

Curso de doctorado

RAZONAMIENTO BAYESIANO

Ejercicios

Nota: Los ejercicios resueltos deben ser enviados a la siguiente dirección antes del **14 de septiembre**:

Prof. Severino Fernández Galán
Dpto. Inteligencia Artificial
E.T.S.I. Informática de la UNED
C/ Juan del Rosal, 16.
28040 Madrid (España)

Se ruega los envíen en copia impresa **en papel**. Si por alguna razón particular debiera enviarlos por correo electrónico, hágalo en formato PDF o Postscript; **no se admitirán los ejercicios enviados en Word.**

Ejercicio 1

Algunos ejemplos de fuentes de incertidumbre en sistemas expertos son los siguientes: *información incompleta*, *información errónea*, *información imprecisa*, *mundo real no determinista*, *modelo incompleto* o *modelo inexacto*. Para cada fuente de incertidumbre citada dé un ejemplo de dominio del mundo real en el que dicha fuente esté presente. Justifique su respuesta y no utilice como ejemplo de dominio el de la medicina.

Ejercicio 2

En cierto problema de diagnóstico se han identificado 5 posibles diagnósticos y 10 hallazgos asociados. Cada diagnóstico o hallazgo puede estar *ausente* o *presente*. ¿Cuántos parámetros probabilistas independientes habría que estimar para poder aplicar el *teorema de Bayes* a la tarea de diagnóstico en este problema? ¿Cuántos parámetros habría que estimar si se decidiera utilizar el *método probabilista clásico*?

Ejercicio 3

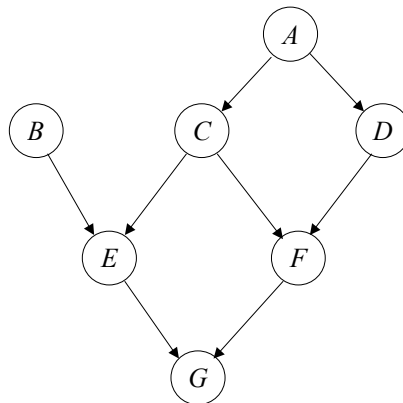
Dé un ejemplo tomado del mundo real en que se violen las hipótesis de exclusividad de diagnósticos y de independencia condicional de hallazgos dado el diagnóstico, establecidas por el *método probabilista clásico*.

Ejercicio 4

Dada la red bayesiana de la figura, justificar si las siguientes afirmaciones son correctas:

- a) $I(A, B)$
- b) $I(B, G | E)$
- c) $I(E, F | A)$
- d) $I(E, D | G)$
- e) $I(C, D | A, E)$
- f) $I(A, G | F)$

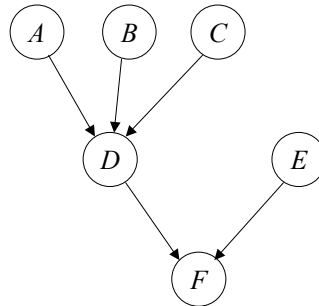
donde $I(X, Y | Z)$ significa “ X e Y son condicionalmente independientes dado Z ”.



Ejercicio 5

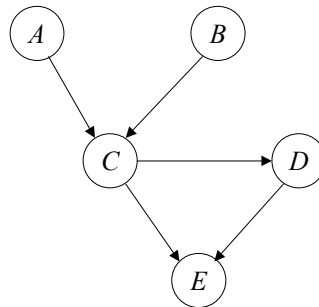
Dada la red bayesiana de la figura, formada por variables binarias, explique cuántos parámetros probabilistas independientes habría que definir en dicha red para cada uno de los siguientes casos:

- a) Los nodos de las dos familias de la red interactúan mediante el *modelo general*.
- b) Los nodos de las dos familias interactúan mediante el modelo de *puerta O probabilista con causas no explícitas*, también denominada *puerta O residual*.



Ejercicio 6

Dada la red bayesiana de la figura y las tablas de probabilidades condicionales asociadas a la misma, determinar $P(b \mid e = \text{presente})$ mediante el método de eliminación de variables.



$P(a)$	$a=\text{presente}$	$a=\text{ausente}$
	0.2	0.8

$P(b)$	$b=\text{presente}$	$b=\text{ausente}$
	0.65	0.35

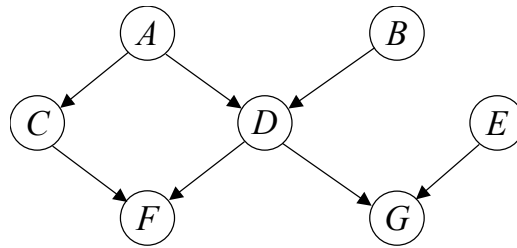
$P(c \mid a, b)$	$c=\text{presente}$	$c=\text{ausente}$
$a=\text{presente}$ $b=\text{presente}$	0.3	0.7
$a=\text{presente}$ $b=\text{ausente}$	0.85	0.15
$a=\text{ausente}$ $b=\text{presente}$	0.7	0.3
$a=\text{ausente}$ $b=\text{ausente}$	0.6	0.4

