Definición formal de red bayesiana

Una red bayesiana es:

- Un conjunto de variables **proposicionales** V
- Un **conjunto de relaciones binarias** definida sobre las variables de V, E
- Una **distribución de probabilidad conjunta** **P** sobre las variables de V

tales que:

- (V, E) forman un grafo **acíclico**, **conexo** y **dirigido** G
- (G, P) cumplen las hipótesis de **independencia condicional**, también llamadas de separación direccional

Hipótesis de independencia condicional

Un grafo acíclico conexo y dirigido $G = (V, E)$ y una distribución de probabilidad conjunta P definida sobre las variables del grafo se dice que cumplen las hipótesis de independencia condicional o separación direccional, si

$$\forall X \in V \text{ y } \forall Y \in V - \{X \cup de(X) \cup pa(X)\}$$

se tiene que

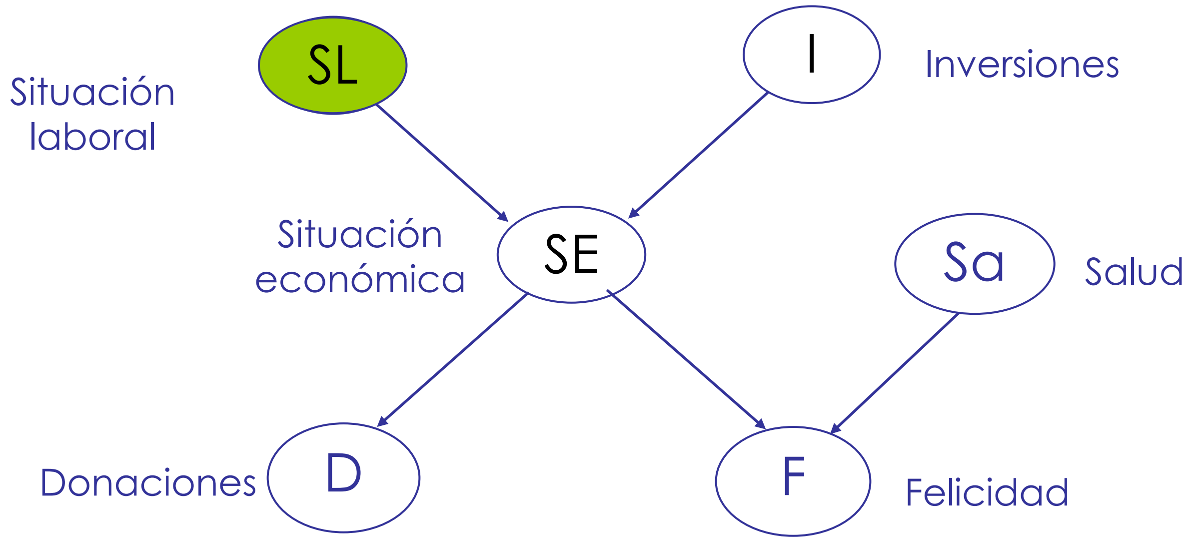
$$X \text{ es independiente de } Y \text{ dado } pa(X)$$

Se cumplen las hipótesis de independencia condicional si

$\forall X \in V$ *y* $\forall Y \in V - \{X \cup de(X) \cup pa(X)\}$ se tiene que X es indep. de Y *dado* $pa(X)$

¿Cómo obtener la lista de independencias que implica una estructura?

$\forall X \in V$ y $\forall Y \in V - \{X \cup de(X) \cup pa(X)\}$ se tiene que X es indep. de Y *dado* $pa(X)$

