Modelado con redes bayesianas

Un modelo es una representación selectiva de la realidad. El proceso de construir modelos se denomina **modelado**, y en el enfoque de redes bayesianas consta de:

- Modelado **cualitativo,** en el que se han de definir;
  - Los nodos o variables (deben tomar un conjunto **EXHAUSTIVO y EXCLUYENTE** de valores).
  - Las relaciones entre los nodos (deben representar **influencia causal**)
- Modelado **cuantitativo,** en el que se especifican los parámetros.
  - Por el teorema fundamental, los parámetros necesarios son las distribuciones de probabilidad condicionada de cada nodo dados sus padres (incluyendo la probabilidad a priori de los nodos sin padre)

Guía para el modelado con redes bayesianas

Paso 1. Identificar la información relevante en el problema (toda la que pueda influir en el proceso de razonamiento)

Paso 2. Representar dicha información como variables y sus determinar sus posibles valores

(asegúrate que sean un conjunto exhaustivo y excluyente).

Paso 3. Identificar las relaciones

(asegúrate que representen relaciones de influencia causal)

Paso 4. Obtener o estimar los parámetros

(asegúrate de que sean coherentes con tipo de relación y el sentido común)

Más sobre modelado: Identificación de las variables

Por ejemplo, **en problemas de diagnóstico**

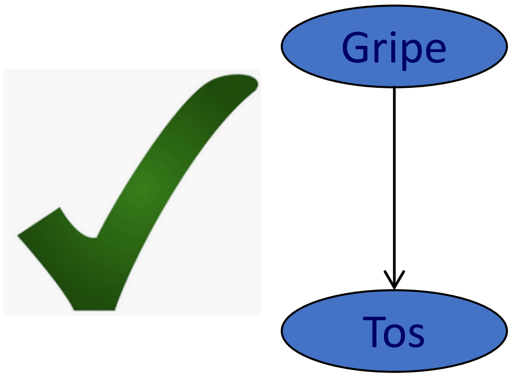
- ¿Cuál es la situación/problema que se plantea?
- ¿Qué posibles causas pueden explicar esta situación?
- ¿Qué otros factores pueden hacer que los problemas o causas ocurran, o impedir que ocurran?
- ¿De qué evidencia se dispone para soportar dichas causas, problemas o factores?

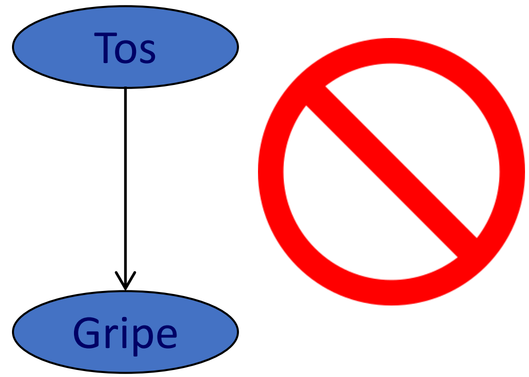
Más sobre modelado: Identificación de las variables

Tipo variable	Breve descripción
Objetivo	Modelan objetos de interés. No observables directamente.
Observación	Modelan la forma de medir variables objetivo. Pueden ser observadas directamente
Factor	Modelan fenómenos que afectan a otras variables del modelo
Promotor	La variable afectada es más probable cuando están presentes.
Inhibidor	La variable afectada es menos probable cuando están presentes.
Requerido	Si no entra en acción, no ocurre la variable afectada.
Preventivo	Si entra en acción, no ocurre la variable afectada.
Auxiliares	Usadas por conveniencia (para simplificar el modelo)

Más sobre modelado: tipos de relaciones

Dos tipos de relaciones:

