

Restricciones Para El Aprendizaje de Redes Bayesianas

Luis M. de Campos Ibañez
Fco. Javier García Castellano

Uncertainty Treatment in Artificial Intelligence Research Group
Department of Computer Science and Artificial Intelligence

Granada University (Spain)



Objetivo



Permitir que los algoritmos de aprendizaje de redes bayesianas incluidos en Elvira pueda utilizar restricciones/conocimiento previo.



Esto Implica:



Estamos trabajando con restricciones duras: La información dada por las restricciones prevalecerá.



La red obtenida debe satisfacer las restricciones.



Los elementos del espacio de búsqueda considerado, en gran medida deben satisfacer las restricciones.

Objetivo

Tipos de Restricciones

- Restricciones de existencia de arcos y/o aristas:

$$\exists x \rightarrow y \text{ o } \exists x \text{---}y$$


- Restricciones de ausencia de arcos y/o aristas:

$$\nexists x \rightarrow y \text{ o } \nexists x \text{---}y$$

- Restricciones de orden parcial (relaciones de precedencia)

$$x < y$$

Se asume la transitividad de esta relación, para que no

 sea necesario especificar toda la relación (si $x < y$ e $y < z$ entonces asumimos que también $x < z$)

Tipos de Restricciones

Representación de las Restricciones

- **Existencia**: Como un elemento de la clase **Graph**. Nunca en una Bnet, puesto que en ésta última no puede haber aristas.
- **Ausencia**: Como un elemento de la clase **Graph**. Nunca en una Bnet, puesto que en ésta última no puede haber aristas y ciclos dirigidos.
- **Orden Parcial**: Como un elemento de la clase **Graph** (en este caso podría ser una Bnet).
Donde $x < y$ se almacena como que existe el arco $x \rightarrow y$ en ese grafo.

Representación de las Restricciones

Verificación de las restricciones

- Suponemos que queremos comprobar si cumple o no las restricciones la red G :

-  Restricciones de existencia G_e

- Si $x \rightarrow y \in G_e$ entonces $x \rightarrow y \in G$
- Si $x \dashrightarrow y \in G_e$ entonces $x \rightarrow y \in G$ o $x \leftarrow y \in G$

-  Restricciones de ausencia G_a

- Si $x \rightarrow y \in G_a$ entonces $x \rightarrow y \notin G$
- Si $x \dashrightarrow y \in G_a$ entonces $x \rightarrow y \notin G$ y $x \leftarrow y \notin G$

-  Restricciones de orden G_o

- $G_o \cup G$ es un dag (no tiene ciclos dirigidos)

Verificación de las Restricciones

Autoconsistencia entre restricciones

Sea G_e , G_a y G_o los grafos que representan las restricciones de existencia, ausencia y orden parcial, respectivamente. Entonces:

- El conjunto de restricciones de existencia es autoconsistente si y sólo si el grafo G_e no tiene ciclos dirigidos.
- El conjunto de restricciones de ausencia siempre es autoconsistente.
- El conjunto de restricciones de orden parcial es autoconsistente si el grafo G_o es un DAG.

Autoconsistencia entre Restricciones

Consistencia entre restricciones

Sea G_e , G_a y G_o los grafos que representan las restricciones de existencia, ausencia y orden parcial, respectivamente. Sea el grafo G_{re} , el que se define como:

$$\{x \rightarrow y \mid x \rightarrow y \in G_e\} \cup \{y \rightarrow x \mid x \rightarrow y \in G_e, x \rightarrow y \in G_a\} \cup \\ \{x \dashrightarrow y \mid x \dashrightarrow y \in G_e, x \rightarrow y \notin G_a, y \rightarrow x \notin G_a\}$$

Entonces las restricciones son consistentes entre si, si y sólo si:

$G_{re} \cap G_a = G_\emptyset$ y $G_{re} \cup G_o$ no tiene ciclos dirigidos

Usamos la convección: $\{x \rightarrow y\} \cup \{x \dashrightarrow y\} = \{x \rightarrow y\}$ y $\{x \rightarrow y\} \cap \{x \dashrightarrow y\} = \{x \rightarrow y\}$

Además, un grafo G es consistente con restricciones (que son consistentes entre si), si y solo si:

$G_{re} \cup G = G$, $G \cap G_a = G_\emptyset$ y $G \cup G_o$ no tiene ciclos dirigidos

Consistencia entre Restricciones

Operadores para restricciones

Sean G_e , G_a y G_o los grafos que representan las restricciones de existencia, ausencia y orden parcial, respectivamente. Sea G un grafo consistente con las restricciones.

