

# MÉTODO PARA CONSTRUIR REDES BAYESIANAS

*María del Carmen Chávez C.\**

*Ricardo Grau Á.\**

*María Matilde García L.\**

## RESUMEN

En el presente trabajo se modelan las relaciones probabilísticas usando las redes bayesianas. La bibliografía consultada muestra que la tendencia actual de las investigaciones en esta técnica se orienta a utilizarlas unidas a otras disciplinas y que también es importante la forma en que se modela una red bayesiana. Se propone un método que utiliza la técnica de segmentación estadística, implementada en el paquete de programas CHAID (Chi-Squared Automatic Interaction Detector) y métodos estadísticos incorporados al SPSS (Statistical Package for Social Science) para obtener modelos de redes bayesianas. Se aplicó el método a un problema de epidemiología para caracterizar los distintos factores de la cardiopatía isquémica, específicamente el infarto de miocardio agudo (IMA). También se aplicó para determinar la influencia de distintas alteraciones orofaciales en la aparición de la retinosis pigmentaria. El método que se propone permite que los expertos en el tema participen en la elección de la mejor topología para la red bayesiana, entre varias alternativas. Para realizar la inferencia en estas redes, se implementó una primera versión del algoritmo de propagación en redes con una estructura de poliárbol.

## ABSTRACT

In this article some probabilistic relations are modeled by using bayesian nets. The consulted literature shows that the present tendencies of this technique are guided to use them joined to other disciplines. It is also important the way in which the bayesian net is modeled.

A method which uses the technique of statistic segmentation is proposed and implemented by using CHAID (Chi-Squared Automatic Interaction Detector). Some statistical methods were joined to SPSS (Statistical Package for Social Science) in order to obtain bayesian nets models.

The method was applied in the field of Epidemiology to characterize the different factors associated with Ischemic Cardiopathy, in particular heart attack. It was applied to determine the influence of different alterations orofacial in the development of Pigmentary Retinosis. The proposed method allows expert in the field of applications, to cooperate in the selection of the best topology for the bayesian network among several alterations.

A first version of the algorithm of propagation in nets with a polytree structure was implemented to make the inference in these nets.

---

\* Grupo de Inteligencia Artificial. Centro de Estudios de Informática, Universidad Central de Las Villas. Santa Clara. Villa Clara. Cuba.

## INTRODUCCIÓN

Las probabilidades condicionadas son útiles para describir las relaciones de dependencia entre variables. Los sistemas basados en el conocimiento que usan la probabilidad se conocen como sistemas probabilísticos, pues su estrategia de razonamiento es la inferencia probabilística. Por ejemplo, el sistema experto PROSPECTOR [DUD80] usa el teorema de Bayes para aconsejar en la exploración de minas. En ocasiones se usa la probabilidad para el cálculo de los valores de certidumbre en sistemas expertos; por ejemplo, el sistema MYCIN [SHO84] en el encadenamiento de reglas tiene en cuenta un análisis de la certidumbre basado en los factores de certeza. En [CAS89] se comienza la implementación de un sistema basado en probabilidades para el caso en el que todas las variables se consideran independientes entre sí.

Otra tendencia ha sido la de emplear redes bayesianas para representar información probabilística. Aunque desde el año 1961 se investiga en el empleo de la inferencia bayesiana en estas redes, no es hasta 1983 cuando se generaliza un método para redes con una sola conexión. Donde por primera vez se utilizan las redes de creencia es en CONVINCER [KIM83]. En la herramienta HUGIN [SPI93][JEN90] se implementan por primera vez métodos de inferencia para redes múltiplemente conectadas. Sistemas más recientes que usan esta tecnología son MUNIN [AND89] y PATHFINDER [HEC91].

Los trabajos revisados muestran que estas técnicas tienden a ser más utilizadas, pero resulta difícil o muy dependiente de los expertos la construcción de un modelo de red bayesiana. Por ello, es conveniente tener una técnica que, a partir de datos históricos, permita sustituir total o parcialmente al experto en la confección de la red y evaluar las probabilidades condicionales.

En el trabajo se muestra un algoritmo para la confección de redes probabilísticas, en particular la determinación automática de la estructura de la red y el cálculo de las probabilidades necesarias a partir de una base de datos previa.