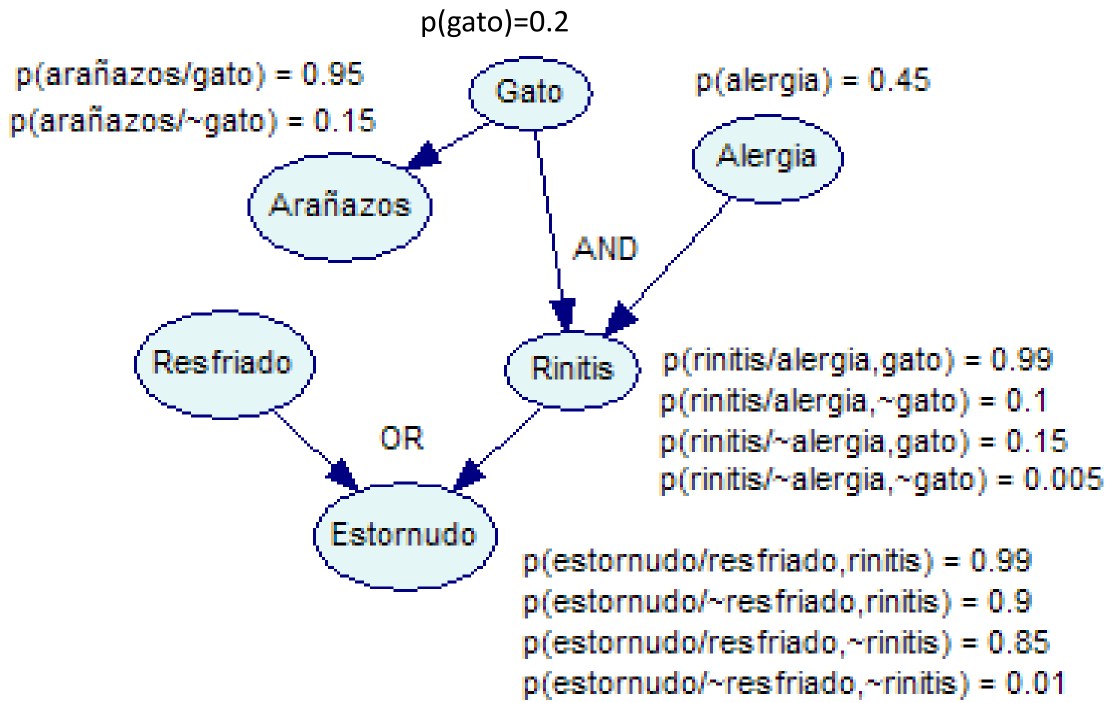




Relación causal	Regla de diagnóstico
Son las que se utilizan en redes bayesianas	NO deben utilizarse en redes bayesianas

Si definimos los enlaces de forma que modelen la **relación causal**, el modelo obtenido es más sencillo de entender y menos complejo en cuanto a parámetros y relaciones.

