Aplicación de Redes Bayesianas Temporales al Diagnóstico y Predicción de Fallos en Generadores de Vapor

Severino Fernández Galán
http://www.ia.uned.es/~seve/
Departamento de Inteligencia Artificial
Universidad Nacional de Educación a Distancia (UNED), España

RESUMEN

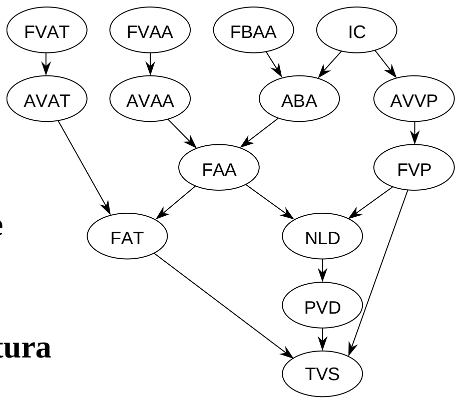
- **Definición de los objetivos del sistema**
- **Análisis de las características principales del dominio**
- **Elección del método de representación de conocimiento adecuado**
- **Descripción del diseño e implementación del sistema**
- **Ejemplo de aplicación**
- **Conclusiones**

OBJETIVOS DEL SISTEMA

- **Modelar el proceso general de propagación de fallos en un generador de vapor de una central termoeléctrica**
- **Ayudar a los técnicos de la central en el diagnóstico y predicción de dichos fallos, producidos a partir de una perturbación inicial en el generador**

CAUSALIDAD

- ***Proceso de propagación de fallos*: secuencia de eventos temporales causalmente interrelacionados**
  - ➤***Eventos iniciales*: fallo en válvulas o bombas**
  - ➤***Eventos intermedios*: apertura de vávulas, aceleración en bombas, incremento en los flujos de agua o vapor…**
  - ➤***Evento final*: decremento en la temperatura del vapor sobrecalentado**



SEVERINO F. GALÁN

INCERTIDUMBRE

- **Fuente de incertidumbre: en cada arco el retardo temporal entre causa y efecto no es constante**

FBAA → ABA

10	10	10	10	11	11	11	11	12	12	12	12	12	12	12	12	13	13	13	13	15	15
15	15	18	18	18	18	20	20	20	20	20	20	20	20	21	21	21	21	21	21	21	21
22	22	22	22	27	27	27	27	29	29	29	29	29	29	29	29	29	29	29	29	31	31
31	31	32	32	32	32																

TIEMPO

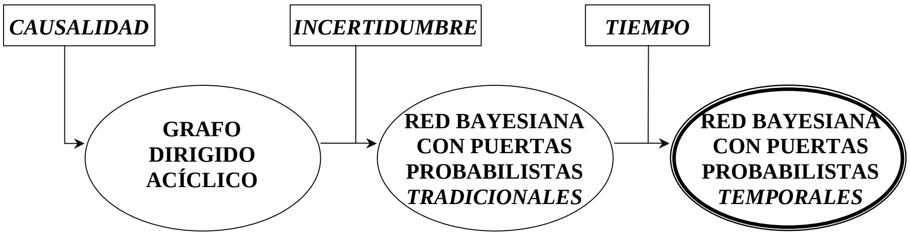
- **En general, la relación causal entre un par de eventos no es instantánea:**

FBAA → ABA: [10, 32]

IC → ABA: [29, 107]

IC → AVVP: [3, 29]

ELECCIÓN DEL MÉTODO DE REPRESENTACIÓN DE CONOCIMIENTO



Aplicación de RBTs al Diagnóstico y Predicción de Fallos en Generadores de Vapor

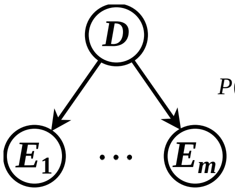
MÉTODO PROBABILISTA CLÁSICO

- **Orígenes: años 60**
- **Objetivo: desarrollo de sistemas de diagnóstico basados en el teorema de Bayes**
- **Ej.: medicina**
  - ➤**Posibles hallazgos (ej.: síntomas):** $e_1, \ldots, e_m$
  - ➤**Posibles diagnósticos:** $d_1, \ldots, d_n$
  - ➤**Se quiere conocer:** $P(d_1, \ldots, d_n \mid e_1, \ldots, e_m)$
  - ➤**Se conoce:** $P(e_1, \ldots, e_m \mid d_1, \ldots, d_n)$
  - ➤**Solución: teorema de Bayes**