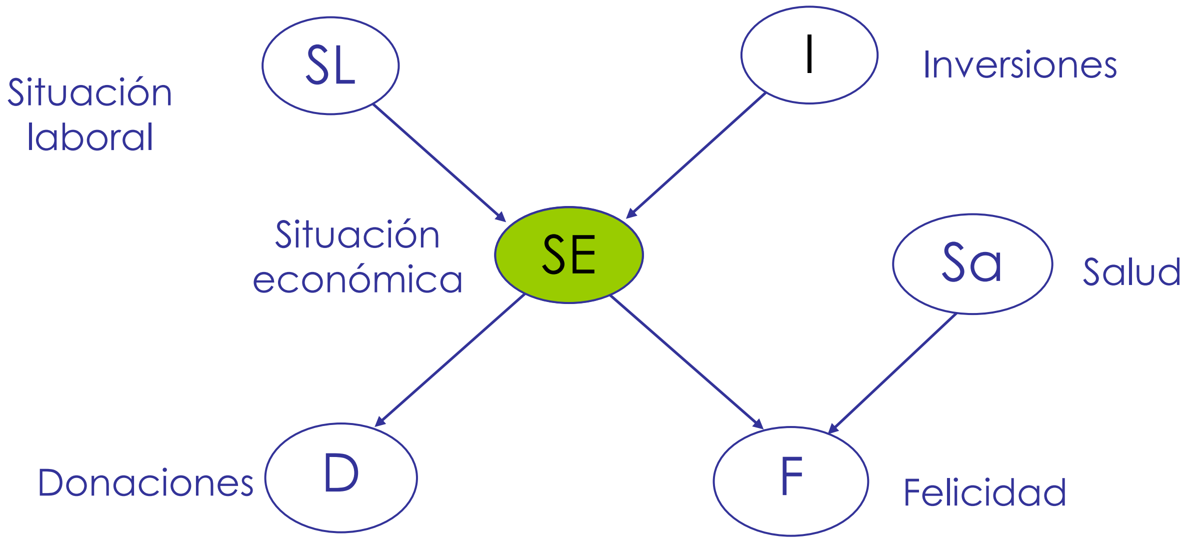


Paso 1. Construimos el conjunto $V - \{X \cup de(X) \cup pa(X)\}$

Paso 2. Construimos el conjunto $pa(X)$

Paso 3: X es independiente de Y *dado* $pa(X)$

$\forall X \in V$ y $\forall Y \in V - \{X \cup de(X) \cup pa(X)\}$ se tiene que X es indep. de Y *dado* $pa(X)$



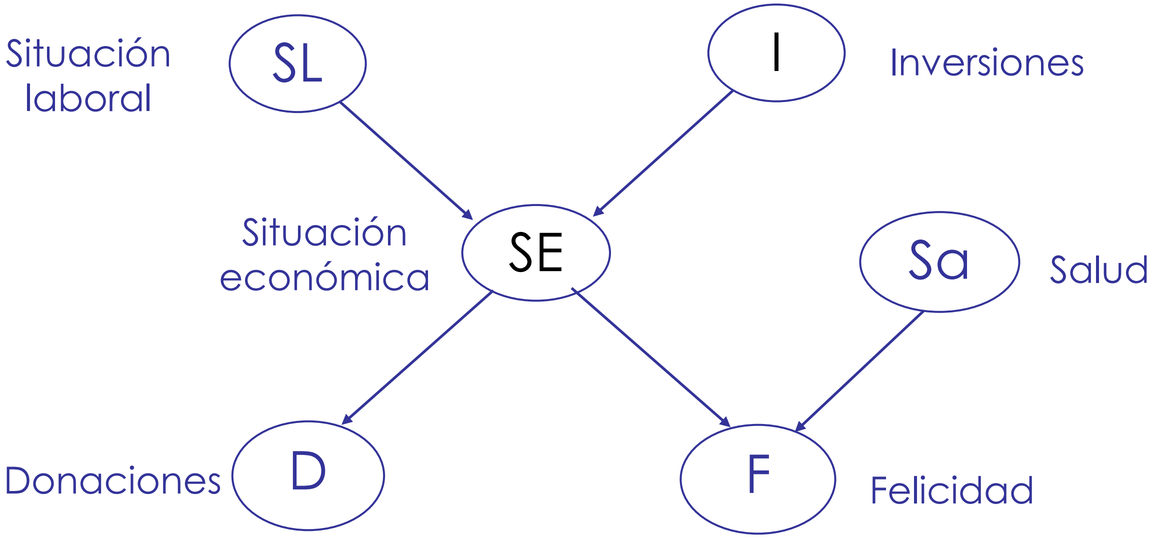
Paso 1. Construimos el conjunto $V - \{X \cup de(X) \cup pa(X)\}$

Paso 2. Construimos el conjunto $pa(X)$

Paso 3: X es independiente de Y *dado* $pa(X)$

¿Cómo obtener la lista completa de independencias de una red?