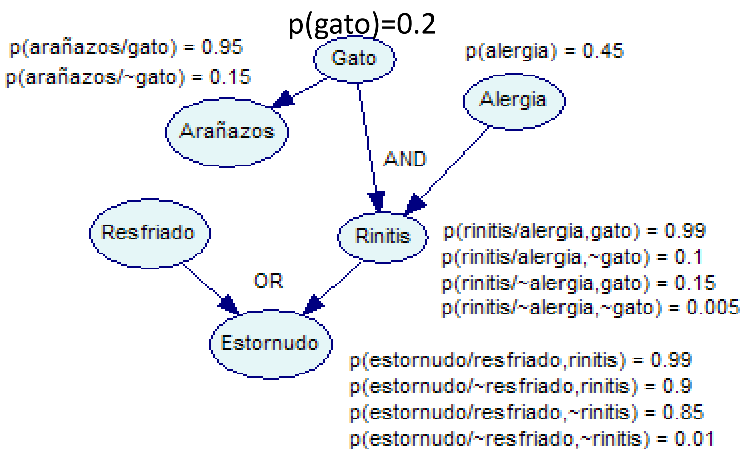
Ejemplo: el problema del estornudo

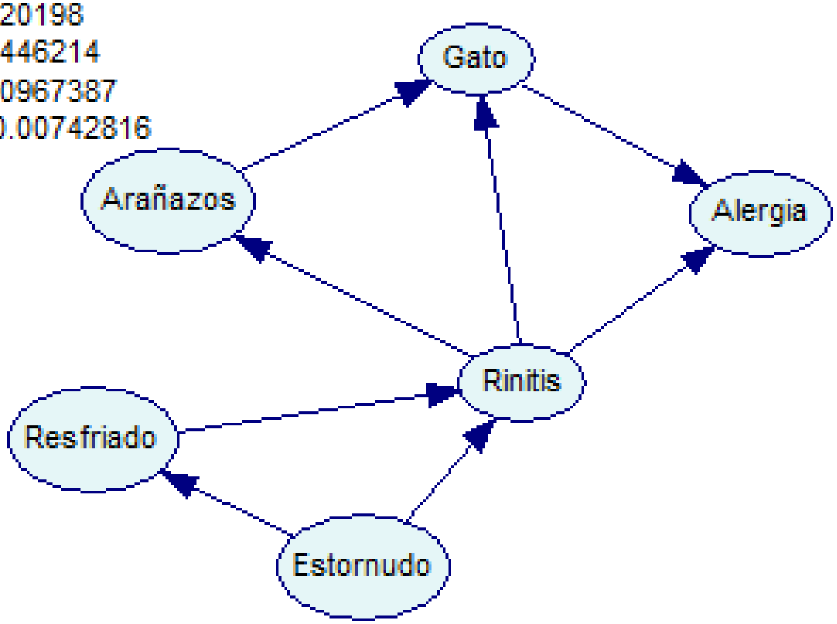


Hemos modelado esta red con relaciones de influencia causal. Esta opción es correcta y resulta más sencilla de comprender y construir para los seres humanos

Ejemplo: el problema del estornudo

p(gato/arañazos,rinitis) = 0.920198
p(gato/arañazos,~rinitis) = 0.446214
p(gato/~arañazos,rinitis) = 0.0967387
p(gato/~arañazos,~rinitis) = 0.00742816

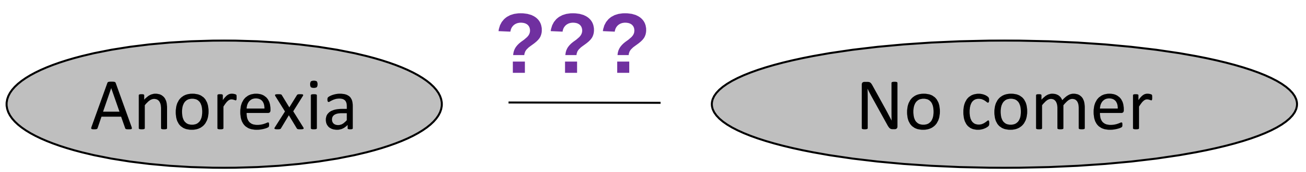




Mediante técnicas matemáticas, es posible invertir los enlaces de la red y obtener una equivalente, cuyas relaciones serían más interpretables como reglas de diagnóstico. Esta red es correcta, pero resulta difícil de comprender (y más aún de especificar) para los seres humanos. Por tanto esta opción, siendo correcta, no es recomendable

Más sobre modelado

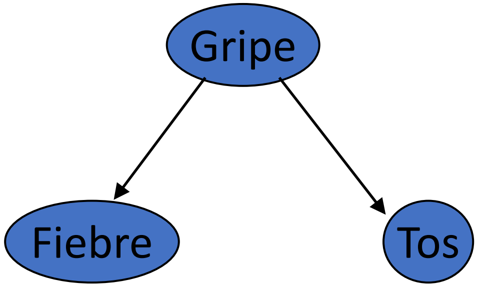
Identificar relaciones de influencia causal no siempre será fácil:



Algunos trucos útiles: **Verificar las independencias del modelo**

Comprobar si las relaciones entre las variables reflejan adecuadamente las dependencias e independencias.

Modelo A



¿Qué independencias refleja esta red?

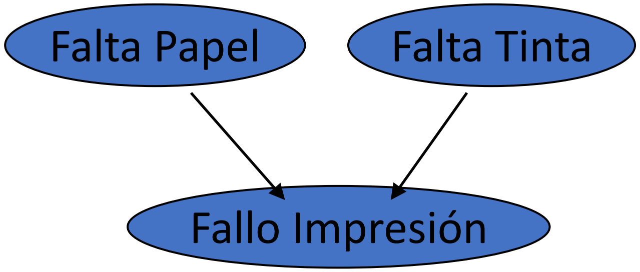
Opción A: Fiebre y Tos son dependientes a priori pero independientes dado Gripe

Opción B: Fiebre y Tos son independientes a priori pero dependientes dado Gripe

Algunos trucos útiles:
Verificar las independencias del modelo

Comprobar si las relaciones entre las variables reflejan adecuadamente las dependencias e independencias.