MÉTODO PROBABILISTA CLÁSICO (cont.)

$$P(d_1,\ldots,d_n \mid e_1,\ldots,e_m) = \frac{P(e_1,\ldots,e_m \mid d_1,\ldots,d_n)\cdot P(d_1,\ldots,d_n)}{\sum_{d'_1,\ldots,d'_n} P(e_1,\ldots,e_m \mid d'_1,\ldots,d'_n)\cdot P(d'_1,\ldots,d'_n)}$$

H1: diagnósticos exclusivos y exhaustivos



$$P(d_i \mid e_1,\ldots,e_m) = \frac{P(e_1,\ldots,e_m \mid d_i)\cdot P(d_i)}{\sum_j P(e_1,\ldots,e_m \mid d_j)\cdot P(d_j)}$$

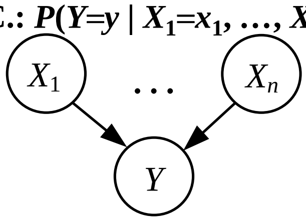
H2: independencia condicional

$$P(d_i \mid e_1,\ldots,e_m) = \frac{P(e_1 \mid d_i)\cdot\ldots\cdot P(e_m \mid d_i)\cdot P(d_i)}{\sum_j P(e_1 \mid d_j)\cdot\ldots\cdot P(e_m \mid d_j)\cdot P(d_j)}$$

- **H1 y H2 no son válidas en casos reales ⇒ MYCIN ⇒ redes bayesianas (Judea Pearl, años 80)**

REDES BAYESIANAS

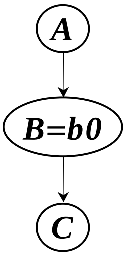
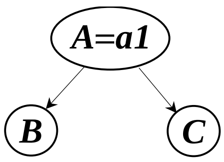
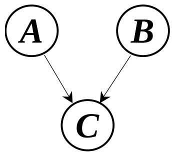
- **Red bayesiana: G.D.A. + separación direccional**
- **Nodo: variable aleatoria con valores exclusivos y exhaustivos**
- **Enlace → T.P.C.: $P(Y{=}y \mid X_1{=}x_1, \ldots, X_n{=}x_n)$**



- **Probabilidad conjunta: $P(n_1,\ldots,n_m)=\prod_i P(n_i \mid pa(n_i))$**
- **Probabilidad a posteriori: $P(n_i \mid e)$ / $e=\cup\{N_j=n_j\}_{j\in\{1,\ldots,m\}-\{i\}}$**
- **Modelos probabilísticos de interacción causal: General, <u>O</u>, Y, <u>MAX</u>, MIN, SUM…**

PROPIEDAD DE SEPARACIÓN DIRECCIONAL

- **S.D.: Refleja la noción humana de causalidad**

Independencia condicional entre hijos y padres de un nodo, dado su valor	**Independencia condicional entre efectos, dado el valor de la causa**	**Independencia a priori de los nodos sin antepasados comunes**
		
***P*(*c* | *a*, *b*0)=*P*(*c* | *b*0)**	***P*(*b* | *a*1, *c*)=*P*(*b* | *a*1)**	***P*(*a* | *b*)=*P*(*a*)**

- **Si estas relaciones no fueran ciertas en un dominio particular, habría que añadir arcos**

PROPAGACIÓN DE LA EVIDENCIA

- **A partir de la probabilidad conjunta: complejidad exponencial con el número de nodos**

$$P(x \mid e) = \frac{P(x,e)}{P(e)} = \alpha \cdot \sum_{x_i \setminus \{x,e\}} P(x_1, \ldots, x_n)$$

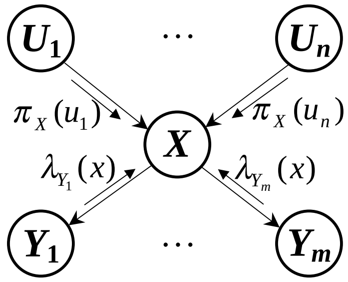
- **En poliárboles,**

$$P(x \mid e) = \frac{P(x,e)}{P(e)} = \alpha \cdot P(x, e_X^+, e_X^-) = \alpha \cdot P(x, e_X^+) \cdot P(e_X^- \mid x, e_X^+)$$

$$= \alpha \cdot P(x, e_X^+) \cdot P(e_X^- \mid x) = \alpha \cdot \pi(x) \cdot \lambda(x)$$

PROPAGACIÓN DE LA EVIDENCIA (cont.)