• **Inserción**: $G' = G \cup \{x \rightarrow y\}$ con $x \rightarrow y \notin G$. Entonces, G' es consistente con las restricciones, si y sólo si:

$$x \rightarrow y \notin G_a \text{ and } x \dashrightarrow y \notin G_a$$

No hay un camino dirigido de y a x en $G \cup G_o$

• **Borrado**: $G' = G \setminus \{x \rightarrow y\}$ con $x \rightarrow y \in G$. Entonces, G' es consistente con las restricciones, si y sólo si:

$$x \rightarrow y \notin G_e \text{ and } x \dashrightarrow y \notin G_e$$

• **Inversión**: $G' = (G \setminus \{x \rightarrow y\}) \cup \{y \rightarrow x\}$ con $x \rightarrow y \in G$. Entonces, G' es consistente con las restricciones, si y sólo si:

$$x \rightarrow y \notin G_e, y \rightarrow x \notin G_a \text{ y } x \dashrightarrow y \notin G_o$$

No hay un camino dirigido de x a y en $(G \setminus \{x \rightarrow y\}) \cup G_o$

Operadores para Restricciones

Experimentos



Varias Redes: Asia, Alarm, Insurance, Hailfinder.



Score+search: Búsqueda Local, métrica Bde.



Medidas estructurales (usando *essential graphs*): número de arcos añadidos (A), número de arcos borrados (D), número de arcos invertidos (I).



Proporción de tiempos (T) respecto al algoritmo sin restricciones.