Modelo B

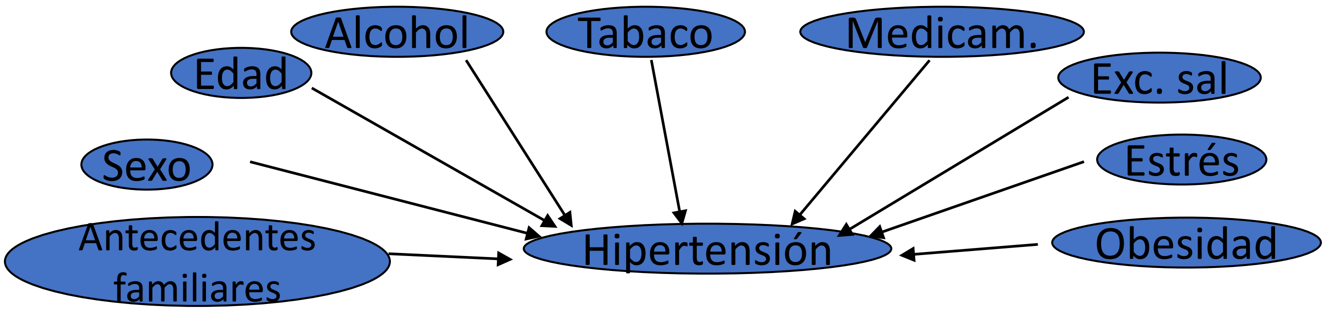


¿Qué independencias refleja esta red?

Opción A: Falta Papel y Falta Tinta son independientes a priori pero dependientes dado Fallo impresión

Opción B: Falta Papel y Falta Tinta son dependientes a priori pero independientes dado Fallo Impresión

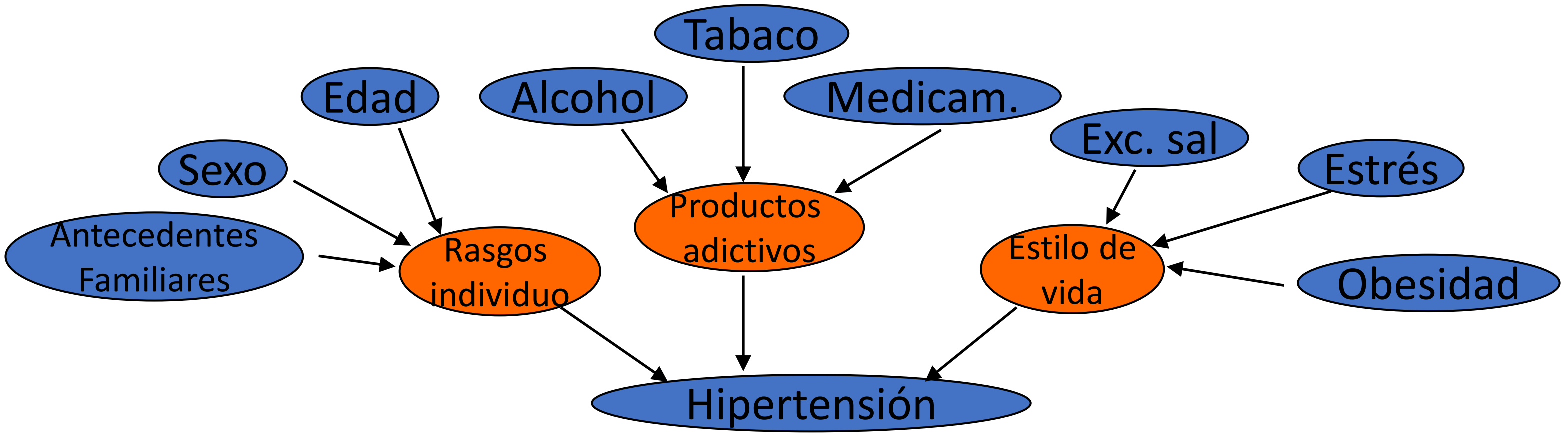
Algunos trucos útiles: estructura



¿Cuántos parámetros tenemos que introducir en este modelo si suponemos todas las variables binarias?