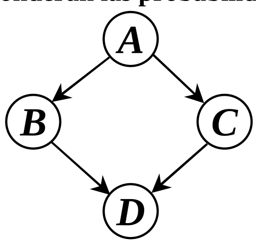
- **Existen fórmulas de carácter recursivo para el cálculo de π y λ en cada nodo.**
- **Dichas fórmulas se prestan a una implementación distribuida**



- **Complejidad: proporcional al camino más largo de la red**

PROPAGACIÓN DE LA EVIDENCIA (cont.)

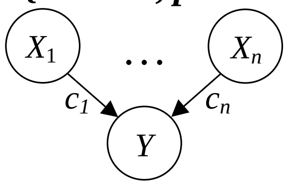
- ***Método de Agrupamiento*: se agrupan variables en nodos de mayor tamaño para que desaparezcan los bucles → árbol de cliques**
- ***Método de Condicionamiento*: se asignan valores a ciertas variables para romper los bucles, se propaga la evidencia para cada posible distribución de dichos valores y se ponderan las probabilidades obtenidas para cada distribución**



- ***Método de Simulación*: se estiman las probabilidades a partir de una serie de ensayos**
  - ➤ **Ventaja: es independiente de la topología de la red**
  - ➤ **Inconveniente: Ensayos inútiles (*Logic Sampling*) → Nuevos métodos**

PUERTA O PROBABILISTA (*NOISY OR-GATE*)

- **Variables binarias → {*ausente, presente*}**
- **X_i: causa *i*, *Y*: efecto**

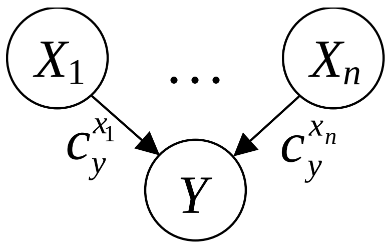


- $\boldsymbol{c_i \equiv P(Y{=}{+}y \mid X_1{=}\neg x_1, \ldots, X_i{=}{+}x_i, \ldots, X_n{=}\neg x_n)}$
- **T.P.C.:** $P(+y \mid \bar{x}) = 1 - \prod_{i \in T_X} (1 - c_i)$

$P(+y \mid x_1, x_2)$	$\neg x_1$	$+x_1$
$\neg x_2$	**0**	c_1
$+x_2$	c_2	$1-(1-c_1)(1-c_2)$

PUERTA MAX PROBABILISTA (*NOISY MAX-GATE)*

- **Variables graduadas → {*ausente* < *grado*1 < … < *grado*g}**
  - ➤**Fiebre → {*ausente*(0) < *leve*(1) < *moderada*(2) < *severa*(3)}**
- **X_i: causa *i*, *Y*: efecto**



- $c_y^{x_i} \equiv P(Y = y \mid X_1 = 0, \ldots, X_i = x_i, \ldots, X_n = 0)$

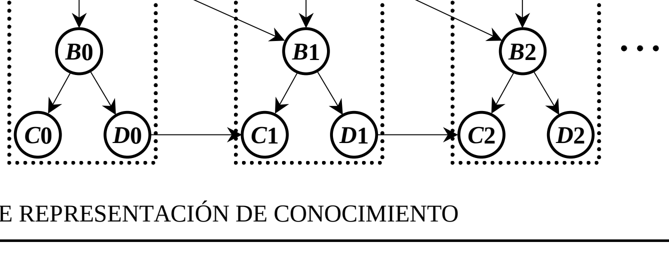
PUERTA MAX PROBABILISTA (cont.)