Porcentajes de restricciones: 10%, 20%, 30%, 40%

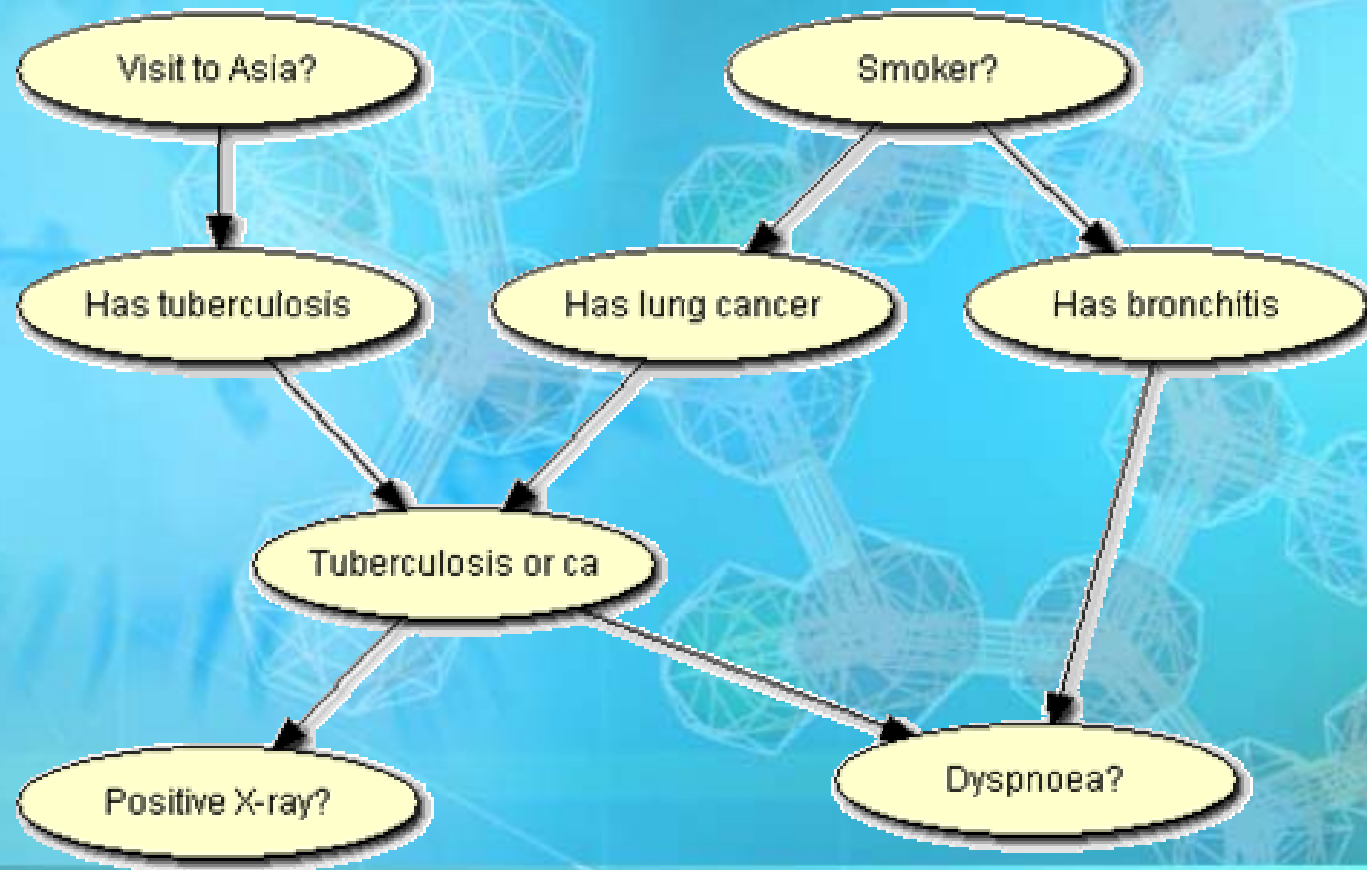


Valores medios 50 ejecuciones (10000 casos) ó 40 ejecuciones (500 casos)

Experimentos

Experimentos: Red Asia

- i** Probabilidad de que un paciente tenga tuberculosis, cáncer de pulmón o bronquitis.
- i** 8 variables, 8 arcos, 10000 casos.