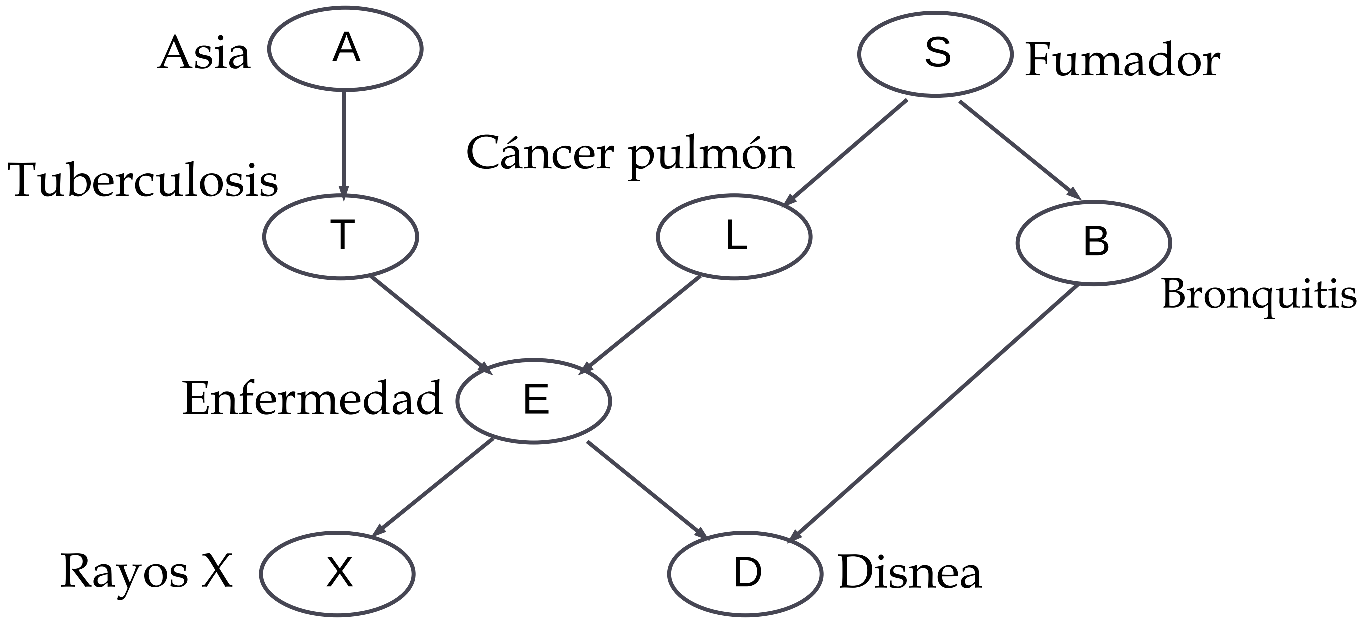
$\forall X \in V$ y $\forall Y \in V - \{X \cup de(X) \cup pa(X)\}$ se tiene que *X* es indep. de *Y dado pa(X)*



Haciendo lo mismo para cada nodo, sacaríamos el listado completo de independencias:

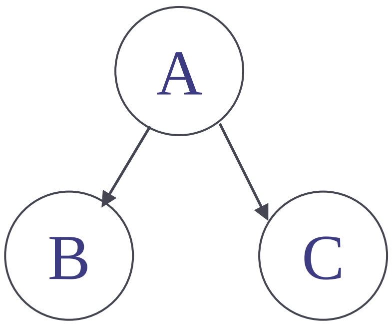
- SL es independiente de I, Sa
- SE es independiente de Sa dados {SL, I}

Ejercicio propuesto: sacar el listado completo de independencias de la siguiente estructura:

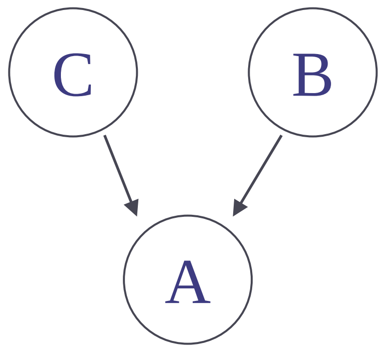


Significado de las relaciones de independencia

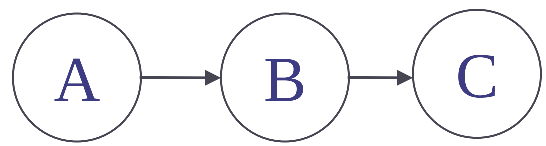
Estudiemos y analicemos el significado de las relaciones de independencia condicional para todas las posibles estructuras de redes con tres nodos