- T.P.C.: $P(y \mid \bar{x}) = \sum_{\bar{z} \mid max\ \bar{z}=y} \prod_i c_{z_i}^{x_i}$

$P(y \mid X_1{=}x_1, X_2{=}x_2)$	$c_0^{x_1}$	$c_1^{x_1}$	$\cdots$	$c_{g_Y-1}^{x_1}$	$c_{g_Y}^{x_1}$
$c_0^{x_2}$	Y=0	Y=1	$\cdots$	Y=g$_Y$–1	Y=g$_Y$
$c_1^{x_2}$	Y=1	Y=1	$\cdots$	Y=g$_Y$–1	Y=g$_Y$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$c_{g_Y-1}^{x_2}$	Y=g$_Y$–1	Y=g$_Y$–1	$\cdots$	Y=g$_Y$–1	Y=g$_Y$
$c_{g_Y}^{x_2}$	Y=g$_Y$	Y=g$_Y$	$\cdots$	Y=g$_Y$	Y=g$_Y$

REDES BAYESIANAS DINÁMICAS

- **Tiempo discreto**
- ***Nodos*: se crea una copia de cada variable para cada instante de tiempo**
- ***Enlaces*: se reproduce un modelo causal estático para cada instante y se establecen arcos entre variables de modelos consecutivos (procesos markovianos)**

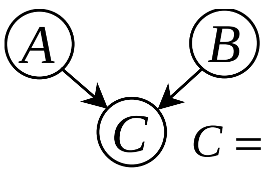


REDES BAYESIANAS DINÁMICAS (cont.)

- **Inconvenientes para nuestro dominio:**
  - ➤ **Se obtienen redes muy complejas estructuralmente**
  - ➤ **Cada variable debería representar un evento irreversible (sólo ocurre una vez a lo largo del tiempo) y no un evento que puede producirse múltiples veces**
  - ➤ **Nuestro dominio no es markoviano ya que puede haber diferentes retardos entre dos nodos evento**

REDES BAYESIANAS CON NODOS TEMPORALES

- **Tiempo dividido en cada nodo en intervalos de diferente duración**
- ***Nodos*: representan un evento o cambio de estado de una variable (como máximo un cambio a lo largo del tiempo)**
  - ➤ **Valor de un nodo: intervalo en que ocurre el evento**
  - ➤ **Cada intervalo está definido relativamente a la ocurrencia de un evento padre**

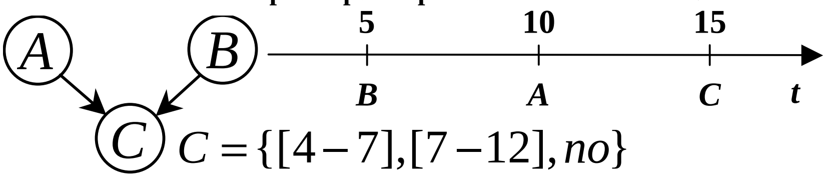


$$C = \{[0-20],[20-50],no\}$$

- ***Enlaces*: representan relaciones causales temporales entre nodos**

REDES BAYESIANAS CON NODOS TEMPORALES (cont.)

- **Inconvenientes:**
  - ➤ **No utiliza modelos canónicos probabilistas adaptados a procesos temporales → siempre el modelo general**
  - ➤ **La ocurrencia en un instante determinado de un nodo evento con padres no constituye evidencia directa → escenarios**
  - ➤ **La definición de los valores de un nodo mediante tiempos relativos a cada uno de sus nodos padre puede producir inconsistencias:**



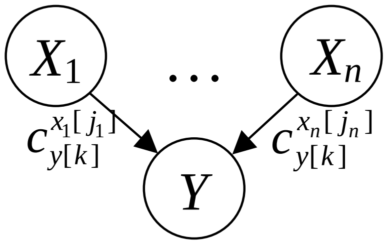
$$C = \{[4-7],[7-12],no\}$$

$$\Downarrow$$

valores no exclusivos

PUERTA O PROBABILISTA TEMPORAL (*TEMPORAL NOISY OR-GATE*)

- **Variable *V*: evento temporal → {*v*[0], *v*[1], …, *v*[t_V], *v*[*nunca*]}**
- ***V*=*v*[*i*] significa "El evento *V* tiene lugar en el instante *i*."**
- **X_i: causa *i*, *Y*: efecto**



- $c_{y[k]}^{x_i[j_i]} \equiv P(y[k] \mid x_1[nunca], \ldots, x_i[j_i], \ldots, x_n[nunca])$

$$j_i \in \{0, \ldots, t_{X_i}\},\ k \in \{0, \ldots, t_Y\}$$

PUERTA O PROBABILISTA TEMPORAL (cont.)