$P(d \mid c)$	$d=\text{presente}$	$d=\text{ausente}$
$c=\text{presente}$	0.25	0.75
$c=\text{ausente}$	0.9	0.1

$P(e \mid c, d)$	$e=\text{presente}$	$e=\text{ausente}$
$c=\text{presente}$ $d=\text{presente}$	0.11	0.89
$c=\text{presente}$ $d=\text{ausente}$	0.3	0.7
$c=\text{ausente}$ $d=\text{presente}$	0.1	0.9
$c=\text{ausente}$ $d=\text{ausente}$	0.9	0.1

Ejercicio 7

Utilice el *algoritmo de agrupamiento con árbol de unión* para hallar las probabilidades a posteriori de las variables de la red bayesiana de la figura, sabiendo que $G=1$.



Las probabilidades asociadas a cada nodo de la red bayesiana son las que aparecen en las tablas siguientes:

a	$p(a)$	b	$p(b)$	a	c	$p(c a)$	a	b	d	$p(d a,b)$	e	$p(e)$	c	d	f	$p(f c,d)$	d	e	g	$p(g d,e)$
0	0.3	0	0.6	0	0	0.25	0	0	0	0.3	0	0.3	0	0	0	0.1	0	0	0	0.22
1	0.7	1	0.4	0	1	0.75	0	0	1	0.7	1	0.7	0	0	1	0.9	0	0	1	0.78
				1	0	0.3	0	1	0	0.7			0	1	0	0.2	0	1	0	0.1
				1	1	0.7	0	1	1	0.3			0	1	1	0.8	0	1	1	0.9
							1	0	0	0.2			1	0	0	0.3	1	0	0	0.1
							1	0	1	0.8			1	0	1	0.7	1	0	1	0.9
							1	1	0	0.4			1	1	0	0.4	1	1	0	0.5
							1	1	1	0.6			1	1	1	0.6	1	1	1	0.5

Ejercicio 8

Un jugador de fútbol ha recibido dos ofertas económicas para fichar por dos equipos, A y B. Las dos ofertas se basan en los resultados deportivos que se consigan a final de temporada, según la siguiente tabla:

Clasificación en la liga	Equipo A	Equipo B
1°	1 millón de euros	2.8738 millones de euros
2°	0.5 millones de euros	2 millones de euros
3°	0.1 millones de euros	1 millón de euros
4°	0.01 millones de euros	0.5 millones de euros
≥ 4°	0.001 millones de euros	0.0001 millones de euros

El jugador tomará la decisión de por qué equipo fichar siguiendo el *principio de máxima utilidad*: elegir la opción que tenga la mayor utilidad esperada. Para ello su representante ha recopilado las estadísticas clasificatorias de los dos equipos en la liga, las cuales aparecen en la siguiente tabla:

Clasificación en la liga	Probabilidad para el equipo A	Probabilidad para el equipo B
1°	0.01	0.001
2°	0.02	0.002
3°	0.04	0.002
4°	0.06	0.033
≥ 4°	0.87	0.962

¿Por qué oferta se decantará el jugador?

Ejercicio 9

Un propietario de locales decide alquilar o no un local de ensayo a grupos musicales en función de si tienen o no algún disco a la venta. Si no se alquila un local a cierto grupo musical, se sabe que el coste medio de mantener el local cerrado hasta que aparezca un

nuevo grupo musical es de 5 unidades. Si se alquila el local y el grupo paga el alquiler, el propietario obtiene un beneficio medio de 500 unidades. En caso de que el grupo sea moroso, se sufren unas pérdidas medias de 80 unidades derivadas del pago de los recibos de agua, luz, etc. Por otro lado, de los archivos que guarda el propietario se deduce que el porcentaje de morosos entre grupos que tienen algún disco a la venta es del 8% y sube al 90% en el caso de grupos sin discos a la venta. También se deduce de los archivos que, de los grupos que acuden en busca de local para alquilar, el 60% tiene algún disco a la venta. Acabados los ensayos, existe una determinada probabilidad de que haya que hacer reparaciones en el local por un valor medio de 60 unidades. Los datos de que dispone el propietario sobre qué tipo de grupos son los que provocan que haya que hacer reparaciones son los siguientes:

$P(\text{destrozos} \mid \text{disco}, \text{paga})$	$+\text{destrozos}$	$\neg\text{destrozos}$
$+\text{disco}$ $+\text{paga}$	0.02	0.98
$+\text{disco}$ $\neg\text{paga}$	0.12	0.88
$\neg\text{disco}$ $+\text{paga}$	0.05	0.95
$\neg\text{disco}$ $\neg\text{paga}$	0.75	0.25

Construya un diagrama de influencia para este problema y evalúe paso a paso el árbol de decisión correspondiente. ¿Cuál es la política de actuación en este problema y la utilidad esperada para el propietario? ¿Hasta qué porcentaje debería subir el número de grupos que, no teniendo disco a la venta, no son morosos, para que se cambie en la política de actuación la decisión de alquilar o no locales a los grupos sin algún disco a la venta?