Se aplica el método en un problema de epidemiología para caracterizar los distintos factores de la cardiopatía izquémica, específicamente el infarto de miocardio agudo (IMA), y se hace una aplicación al diagnóstico para conocer cómo distintas alteraciones orofaciales influyen en la aparición de la retinosis pigmentaria.

### Redes bayesianas

Este tipo de redes se define como un grafo que cumple lo siguiente:

- Los nodos de la red están formados por un conjunto de variables aleatorias que se denotan con la letra  $X$  o con subíndices  $X_1, X_2, \dots, X_n$ . En principio, estas variables pueden ser rasgos, pero puede ocurrir también que un mismo rasgo tenga que ser descompuesto en varias variables aleatorias. Por ejemplo, si el rasgo es multivaluado y se desean las variables aleatorias dicotómicas.
- Cada par de nodos se conecta entre sí mediante un conjunto de enlaces o flechas. El significado de una flecha que vaya del nodo  $X$  al nodo  $Y$  es el de que  $X$  ejerce una influencia directa sobre  $Y$ ; en términos de probabilidades esto significa que hay una dependencia condicional de  $Y$  con respecto a  $X$ , y que la probabilidad de  $Y$  es diferente de la probabilidad de  $Y$  dado  $X$ .
- Por cada nodo hay una tabla de probabilidad condicional que sirve para cuantificar los efectos de los padres sobre el nodo. Los padres de un nodo son aquellos nodos cuyas flechas apuntan hacia éste.
- El grafo no tiene ciclos dirigidos (por lo tanto es un grafo acíclico dirigido (GAD)). Esto significa que no se presentan ambigüedades en el encadenamiento de probabilidades condicionales por el hecho de tener influencias directas cíclicas.

Vista la red de creencias como el grafo junto con las tablas de probabilidad condicional, ella puede ser interpretada como una representación de la función de distribución de probabilidad conjunta<sup>1</sup> (DPC) de la

1 Dado un  $n+1$ -plus  $(X_1, X_2, \dots, X_n, Y)$  de variables aleatorias, se llama distribución de probabilidad conjunta a la función  $F[x_1, x_2, \dots, x_n, y] = \text{prob}[X_i \leq x_i, i = 1, \dots, n, Y \leq y]$ . Dicha probabilidad no puede calcularse en términos de las distribuciones individuales de  $X_1, X_2, \dots, X_n, Y$ , a menos que haya independencia.

clase y de todos los rasgos (variables). El árbol en sí codifica un conjunto de aseveraciones de independencia condicional. Las tablas de probabilidades condicionales completan la caracterización de la distribución conjunta.

El árbol es importante para construir la red en sí. Los valores que aparecen en las tablas de probabilidades condicionales son imprescindibles en el procedimiento de inferencia. Esta representación es a lo que algunos autores llaman *I-mapa minimal* de la distribución conjunta [Cas96][GAR90].

Formalmente, esta representación de la DPC permite definir un modelo de red bayesiana, como un par  $(D, P)$ , donde  $D$  es un grafo acíclico dirigido (GAD),  $P = \{p(x_1 | \Pi_1), \dots, p(x_n | \Pi_n)\}$  es un conjunto de  $n$  sentencias o tablas de probabilidad condicional CPDs (Conditional Probability Distribution), una para cada variable  $X_i$ , y  $\Pi_i$  es el conjunto de padres del nodo  $X_i$  en  $D$ .

## MODELOS QUE USAN REDES BAYESIANAS

Analizando los pasos de la ingeniería del conocimiento y el algoritmo para construir redes bayesianas [STU96], una vez establecidas las variables que intervienen en el modelo, se definen las relaciones de independencia condicional de estas variables; sin embargo, cuando se trabaja con los expertos, en ocasiones resulta difícil responder a preguntas como: ¿qué variables están relacionadas?, ¿en qué orden están relacionadas?. Una vez obtenida la topología de la red, cuáles son las tablas de probabilidades asociadas. Haciendo un análisis de todo esto se propone un método para construir redes bayesianas.