- **TPC: Puerta O Probabilista Temporal $\rightarrow$ Puerta MAX probabilista ($\boldsymbol{nunca < t_Y < t_Y - 1 < \ldots < 1 < 0}$)**

$P(y[k] \mid x_1[j_1], x_2[j_2])$	$c_{y[nunca]}^{x_1[j_1]}$	$c_{y[t_Y]}^{x_1[j_1]}$	$\cdots$	$c_{y[1]}^{x_1[j_1]}$	$c_{y[0]}^{x_1[j_1]}$
$c_{y[nunca]}^{x_2[j_2]}$	$y[nunca]$	$y[t_Y]$	$\cdots$	$y[1]$	$y[0]$
$c_{y[t_Y]}^{x_2[j_2]}$	$y[t_Y]$	$y[t_Y]$	$\cdots$	$y[1]$	$y[0]$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$c_{y[1]}^{x_2[j_2]}$	$y[1]$	$y[1]$	$\cdots$	$y[1]$	$y[0]$
$c_{y[0]}^{x_2[j_2]}$	$y[0]$	$y[0]$	$\cdots$	$y[0]$	$y[0]$

RED PARA MODELADO DE FALLOS EN UN GENERADOR DE VAPOR

- **Horizonte o rango temporal**: 12 minutos
- **Unidad temporal**: 20 segundos
  - ➤ $FAT = fat[3]$ significa “incremento en el flujo de agua de atemperación producido entre los segundos 41 y 60”
- **Parámetros independientes por arco**:

$$c_{y[k]}^{x_i[j_i]} \quad j_i \in \{1, \ldots, 36, nunca\},\ k \in \{1, \ldots, 36\}$$

$$(36+1)\cdot 36 = 1332 \longrightarrow 1296 \longrightarrow 666 \longrightarrow 36$$

$$c_{y[k]}^{x_i[nunca]} = 0 \qquad c_{y[k]}^{x_i[j_i > k]} = 0 \qquad c_{y[k+\Delta t]}^{x_i[j_i+\Delta t]} = c_{y[k]}^{x_i[j_i]}$$

ADQUISICIÓN DE LOS PARÁMETROS DE LA RED

- **Obtener $c_{y[k]}^{x_i[j_i]}$ requeriría averiguar:**
  - **Dado que X_i tiene lugar en un cierto <u>intervalo</u> de 20 sgs., ¿cuál es la probabilidad de que su efecto Y ocurra en ese mismo intervalo, si el resto de sus causas están ausentes?, ¿y la probabilidad de que Y ocurra en el siguiente intervalo? y así sucesivamente.**
- **Disponíamos de los parámetros $\tilde{c}_Y^{X_i}(\Delta t)$:**
  - **Dado que X_i tiene lugar en un cierto <u>instante</u>, ¿cuál es la probabilidad de que su efecto Y ocurra en el siguiente intervalo de 20 sgs., si el resto de sus causas están ausentes?, ¿y la probabilidad de que Y ocurra en el intervalo siguiente al siguiente? y así sucesivamente.**

$$c_{y[j_i+\Delta t]}^{x_i[j_i]} = \frac{\tilde{c}_Y^{X_i}(\Delta t) + \tilde{c}_Y^{X_i}(\Delta t + 1)}{2}$$

IMPLEMENTACIÓN

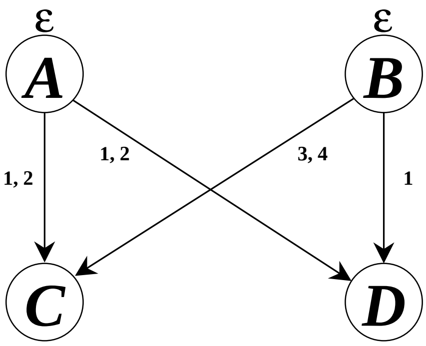
- Entorno de desarrollo utilizado: Elvira (creado conjuntamente por varias universidades españolas)

http://www.ia.uned.es/~fjdiez/bayes/elvira/

- Tiempo empleado en inferencia: por debajo del minuto a partir de métodos exactos