Cola con cola

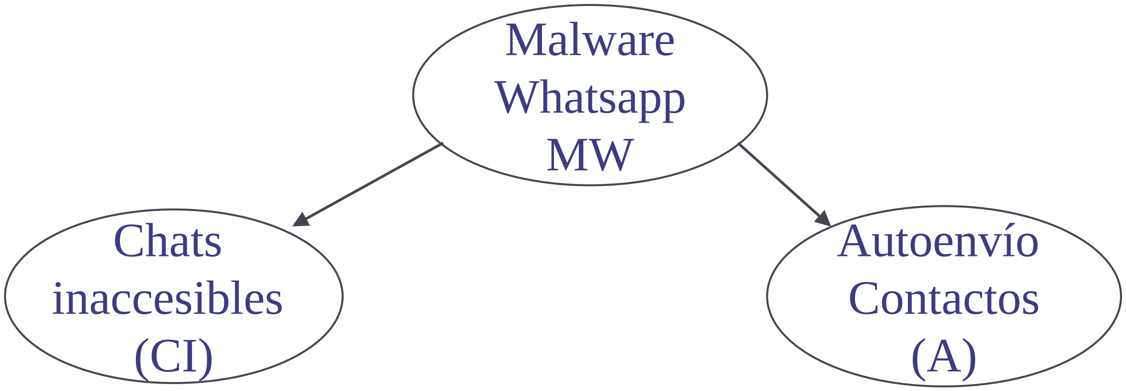


Cabeza con cabeza



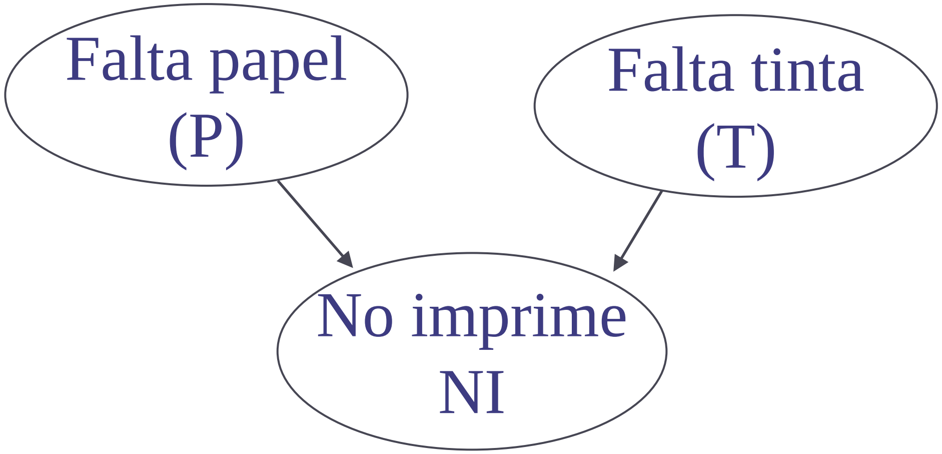
Cabeza con cola

Cola con cola



Cola con cabeza

Cabeza con cabeza



Teorema fundamental:

Dada una red bayesiana, su distribución de probabilidad puede expresarse como:

$$P(x_1, \dots, x_n) = \prod_{x_i} P(x_i \,/\, pa(x_i))$$

Método probabilístico clásico

Individuos	Edad	Obesidad	Hernia	Indigestión	Vómitos
Individuo 1	Mayor_50	no	no	no	no
Individuo 2	Mayor_50	no	no	no	no
Individuo 3	Mayor_50	no	no	no	no
Individuo 4	Mayor_50	no	no	no	no
Individuo 5	Mayor_50	no	sí	no	sí
Individuo 6	Mayor_50	sí	sí	no	sí
Individuo 7	Mayor_50	sí	sí	no	sí
Individuo 8	Mayor_50	sí	sí	no	si
Individuo 9	Menor_50	no	no	no	no
Individuo 10	Menor_50	no	no	no	no
Individuo 11	Menor_50	no	no	no	no
Individuo 12	Menor_50	no	no	no	no
Individuo 13	Menor_50	no	no	no	no
Individuo 14	Menor_50	no	no	no	no
Individuo 15	Menor_50	no	no	no	no
Individuo 16	Menor_50	no	no	no	no
Individuo 17	Menor_50	no	no	sí	sí
Individuo 18	Menor_50	sí	no	no	sí
Individuo 19	Menor_50	sí	no	sí	sí
Individuo 20	Menor_50	sí	no	no	no