### Método para construir redes bayesianas

#### *Paso I. Obtención de la topología de la red*

Para obtener la arquitectura de la red se aplica el paquete de programas CHAID (Chi-squared Automatic Interaction Detector), el cual surge como una técnica de segmentación. Es particularmente útil en todos aquellos problemas en que se quiera subdividir una población a partir de una variable dependiente y de posibles variables predictoras, que cambien esencialmente los valores de la variable dependiente en

cada una de las subpoblaciones o segmentos. Ejemplos típicos asociados con su origen son, por ejemplo, los problemas de estudio de mercado. En estos casos la variable dependiente puede ser la aceptación o no de un producto, y las variables predictoras, un conjunto de características de la población que pueden influir en esta aceptación o no. La técnica de CHAID es capaz, en este caso, de segmentar la población en grupos, de acuerdo con determinados valores de esas variables predictoras y sus interacciones, que distinguen de forma óptima, en algún sentido, diferencias esenciales en el comportamiento de la variable dependiente [CHA94].

Desde esta formulación inicial, se concibió la posibilidad de la aplicación a diversas investigaciones de salud. La más típica de ellas es precisamente en epidemiología, en el estudio de los factores de riesgo asociados a una enfermedad. En tal caso, la variable dependiente puede ser simplemente la variable que distingue un grupo de enfermos y sanos, y las variables predictivas, los posibles factores de riesgo.

Más que segmentar la población en este caso, la técnica de CHAID se usa para:

- Conocer cuáles, entre decenas de variables (o rasgos) pueden ser eliminadas.
- Para comprender el orden de importancia de los rasgos (en las investigaciones epidemiológicas, para comprender el orden de los factores de riesgo en la caracterización de la enfermedad y en particular ayudar a detectar posibles factores confundidores o modificadores de riesgo).
- Para entender cómo interactúan los rasgos unos con otros (en las investigaciones epidemiológicas para entender cómo ciertos factores de riesgo interactúan con otros).
- Simplificar las crostabulaciones combinando categorías de variables predictoras que no difieren significativamente.

Esto último es de lo más interesante. CHAID combina categorías de una variable predictora que no difieren significativamente. De esta forma se resuelve, por ejemplo, el problema de cómo jerarquizar la edad (considerada como un posible factor de riesgo) para

obtener una tabla de contingencia significativa con la enfermedad. Simplemente si se hace en 10 o 12 categorías, CHAID se ocupará de unir las categorías consecutivas que no difieren significativamente, y el resultado final mostrará muchos menos rangos de edad, evidenciando las que constituyen un verdadero factor de riesgo. Lo mismo es capaz de hacer con variables, incluso nominales (por ejemplo la raza o color de la piel), e incluso con variables que tienen un valor perdido, asociando éste a la categoría de la variable con respecto a la cual los casos son más parecidos en su comportamiento. [ZAM97]

El CHAID permite obtener un árbol en forma automática con las características mencionadas y brinda la estructura de la red bayesiana con sólo invertir el orden del árbol. Sin embargo, no se descarta la participación de los expertos, pues este software permite obtener distintos árboles si se cambian parámetros o se hace interactivamente con el usuario; esto permite que se tenga en cuenta al experto a la hora de seleccionar la topología más adecuada.

### **Paso II. Obtención de las tablas de probabilidades**

Para la obtención de las tablas de probabilidad se propone seguir estos pasos:

- Calcular las probabilidades de todas las variables que son nodos, en el grafo utilizando el comando FRECUENCIAS del paquete estadístico SPSS (Statistical Package for Social Science); previamente deben ser ponderadas las variables, utilizando como pesos los valores de las tasas de la variable dependiente.

*En el SPSS las variables peso determinan el número de veces que cada caso es incluido en un análisis. Cuando se usa una variable peso, la contribución de un caso en el análisis es pesada en proporción a la población de casos representados en la muestra. Ello puede ser útil, por ejemplo, en un estudio de casos-controles de una enfermedad rara, en el cual se desee tener tamaños proporcionales del grupo de casos y de los controles a los de la población general. Necesitaría probablemente 200 individuos sanos por cada enfermo. Si usted colecciona datos de 150 enfermos y 150 sanos, puede pesar los primeros por "1" y los segundos por "200", y obtendría los resultados equivalentes a los de una muestra repre-*

*sentativa de la población. En el CHAID también se puede trabajar con variables peso, introduciéndolas como frecuencia.*

- Para cada nodo del grafo que tenga más de un padre, unir estas variables padres y volver al paso 1.
- Una vez calculadas las distribuciones de probabilidad para las variables individuales, y para las variables unidas (unión de las variables que son padres de un mismo nodo del grafo), utilizar tablas de contingencia simples (tabla que muestra el número de casos que hay en cada combinación de las categorías de dos variables categóricas [JOB92]) para calcular las probabilidades condicionadas.