EJEMPLO

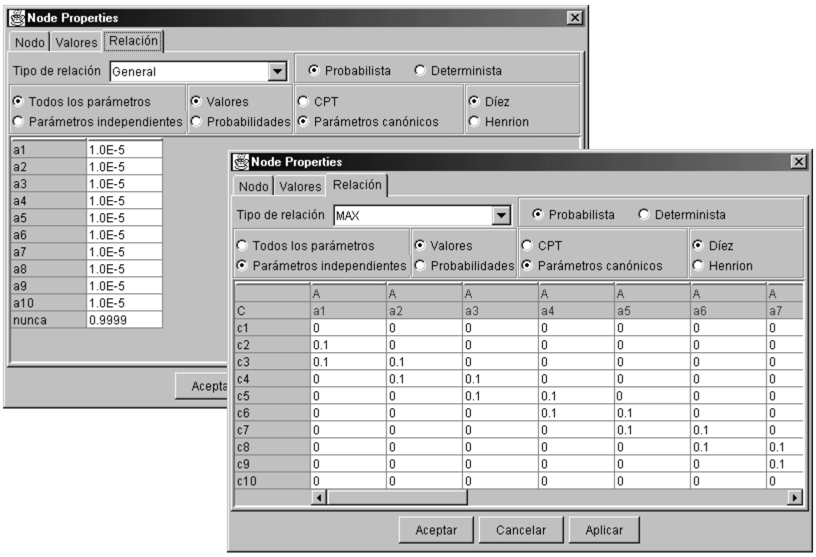
RANGO TEMPORAL: {1, …,10}



***c*[6] ⇒ *a*[4], *a*[5], *b*[2], *b*[3]**

***c*[6], *d*[7] ⇒ *a*[5]**

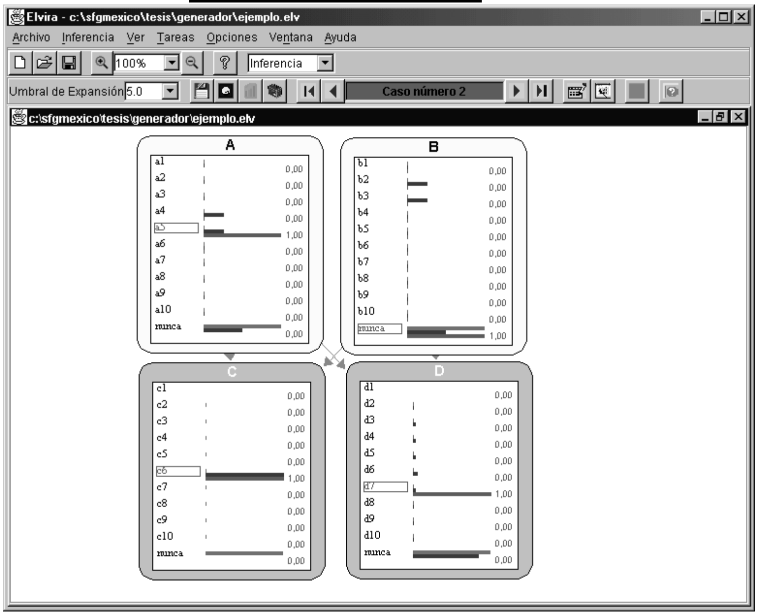
EJEMPLO (cont.)



a1	1.0E-5
a2	1.0E-5
a3	1.0E-5
a4	1.0E-5
a5	1.0E-5
a6	1.0E-5
a7	1.0E-5
a8	1.0E-5
a9	1.0E-5
a10	1.0E-5
nunca	0.9999

	A	A	A	A	A	A	A
C	a1	a2	a3	a4	a5	a6	a7
c1	0	0	0	0	0	0	0
c2	0.1	0	0	0	0	0	0
c3	0.1	0.1	0	0	0	0	0
c4	0	0.1	0.1	0	0	0	0
c5	0	0	0.1	0.1	0	0	0
c6	0	0	0	0.1	0.1	0	0
c7	0	0	0	0	0.1	0.1	0
c8	0	0	0	0	0	0.1	0.1
c9	0	0	0	0	0	0	0.1
c10	0	0	0	0	0	0	0

EJEMPLO (cont.)



CONCLUSIONES

RED BAYESIANA $\xrightarrow{\textit{Causalidad, Incertidumbre}}$ **MODELO DE PROPAGACIÓN DE FALLOS EN UN GENERADOR DE VAPOR**

Tiempo

PUERTA O PROBABILISTA TEMPORAL