Opción A: 10
Opción B: $9 + 2^9$
Opción C: $2^{10} - 1$

Algunos trucos útiles: estructura

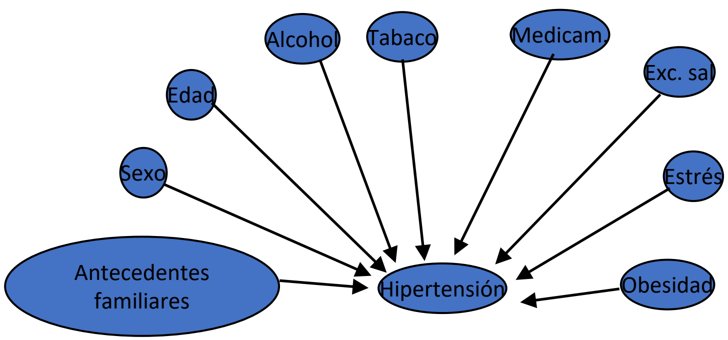


¿Cuántos parámetros tenemos que introducir en este modelo si suponemos todas las variables binarias?

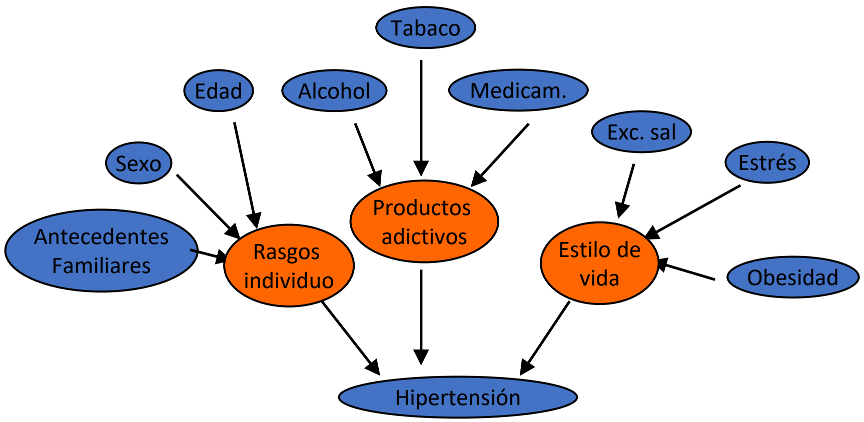
Opción A: 13
Opción B: $9 + 4 \cdot 2^3$
Opción C: $2^{13} - 1$

Algunos trucos útiles: estructura

¿Cuál de estos dos redes bayesianas es más compleja?



Opción A



Opción B

Número de parámetros requeridos:

Opción A: $9 + 2^9 = 9 + 64 \cdot 2^3$

Opción B: $9 + 4 \cdot 2^3$

Introducir nodos intermedios es una buena idea para reducir la complejidad

Más sobre modelado: **parámetros**

- **Especificación directa** de los parámetros (con ayuda de expertos, costoso)
  - En ausencia de información, se suele suponer equiprobabilidad
  - En muchos casos, se podrán utilizar relaciones tipo AND u OR, con su correspondiente ruido

- **Aprendizaje** (si existen bases de datos)
  - *Aprendizaje de parámetros* (si se dispone de la estructura);
  - *Aprendizaje estructural* (se aprende tanto la estructura como los parámetros)

- **Combinaciones** de los dos anteriores (iterativamente)

Algunos trucos útiles: **modelos canónicos**

- Son modelos que ayudan a simplificar la especificación de las tablas de probabilidades condicionadas, para ciertos tipos de relación causal entre variables.
- Los más utilizados son los modelos NOISY-OR y NOISY-AND.