Cuando se tiene elaborado el modelo de red bayesiana, para realizar la inferencia se utilizan algoritmos para la propagación de evidencias [CAS96]. Hasta ahora sólo se dispone del software para cuando la estructura de la red es un políárbo.

Haciendo uso del método se han realizado algunas aplicaciones, y a continuación se exponen dos de estas.

### **Ejemplo de aplicación a un problema de epidemiología**

Con el objetivo de caracterizar los riesgos de la cardiopatía izquémica, en 436 pacientes del municipio de Camajuaní se midieron las siguientes características a cada individuo de la muestra: grupo (enfermos o sanos), sexo, edad, raza, antecedentes de IMA (infartos del miocardio agudos), angina de pecho, otras formas de cardiopatías, hábitos de fumar, tiempo de fumador, si fuma específicamente cigarrillos, si fuma específicamente tabacos, si fuma específicamente en pipa, cantidad de cigarros, padecimiento de hipertensión arterial, grado de la hipertensión arterial, tiempo de padecimiento de la hipertensión arterial, antecedentes de diabetes mellitus, tipo de diabetes, tiempo de padecimiento de la diabetes, obesidad, sedentarismo, práctica de ejercicios físicos, si pertenece al círculo de abuelos, hiperlipoproteinemia, tipo de hiperlipoproteinemia, hábitos de ingestión de bebidas alcohólicas, estrés en el hogar, estrés en el trabajo, estrés en otro lugar.

Se estudiaron 324 enfermos y 112 sanos, conformando así los grupos de la muestra, que se distinguen por la variable dependiente Grupo (codificada con 1 para los en-

fermos y 2 para los sanos). El resto de las variables son las variables predictoras, que representan posibles factores de riesgo, o protectoras. La mayoría de estas variables son binarias o múltiples, algunas tienen nivel de medición ordinal, y otras son en principio continuas. Para el trabajo con el CHAID todas ellas serán discretizadas, estableciendo rangos en los casos necesarios (por ejemplo, en la edad o el tiempo de padecimiento de una enfermedad); pero el número de categorías en que ellas deben ser separadas no es motivo de preocupación inicial. Pueden ser muchas (prácticamente hasta 31 categorías), y la técnica de CHAID se ocupará de unir las que no difieran significativamente desde el punto de vista de la presencia o no de la enfermedad. En la figura 1 se muestra un ejemplo simple de un árbol unilateral, obtenido por el CHAID, para obtener la estructura de la red bayesiana, que se muestra en la figura 2. El árbol de decisión se utiliza para saber cuáles variables quedan en el modelo y cómo están condicionadas, aunque realmente lo que se hace es invertir el árbol; en este caso se obtiene un poliárbol donde la variable que más influye en un IMA es la angina de pecho, le siguen los antecedentes de IMA, otras formas de cardiopatía y sedentarismo.