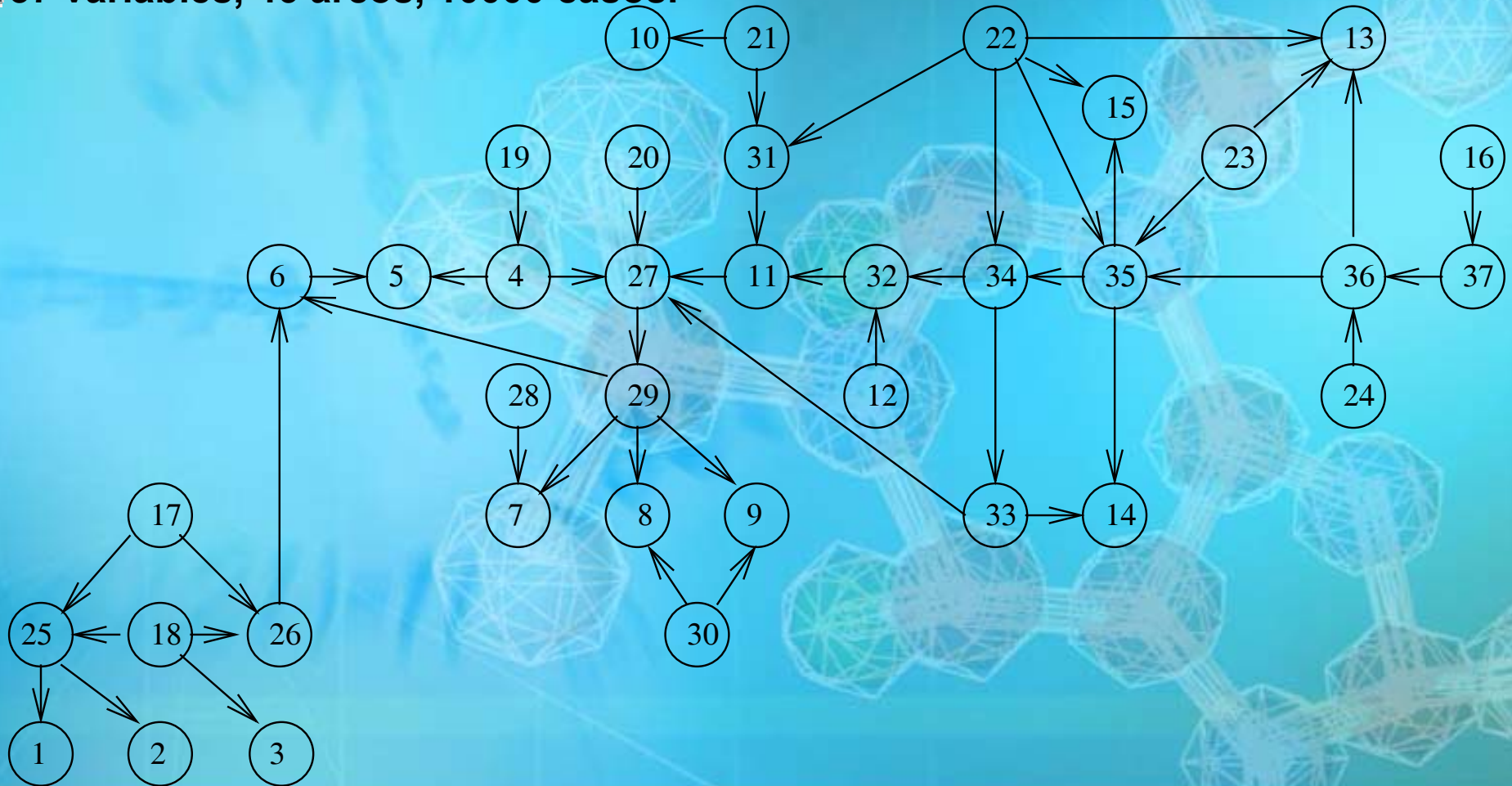


Experimentos: Red Asia

Experimentos: Red Alarm

i Alarm Monitoring System.

i 37 variables, 46 arcsos, 10000 casos.



Experimentos: Red Alarm

Experimentos: Red Insurance

1 Evaluar riesgos en seguros de coches.

1 27 variables, 52 arcos, 10000 casos.



Experimentos: Red Insurance

Experimentos: Red Hailfinder

1 Predicción de granizo severo en verano para la región noreste de Colorado.

1 56 variables, 66 arcos, 10000 casos.



Experimentos: Red Hailfinder

Experimentos: Resultados Asia

	G_e, G_a, G_o					only G_e					only G_a					only G_o				
%	BDeu	A	D	I	T	BDeu	A	D	I	T	BDeu	A	D	I	T	BDeu	A	D	I	T
10%	-2258.48	1.8	0.9	1.3	76	-2257.42	1.8	0.8	2.4	76	-2260.59	1.8	1.0	2.2	76	-2257.65	1.9	1.0	2.4	77
20%	-2256.95	1.5	0.8	0.2	56	-2256.94	1.8	0.8	1.6	57	-2260.34	1.6	1.1	1.1	56	-2257.37	1.9	1.1	1.9	58
30%	-2256.71	1.1	0.5	0.0	43	-2256.69	1.9	0.5	0.8	44	-2258.96	1.4	1.1	0.7	43	-2256.76	2.0	1.1	1.0	42
40%	-2256.87	0.7	0.5	0.0	28	-2256.61	1.9	0.5	0.4	29	-2260.00	0.9	1.0	0.4	28	-2256.59	2.0	1.1	0.8	30
0%	-2257.90	2	1	3		running time: 0.51 sec.					BDeu true network: -2257.55									

Table 1. Results obtained for Asia.