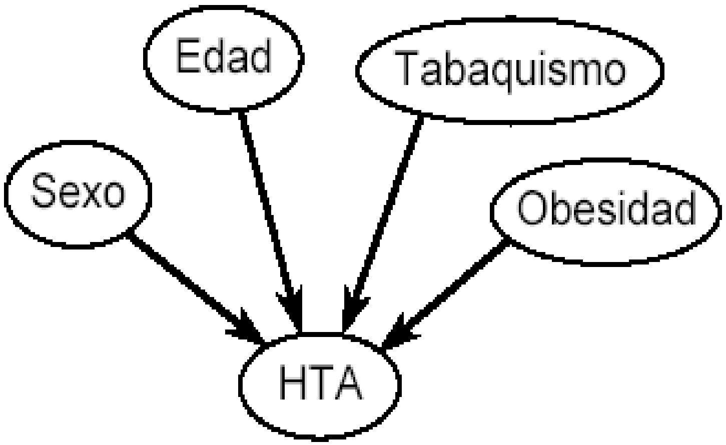
Modelo canónico: la puerta OR

Hipótesis

1. Cada una de las causas, por sí misma, puede producir el efecto y basta con que una de ellas esté presente para que el efecto ocurra
2. Cuando todas las causas están ausentes, el efecto está ausente
3. No hay interacción entre las causas

Si se cumplen las hipótesis de la puerta OR a partir de las **probabilidades condicionadas por un solo padre** podemos calcular las probabilidades condicionadas por todos los padres.

Ejemplos

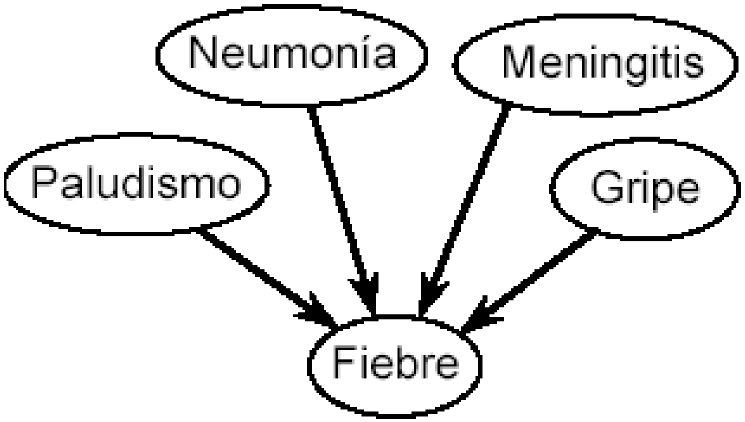


¿Puede modelarse con una puerta OR?

Opción A: SÍ

Opción B: NO

Ejemplos



¿Puede modelarse con una puerta OR?

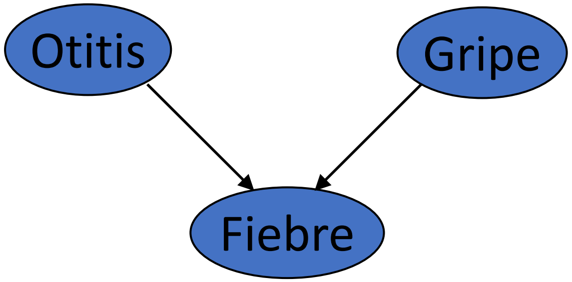
Opción A: SÍ

Opción B: NO

Modelo OR

Sea c_x = probabilidad de que la causa x provoque el efecto cuando todas las otras causas están ausentes.

Conocidas todas las c_x podemos calcular las probabilidades condicionadas por los padres



Conocidos $c_g = P(+f/+g, \neg o)$=0.8 y $c_o = P(+f/\neg g,+o)$=0.6, los parámetros de la red se pueden calcular así:

P(+f/+g, +o)

P(+f/¬g, +o)

P(+f/+g, ¬o)

P(+f/ ¬g, ¬o)

Si queremos incluir en el modelo que *es posible que otras causas no determinadas provoquen también fiebre* necesitamos además de las probabilidades de que cada causa provoque el efecto por separado un factor de ruido

Modelo canónico: la puerta NOISY-OR

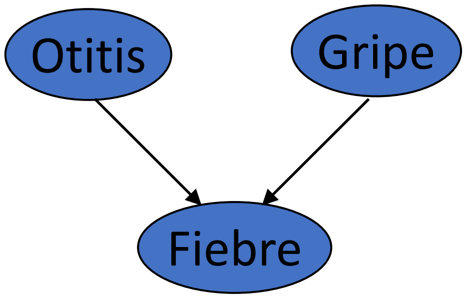
Hipótesis

1. Cada una de las causas, por sí misma, puede producir el efecto y basta con que una de ellas esté presente para que el efecto ocurra
2. Hay causas desconocidas no incluidas en la red que puede producir el efecto cuando todas las causas especificadas en la red están ausentes (ruido)
3. El efecto solo se produce en presencia de alguna de las causas modeladas o de ruido
4. No hay interacción entre las causas

Si se cumplen las hipótesis de la puerta NOISY-OR: a partir de las probabilidades condicionadas por un solo padre y el factor de ruido podemos calcular las probabilidades condicionadas por todos los padres.