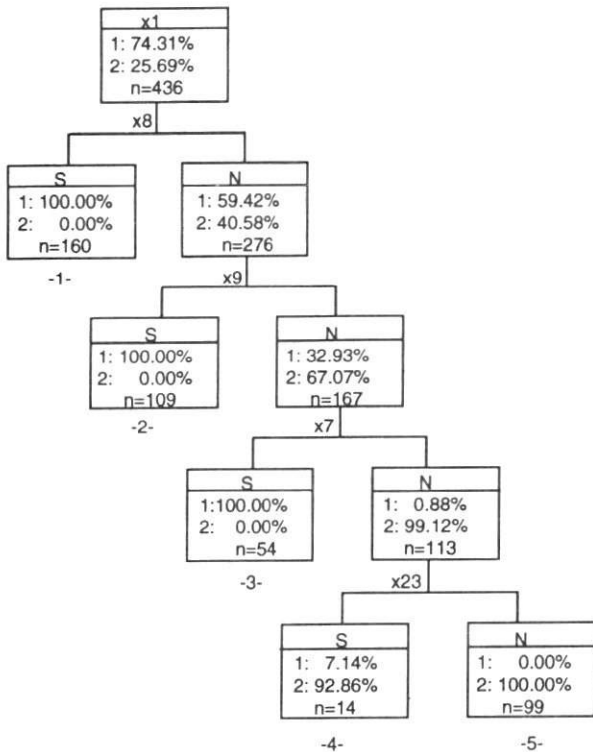


Figura 1. Árbol que ofrece el CHAID para el ejemplo de IMA

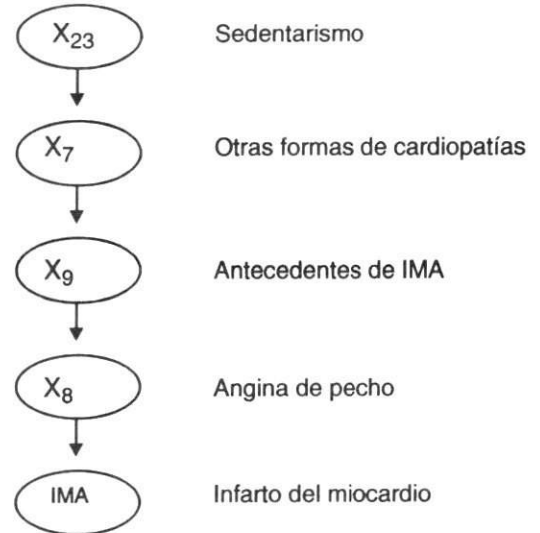


Figura 2. Estructura del modelo de RB obtenida para el diagnóstico de un IMA

Para formar el modelo faltarían las tablas de probabilidades asociadas, y para ello se aplicó el paso II del método propuesto. Supóngase que se dispone de la tasa histórica de IMA, con los valores:

$$p(\text{IMA} = \text{Si}) = 0.67$$

$$p(\text{IMA} = \text{No}) = 0.33$$

Con estos valores se determinan los pesos de ponderación y con el SPSS se calculan las probabilidades individuales (previa ponderación)

$$p(X_8), p(X_9), p(X_7), p(X_{23})$$

X <sub>23</sub>	p(X <sub>23</sub> )
Si	0.369
No	0.631

X <sub>7</sub>	p(X <sub>7</sub> )
Si	0.161
No	0.839

X <sub>8</sub>	p(X <sub>8</sub> )
Si	0.296
No	0.704

X <sub>9</sub>	p(X <sub>9</sub> )
Si	0.423
No	0.577

- Se definen las variables con nodos padres. En este caso hay un poliárbol, en el que, además, cada nodo tiene un único padre.
- Se calculan las probabilidades en cada nodo condicionadas a su padre. Ello se logra con tablas de contingencia ponderadas.

IMA	$X_8$	$P(IMA/X_8)$
Si	Si	1
Si	No	0.748
No	Si	0
No	No	0.252

$X_9$	$X_7$	$P(X_9/X_7)$
Si	Si	0.024
Si	No	0.349
No	Si	0.976
No	No	0.651

$X_8$	$X_9$	$P(X_8/X_9)$
Si	Si	0.027
Si	No	0.590
No	Si	0.973
No	No	0.41

$X_7$	$X_{23}$	$P(X_7/X_{23})$
Si	Si	0.204
Si	No	0.131
No	Si	0.796
No	No	0.869

Como el modelo de red bayesiana obtenido tiene una estructura de poliárbol, se puede realizar la inferencia con el software que implementa el algoritmo de propagación exacta en poliárboles.

Al hacer corridas del algoritmo se tienen los siguientes resultados: si se conoce que se tiene antecedentes de IMA, la probabilidad de que ocurra un IMA es 0.8933, si la evidencia es que se tienen otras formas de cardiopatía, la probabilidad de IMA es 0.7548; sin embargo, si estas dos cosas han ocurrido simultáneamente, la probabilidad de ocurrir un IMA se mantiene en 0.7548. La probabilidad de que ocurra un IMA es 1 solamente en el caso que se conozca que ocurrió una angina de pecho, lo que resulta evidente si la variable objetivo (IMA) depende directamente de la variable predictora que se menciona.