Calcular la distribución conjunta de las variables

P(Edad= Mayor_50; Obesidad=no; Hernia=no; Indigestión=no, Vómitos=no) = 4/20
P(Edad= Mayor_50; Obesidad=no; Hernia=sí; Indigestión=no, Vómitos=sí) = 1/20
P(Edad= Mayor_50; Obesidad=sí; Hernia=sí; Indigestión=no, Vómitos=sí) = 3/20
P(Edad= Menor_50; Obesidad=no; Hernia=no; Indigestión=no, Vómitos=no) = 8/20
P(Edad= Menor_50; Obesidad=no; Hernia=no; Indigestión=si, Vómitos=si) = 1/20
P(Edad= Menor_50; Obesidad=sí; Hernia=no; Indigestión=no, Vómitos=sí) = 1/20
P(Edad= Menor_50; Obesidad=sí; Hernia=no; Indigestión=si, Vómitos=sí) = 1/20
P(Edad= Menor_50; Obesidad=sí; Hernia=no; Indigestión=no, Vómitos=no) = 1/20

A partir de la distribución conjunta calcular:

- Las distribuciones marginales:

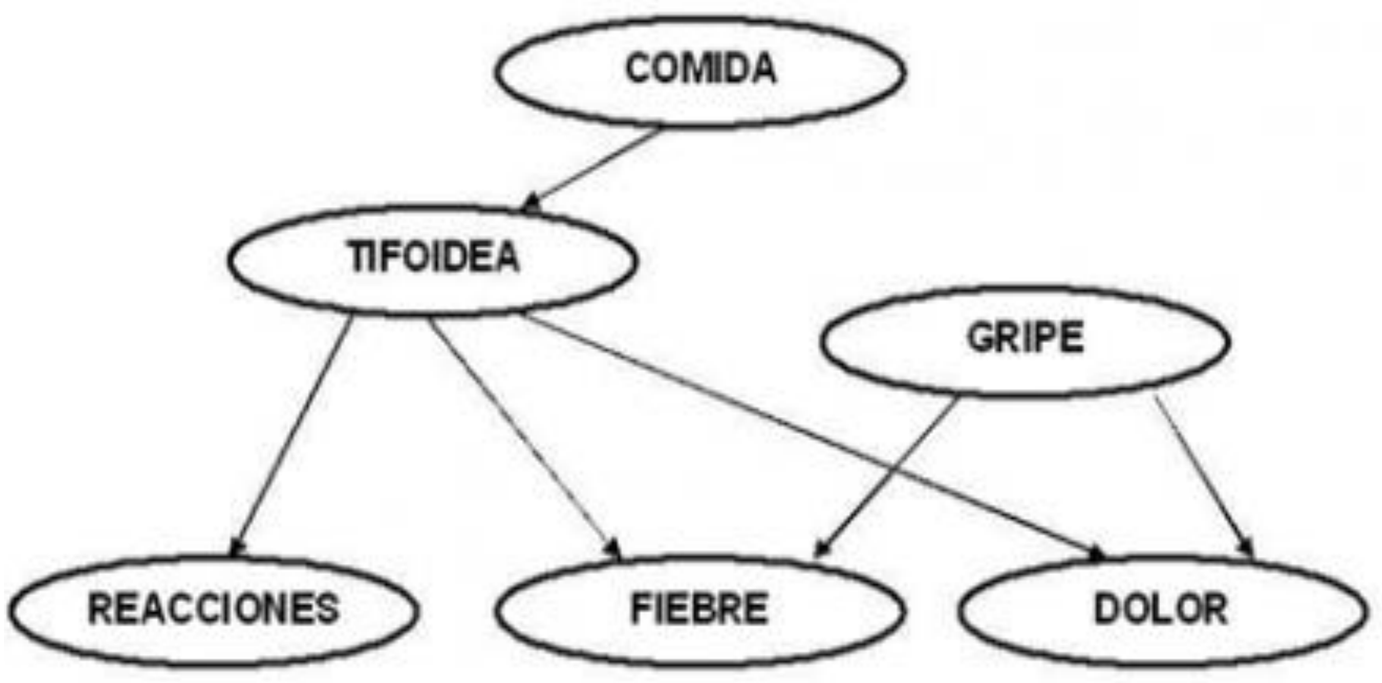
  P(Edad = Mayor_50) y P(Edad = Mayor_50, Hernia= sí) **8/20** **4/20**

- P(Hernia=si/Vómitos=si) (diagnóstico) **4/7**

- P(Vómitos=si/Obesidad=si, Hernia=si) (predicción) **1**

¿Cómo aplicar el Teorema fundamental a una red?