Modelo NOISY-OR (puerta OR ruidosa)

Veamos como a partir de $c_{g,}$ c_o y c_r, se determinan las probabilidades necesarias para el modelo



c_o = P(+f/¬g, +o , ¬r) = 0.6

c_g = P(+f/+g, ¬o , ¬r) = 0.8

c_r = P(+f/¬g, ¬o , +r) = 0.01

Parámetros necesarios en la red:

P(+f/+g, +o)

P(+f/¬g, +o)

P(+f/+g, ¬o)

P(+f/ ¬g, ¬o)

Por las hipótesis 1 y 4:

P(+f/+g,+o,¬r)= P(+f/+g,¬o,¬r) + P(¬f/+g, ¬o, ¬r)P(+f/¬g,+o,¬r)= 0.8+0.2*0.6 =0.92

Por la hipótesis 3:

P(+f/¬g, ¬o, ¬r) = 0

Modelo NOISY-OR (puerta OR ruidosa)

P(+f/¬g, ¬o) = P(+f/¬g, ¬o, +r) + P(+f/¬g, ¬o, ¬r) = $\mathbf{c_r}$ + 0 = $\mathbf{c_r}$

Las otras probabilidades necesarias en la red se calculan mediante:

P(+f/+g,+o) =

P(+f/+g,¬o) = **+**

P(+f/¬g, +o)=

Probabilidad de fiebre, dada cierta combinación de causas

+

Si dada esa combinación no hay fiebre, que actúe la causa no modelada (ruido)

La puerta OR para más de dos variables

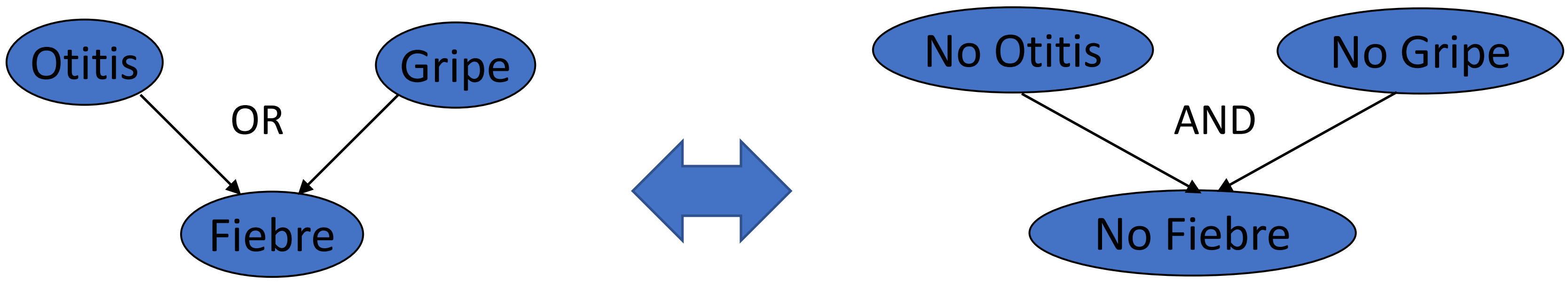
En general, si en un modelo tenemos un efecto X y $U_1,\ldots,U_n$ son las causas posibles, si denotamos por c_i a la probabilidad que tiene cada causa de producir el efecto por separado y por q_i a la probabilidad complementaria ($q_i = 1\text{-}c_i$), entonces:

$$P(\neg\, x/U_1\,, \ldots, U_n) = \prod_{i \in T_U} q_i$$

donde T_u = conjunto de causas de X que están presentes.