#### *Un ejemplo de aplicación al diagnóstico*

Se analiza ahora un problema en el que se pretende averiguar cómo diferentes alteraciones orofaciales influyen en la aparición de la retinosis pigmentaria. Se quiere saber cuáles de estas alteraciones internas o externas caracterizan realmente a la enfermedad, distinguiéndolas integralmente de las alteraciones que ocasionalmente pueden presentar los individuos sanos. Se pretende, además, encontrar la forma idónea de examinar sucesivamente dichas alteraciones con fines de diagnóstico.

En una primera fase del análisis, en la distinción del grupo de enfermos y sanos, se emplearon técnicas de

análisis discriminante. Para la aplicación rigurosa de estas técnicas, fue necesario corregir los niveles de medición de algunas variables no dicotómicas ni continuas (sin distribución normal, tales como tipo facial, tipo de perfil, y la forma del paladar blando). El análisis discriminante permitió determinar, en primer lugar, que las variables estudiadas (alteraciones internas y externas) son capaces de distinguir, en su conjunto, individuos enfermos y sanos, con alta significación. Este hecho no es trivial, pues pudiera suceder que en cada una de las variables hubiese una cierta diferencia, pero sólo el conjunto de ellas es capaz de identificar bastante bien los grupos. Este es el caso.

En la segunda etapa de este procesamiento se perfecciona el análisis de distinción de los grupos, utilizando la técnica de regresión logística. Ella, en principio, es mejor, porque no está sujeta a las restricciones sobre la distribución de las variables originales, y sus resultados son más fáciles de interpretar por médicos o epidemiólogos, ya que los parámetros estimados se analizan como factores de riesgo. [ZAM97]

En una tercera etapa se analizó que estas anomalías no son independientes, que están correlacionadas y pueden influir en interacción. Cuando este es el caso, el proceso de diagnóstico puede simplificarse si se atiende a estas interacciones, y se organiza un árbol que ordena y optimiza el diagnóstico. Dicho árbol se determina por criterios de razón de verosimilitud chi-cuadrado (técnicas de CHAID), y aparece reflejado en la Figura 3; en [ZAM97] se hace un análisis de estos resultados. En este caso se

utiliza el árbol para construir un modelo de red bayesiana, teniendo en cuenta que el mismo evidencia las variables más significativas y su interacción; la red obtenida se muestra en la figura 4. En este caso, cuando se va a invertir el árbol la, variable que influye directamente sobre la variable dependiente (grupo) es el tercio superior (TS), y después le siguen presencia de lengua de pájaro (LP) e hipertelorismo (H); en el grafo esto se manifiesta en que el nodo que representa al tercio superior tiene dos nodos padres, que representan cada una de estas variables. En este caso hay un camino del árbol en el que la variable LP depende de la variable H, por lo que se tiene un enlace desde el nodo LP hasta el nodo H; la otra variable en importancia es el paladar blando normal (PBN)

Para formar el modelo se calculan las tablas de probabilidades asociadas y se supone que se dispone de la tasa histórica de  $p(\text{grupo})$ , con los valores:

$$p(\text{grupo} = \text{Enf.}) = 0.001 \quad p(\text{grupo} = \text{Sanos}) = 0.999$$

Con estos valores se determinan los pesos de ponderación y con el SPSS se calculan las probabilidades individuales:  $p(\text{TS})$ ,  $p(\text{LP})$ ,  $p(\text{H})$ ,  $p(\text{BPN})$