La distribución conjunta sería P(C,T,G,R,F,D)
La distribución condicionada de cada nodo dados sus padres sería:



P(C,T,G,R,F,D) = P(C)*P(T/C) *P(G) *P(R/T) *P(F/T,G) *P(D)

P(¬c,+t,¬g,+r,+f,+d) = P(¬c)*P(+t/¬c) *P(¬g) *P(+r/+t) *P(+f/+t, ¬g) *P(+d/+t, ¬g))

Importancia del teorema de independencia condicional

Dada una red bayesiana, su distribución de probabilidad conjunta puede expresarse como:

$$P(x_1, ..., x_n) = \prod P(x_i/pa(x_i))$$

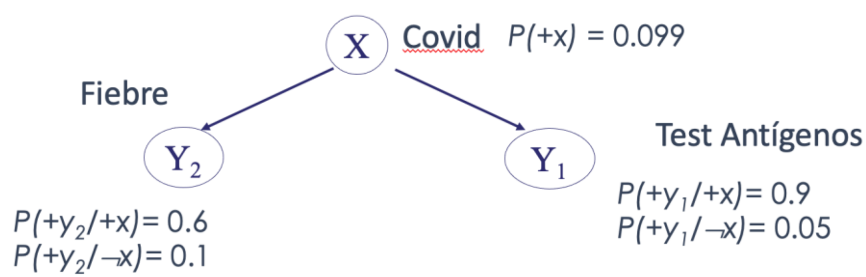
Ejemplo:

En el enfoque antiguo, necesitábamos la distribución conjunta P(X1, X2,...., X9)

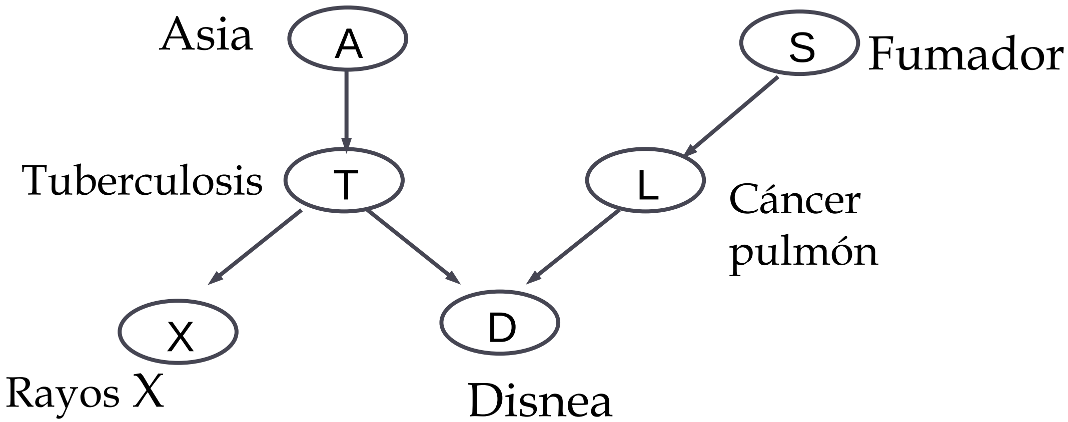
El teorema nos dice que esos valores podemos calcularlos a partir de P(X1), P(X2/X1), ... P(X9/X8)

Además, para el experto esos valores tienen más sentido

(prevalencia, falsos positivos, falsos negativos, etc).



Ejercicio de repaso completo

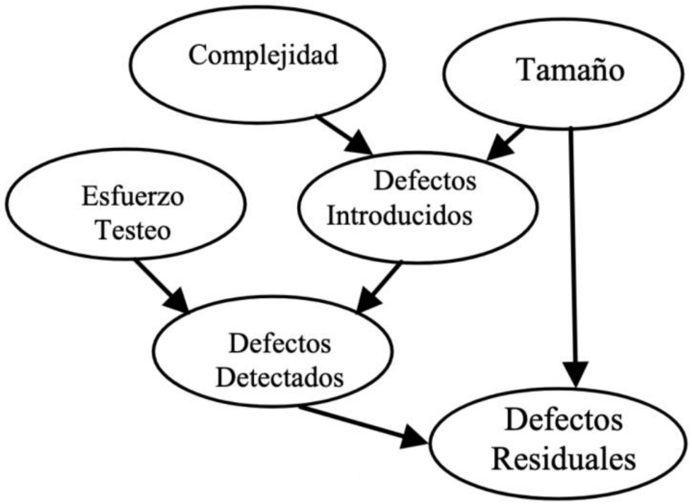


Se pide:

- ¿Qué independencias/dependencias entre las variables de la red implican las hipótesis de independencia condicional?
- Si suponemos ciertas las hipótesis de independencia condicional, ¿cuántas probabilidades sería necesario especificar?
- Si no podemos suponer las hipótesis de independencia condicional, ¿qué probabilidades deberíamos pedir al experto? ¿Cuántos valores son, en total?
- ¿Cómo podemos calcular la probabilidad conjunta a partir de las condicionadas? Aplicando el teorema de factorización, indica cómo se calcularía la probabilidad de que el paciente fume dado que tiene disnea (mirad el ejemplo 2.6 en pág. 29)

Ejercicio propuesto

Dada la siguiente red sobre detección de defectos en ingeniería del software:



DOI: 10.4272/978-84-9745-204-5.CH10 · Corpus ID: 63607562

Redes bayesianas en la ingenieria del software

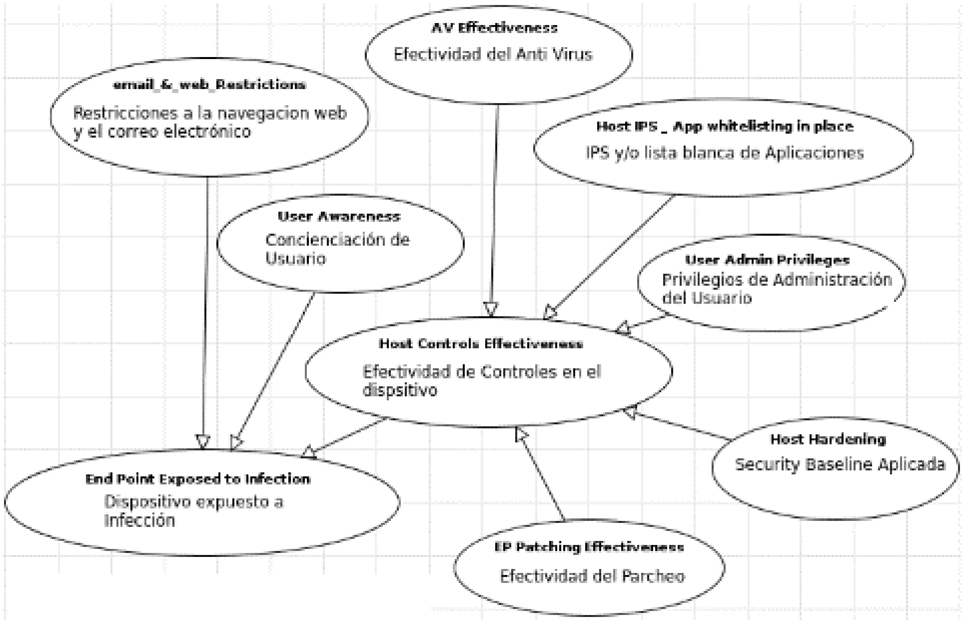
P. J. Gonzalez, I. Román, José Javier Dolado Cosín · Published 2004 · Computer Science

Se pide:

- ¿Qué independencias/dependencias entre las variables de la red se tienen si sabemos que la red es bayesiana?
- En ese caso, ¿cuántas probabilidades sería necesario especificar?. Da una estimación para estos valores que sea coherente con el sentido común.
- Si no sabemos que la red es bayesiana, ¿qué probabilidades deberíamos pedir al experto? ¿Cuántos valores en total?
- Aplica el teorema fundamental a esta red

Ejercicio propuesto

Dada la siguiente red sobre infección por virus informáticos:



Infección por virus informáticos: Una aplicación de las redes Bayesianas (parte 1)

septiembre 17, 2015 · por Miguel A. Sánchez

Se pide:

- ¿Qué independencias/dependencias entre las variables de la red se tienen si la red es bayesiana?
- Si suponemos que es una red bayesiana y que todas las variables son binarias, ¿cuántas probabilidades sería necesario especificar?
- ¿Cómo se aplicaría el teorema fundamental a esta red?