- El número de arcos borrados, añadidos e invertidos decrece según el número de restricciones de existencia, ausencia y de orden parcial, respectivamente, crece.
- El uso de un tipo particular de restricciones NO tiende a disminuir las diferencias estructurales asociados a los otros tipos de restricciones
- Las redes encontradas tienen peor score que la red óptima
- Tarda menos tiempo

Experimentos: Resultados Asia

Experimentos: Resultados Alarm

%	G_e, G_a, G_o					only G_e					only G_a					only G_o				
	BDeu	A	D	I	T	BDeu	A	D	I	T	BDeu	A	D	I	T	BDeu	A	D	I	T
10%	-108551	1.6	1.2	1.4	77	-108666	3.0	1.5	2.0	78	-108773	4.1	1.7	2.4	77	-108758	4.1	1.7	3.1	78
20%	-108550	0.9	1.0	1.0	60	-108613	2.4	1.4	2.3	61	-108806	3.4	1.7	2.5	60	-108739	3.7	1.3	2.5	61
30%	-108513	0.3	0.8	0.4	46	-108562	1.8	1.0	1.9	47	-108788	2.6	1.5	2.3	46	-108718	3.3	1.1	2.1	47
40%	-108504	0.2	0.7	0.3	33	-108486	0.9	0.8	2.0	35	-108782	1.8	1.3	1.8	34	-108675	2.8	1.1	1.9	35
0%	-108828	5	2	3		running time: 2.53 min.					BDeu true network: -108452									

Table 2. Results obtained for Alarm.

- El número de arcos borrados, añadidos e invertidos decrece según el número de restricciones de existencia, ausencia y de orden parcial, respectivamente, crece.
- El uso de un tipo particular de restricciones tiende a disminuir las diferencias estructurales asociados a los otros tipos de restricciones
- Las redes encontradas tienen peor score que la red óptima
- Tarda menos tiempo

Experimentos: Resultados Alarm

Experimentos: Resultados Insurance

%	G_e, G_a, G_o					only G_e					only G_a					only G_o				
	BDeu	A	D	I	T	BDeu	A	D	I	T	BDeu	A	D	I	T	BDeu	A	D	I	T
10%	-132369	3.1	8.7	9.0	79	-132406	4.2	8.8	10.7	79	-132410	4.6	9.7	10.6	79	-132417	5.3	10.0	10.5	79
20%	-132271	1.4	7.4	6.4	58	-132429	3.4	8.0	8.5	59	-132434	3.0	9.3	8.5	59	-132362	4.6	9.8	9.4	59
30%	-132225	0.5	6.0	4.4	42	-132313	1.8	6.4	6.8	43	-132509	2.5	9.3	8.4	43	-132309	3.8	9.8	8.4	44
40%	-132233	0.2	5.1	3.8	31	-132409	1.6	5.7	5.2	32	-132308	1.4	8.9	6.4	31	-132257	3.3	9.5	7.5	33
0%	-132488	6	10	11		running time: 1.60 min.					BDeu true network: -132512									

Table 3. Results obtained for Insurance.

- El número de arcos borrados, añadidos e invertidos decrece según el número de restricciones de existencia, ausencia y de orden parcial, respectivamente, crece.
- El uso de un tipo particular de restricciones tiende a disminuir las diferencias estructurales asociados a los otros tipos de restricciones
- Las redes encontradas tienen mejor score que la red óptima
- Tarda menos tiempo

Experimentos: Resultados Insurance

Experimentos: Resultados Hailfinder

%	G_e, G_a, G_o					only G_e					only G_a					only G_o				
	BDeu	A	D	I	T	BDeu	A	D	I	T	BDeu	A	D	I	T	BDeu	A	D	I	T
10%	-498306	13.1	10.0	11.6	80	-497977	14.8	10.4	15.2	80	-498146	16.1	11.5	17.3	80	-498245	17.4	12.7	15.8	81
20%	-498354	10.1	8.4	4.7	64	-498098	13.2	9.1	8.7	64	-498475	14.2	11.2	14.3	64	-498444	17.1	12.7	12.5	65
30%	-498424	7.0	6.6	3.2	51	-498219	12.1	8.0	5.5	52	-498535	11.2	10.4	12.4	51	-498726	16.8	12.6	9.5	53
40%	-498550	5.0	5.4	1.0	41	-498347	11.0	6.9	3.7	42	-498516	8.5	9.3	8.9	41	-498722	15.9	12.5	7.5	43
0%	-497904	17	12	19		running time: 6.53 min.					BDeu true network: -503095									

Table 4. Results obtained for Hailfinder.