Como en el caso anterior, el modelo se puede generalizar al caso de que haya cierto ruido (práctica 2: tres causas)

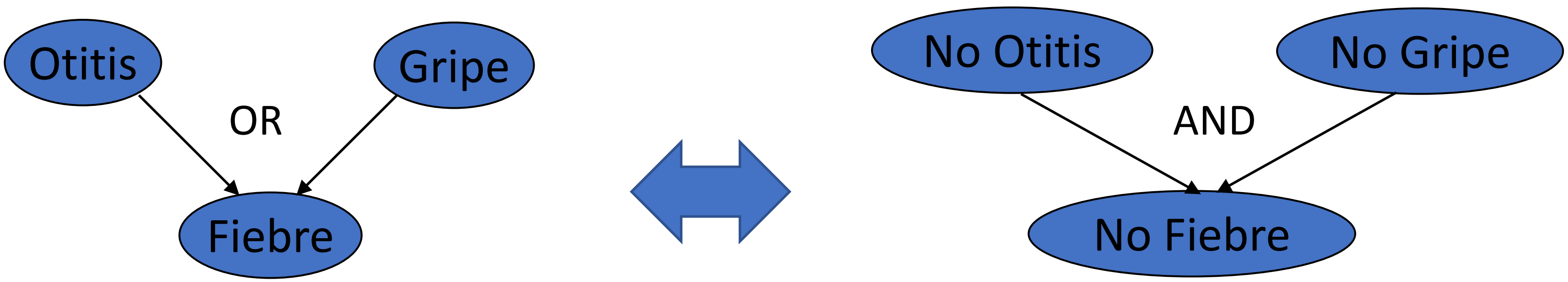
Modelo AND: Relación con Modelo OR



Leyes de Morgan:

$$\neg(A\ and\ B) \equiv \neg A\ or\ \neg B$$
$$\neg(A\ or\ B) \equiv \neg A\ and\ \neg B$$

Ejercicio 1



Teniendo en cuenta las relaciones entre los dos modelos, determina los parámetros del modelo AND a partir de los datos $c_{\neg g} = P(\neg f / \neg g, +o) = 0.2$ y $c_{\neg o} = P(\neg f / +g, \neg o) = 0.4$.

Recuerda: los parámetros a determinar son

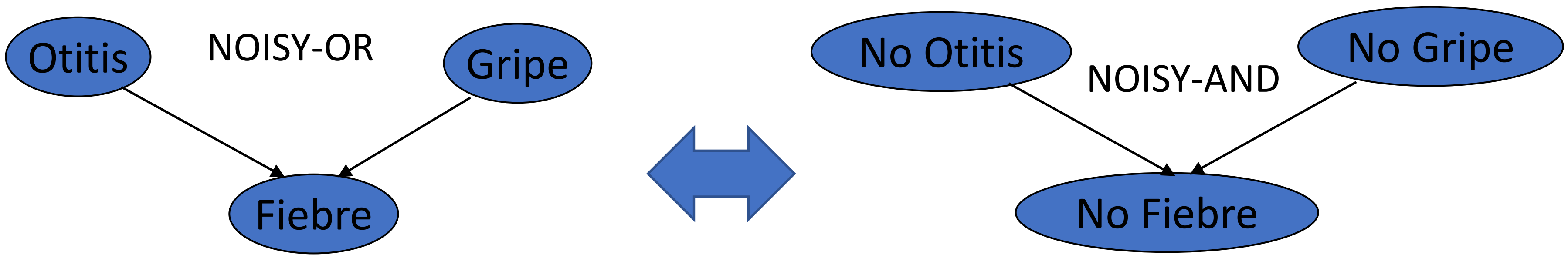
$$P(\neg f / +g, +o)$$

$$P(\neg f / \neg g, +o)$$

$$P(\neg f / +g, \neg o)$$

$$P(\neg f / \neg g, \neg o)$$

Ejercicio 2



Teniendo en cuenta las relaciones entre los dos modelos, determina los parámetros del modelo NOISY AND a partir de los datos $c_{\neg g} = P(\neg f/\neg g, +o)=0.2$, $c_{\neg o} = P(\neg f/+g,\neg o)=0.4$ y $c_{r*} = P(\neg f/\neg g, \neg o)=0.9$.

Recuerda: los parámetros a determinar son

$$P(\neg f/+g, +o)$$

$$P(\neg f/\neg g, +o)$$

$$P(\neg f/+g, \neg o)$$

$$P(\neg f/ \neg g, \neg o)$$