcleo. Considera las variables booleanas A (alarma), D_A (alarma defectuosa), D_I (indicador defectuoso) y las variables enteras I (lectura del indicador) y T (temperatura real del núcleo).

- Dibuja 3 redes bayesianas válidas diferentes que capturen el comportamiento del proceso, entre ellas incluye aquella que capture el mayor número de independencias condicionales con el menor número de arcos. Discute el modelo representado por cada una de las redes.
- De las redes dibujadas, escribe sus distribuciones conjuntas en términos de sus probabilidades condicionales.
- Asumiendo que las variables enteras I y T pueden tomar un máximo de 100 valores, ¿cuál sería el número de valores necesarios en cada nodo y el total en cada una de las redes?

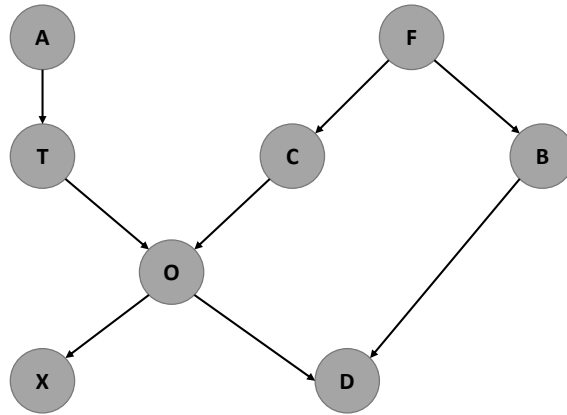


Figura 1: Modelo gráfico de diagnóstico de enfermedades de pulmón.

Ejercicio 2

A una clínica le concierne el diagnóstico de enfermedades de pulmón. Como se puede ver en el modelo de la Figura 1, una visita a Asia (A) hace que la probabilidad de tener tuberculosis (T) aumente. Los nodos en la gráfica tienen el siguiente significado:

Etiqueta	Significado
A	Visita a Asia
T	Tuberculosis
C	Cáncer de pulmón
O	Tuberculosis o Cáncer de pulmón
F	Fumador
X	Prueba de rayos X positiva
B	Bronquitis
D	Disnea

Di si las siguientes relaciones de independencia condicional son verdaderas o falsas y explica por qué:

- $T \perp\!\!\!\perp F \mid D$
- $C \perp\!\!\!\perp B \mid F$
- $A \perp\!\!\!\perp F \mid C$
- $A \perp\!\!\!\perp F \mid C, D$

Ejercicio 3

Imagina una clínica que ayuda a pacientes con ébola en un área afectada por el virus. La red de la Figura 2 intenta capturar la dinámica por la cual las personas que sufren los síntomas pueden llegar a esta clínica y ver a un especialista. Existe la posibilidad de que alguien con ébola ($E = \text{verdadero}$) muestre síntomas, por ejemplo sangrado ($S = \text{verdadero}$) fiebre ($F = \text{verdadero}$) y visite la clínica ($V = \text{verdadero}$). El sangrado incrementa el riesgo de complicaciones ($C = \text{verdadero}$) y la persona puede ser llevada a ver un doctor especialista ($D = \text{verdadero}$).

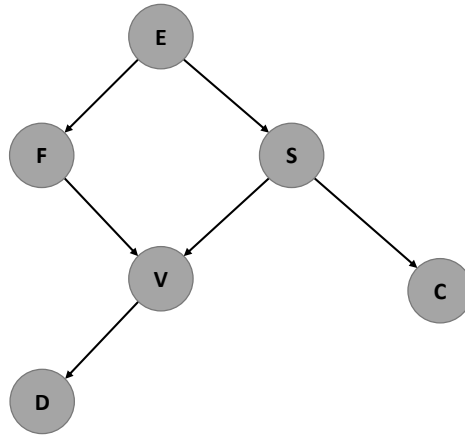


Figura 2: Modelo gráfico de detección de ébola.

- $P(E = \text{verdadero}) = 0.01$
- $P(F = \text{verdadero} | E = \text{verdadero}) = 0.6$
- $P(F = \text{verdadero} | E = \text{falso}) = 0.1$
- $P(S = \text{verdadero} | E = \text{verdadero}) = 0.8$
- $P(S = \text{verdadero} | E = \text{falso}) = 0.05$
- $P(V = \text{verdadero} | F = \text{verdadero}, S = \text{verdadero}) = 0.8$
- $P(V = \text{verdadero} | F = \text{verdadero}, S = \text{falso}) = 0.5$
- $P(V = \text{verdadero} | F = \text{falso}, S = \text{verdadero}) = 0.7$
- $P(V = \text{verdadero} | F = \text{falso}, S = \text{falso}) = 0.0$
- $P(C = \text{verdadero} | S = \text{verdadero}) = 0.75$
- $P(C = \text{verdadero} | S = \text{falso}) = 0.1$
- $P(D = \text{verdadero} | V = \text{verdadero}) = 0.6$
- $P(D = \text{verdadero} | V = \text{falso}) = 0.0$

Con estas probabilidades y el modelo gráfico de la Figura 2, realiza lo siguiente:

- a. Escribe la distribución conjunta de la red bayesiana en función de las probabilidades condicionales.
- b. Si un paciente es llevado al doctor ($D = \text{verdadero}$), usando un paquete de software calcula la probabilidad de que no tenga ébola ($P(E = \text{falso} | D = \text{verdadero})$).
- c. Convierte a la red bayesiana en un modelo gráfico no dirigido (campo aleatorio de Markov) y dibújalo. Captura tantas relaciones de independencia condicional como sea posible.

d. Debido a una campaña de concientización de la salud, las personas son alentadas a visitar la clínica en caso de que tengan fiebre. Esto incrementa la cantidad de visitas de personas con fiebre sin importar el estado de cualquier otra variable.

- ¿Qué probabilidades condicionales en la red se modifican debido a este cambio y en qué sentido?
- Describe cualquier efecto que esto tenga en la proporción de personas con complicaciones que visiten la clínica. Menciona exactamente qué probabilidades condicionales usaste para llegar a tu conclusión.

e. Asume que alguien que no tiene fiebre va al doctor, ¿qué relación de independencia condicional existe en la distribución que no puede ser descubierta a través del grafo solamente?

Opcional Calcula $P(V|E = \textit{verdadero})$ usando el método de eliminación de variables, describiendo de forma detallada el procedimiento que seguiste (2 puntos extras).