- El número de arcos borrados, añadidos e invertidos decrece según el número de restricciones de existencia, ausencia y de orden parcial, respectivamente, crece.
- El uso de un tipo particular de restricciones tiende a disminuir las diferencias estructurales asociados a los otros tipos de restricciones
- Las redes encontradas tienen mejor score que la red óptima
- Tarda menos tiempo

Experimentos: Resultados Hailfinder

Experimentos: Resultados 500 casos

%	G_e, G_a, G_o				only G_e				only G_a				only G_o			
	BDeu	A	D	I	BDeu	A	D	I	BDeu	A	D	I	BDeu	A	D	I
Asia																
10%	-1075.94	0.0	0.8	0.6	-1075.94	0.0	0.8	0.6	-1075.36	0.0	1.0	0.0	-1075.36	0.0	1.0	0.0
20%	-1075.74	0.0	0.6	0.3	-1075.87	0.0	0.6	0.4	-1075.36	0.0	1.0	0.0	-1075.36	0.0	1.0	0.0
30%	-1075.75	0.0	0.6	0.3	-1075.75	0.0	0.6	0.3	-1075.36	0.0	1.0	0.0	-1075.36	0.0	1.0	0.0
40%	-1075.51	0.0	0.6	0.0	-1075.51	0.0	0.6	0.0	-1075.36	0.0	1.0	0.0	-1075.36	0.0	1.0	0.0
0%	-1075.36	0	1	0	BDeu true network: -1075.69											
Alarm																
10%	-5990	9.7	4.0	12.7	-5972	10.4	3.8	15.6	-5998	11.0	5.0	19.3	-6002	13.2	4.7	14.6
20%	-5971	6.2	3.3	7.6	-5959	9.1	3.2	12.1	-5990	9.2	4.6	14.2	-6005	12.4	5.0	12.2
30%	-5946	4.4	2.2	5.8	-5949	7.7	2.6	10.2	-5984	7.3	4.6	10.8	-6003	11.3	4.9	10.2
40%	-5943	3.2	1.6	4.4	-5946	7.1	2.0	8.4	-5989	6.5	4.6	8.8	-5986	9.8	4.6	8.5
0%	-5986	11	5	22	BDeu true network: -5935											
Insurance																
10%	-7262	4.8	17.6	8.8	-7270	5.6	18.2	9.0	-7274	7.2	20.6	7.1	-7270	8.0	20.9	7.8
20%	-7280	3.4	15.2	7.8	-7286	4.6	16.2	8.7	-7270	5.9	19.6	7.4	-7270	7.9	20.7	7.4
30%	-7310	2.0	12.2	6.0	-7325	3.6	13.4	8.0	-7264	4.6	18.7	6.2	-7268	7.8	20.6	7.2
40%	-7353	1.4	9.9	6.2	-7358	3.1	11.2	6.8	-7273	3.9	18.0	5.8	-7265	7.8	20.4	6.9
0%	-7270	8	21	8	BDeu true network: -7592											
Hailfinder																
10%	-27270	13.3	22.2	11.1	-27231	15.4	22.5	13.0	-27202	14.7	23.8	11.4	-27183	16.8	25.2	11.3
20%	-27408	11.0	19.9	9.2	-27330	14.2	19.9	12.1	-27244	12.5	23.3	10.4	-27190	16.4	25.3	10.7
30%	-27582	8.7	17.2	8.0	-27461	13.3	17.4	12.2	-27280	10.3	22.2	9.8	-27198	16.3	25.6	10.2
40%	-27781	7.0	14.2	7.6	-27649	12.4	14.6	11.4	-27309	8.0	21.2	9.2	-27204	16.2	26.0	9.6
0%	-27171	17	25	12	BDeu true network: -29347											

Table 5. Results obtained using data sets with only 500 cases.

Las redes aprendidas tienen mejor score que la óptima (excepto en Alarm)

Experimentos: Resultados 500 casos

Conclusiones

Se han definido tres tipos de restricciones (duras) para redes bayesianas:

Existencia

Ausencia

Orden Parcial

Se ha visto que el uso de información adicional en forma de restricciones mejora la estructura de la red, realizando el aprendizaje en menos tiempo

Estamos trabajando en métodos score+search que no busquen directamente en DAGs (RPDAGs) o algoritmos que utilicen tests de independiencia (PCLearning)

Conclusiones