TS	P(TS)
1	0.398
2	0.398
3	0.172
4	0.031

LP	P(LP)
Si	0.160
	0.840

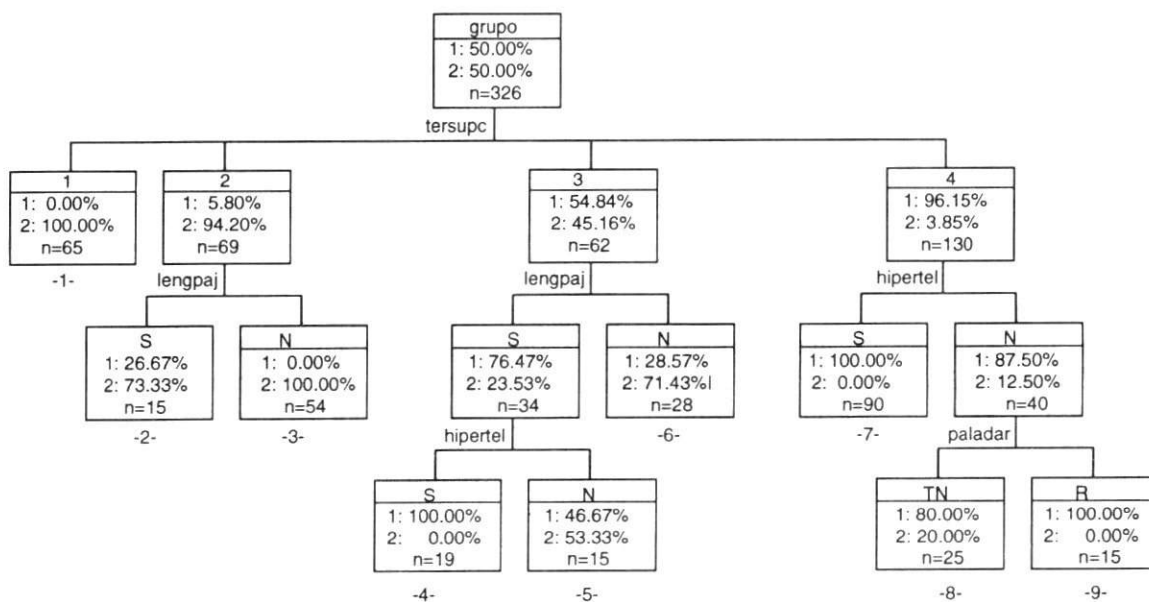


Figura 3. árbol que ofrece el CHAID para el ejemplo de retinosis pigmentaria

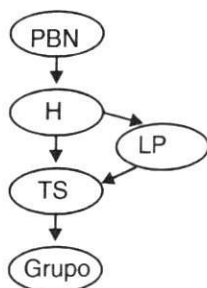


Figura 4. Red Bayesiana obtenida a partir del árbol que ofrece el CHAID

H	P(H)
Si	0.246
No	0.754

PBN	p(PBN)
NT	0.754
R	0.246

Ahora se definen las variables con nodos padres (en este caso el único nodo que tiene dos padres es TS). Como H toma los posibles valores codificados 1 y 2

y también LP, la combinación de las variables H y LP puede lograrse aritméticamente por la expresión:

$$H-LP = 10 * H + LP$$

Se calculan las probabilidades en cada nodo condicionada a su padre

Grupo	TS	P(grupo/TS)
Sanos	1	1
Sanos	2	1
Sanos	3	0.999
Sanos	4	0.976
Enf.	1	0
Enf.	2	0
Enf.	3	0.001
Enf.	4	0.024

LP	H	P(LP/H)
Si	Si	0.002
Si	No	0.212
No	Si	0.998
No	No	0.788

H	PBN	P(H/PBN)
Si	NT	0.155
Si	R	0.525
No	NT	0.845
No	R	0.475

P(TS/H-LP)	H-LP			
	Si Si	Si No	No Si	No No
1	0	0.475	0.192	0.423
2	0.021	0.325	0.423	0.423
3	0.202	0.200	0.308	0.124
4	0.777	0	0.078	0.031

En este ejemplo se pudo obtener el modelo de red bayesiana, sin embargo, no se puede realizar el proceso

de inferencia, debido a que el grafo asociado a la red no es un políárbol y hasta ahora sólo se ha implementado el algoritmo de inferencia para la propagación exacta de evidencias en un políárbol.

### CONCLUSIONES

Se demuestra como puede ser utilizada la técnica de CHAID y la estadística en el diseño de redes bayesianas, como un sustituto total o parcial de los expertos.

Se realizan aplicaciones en medicina que verifican el método propuesto.

### BIBLIOGRAFÍA

[AHA96] AHA DAVID W., CHANG LI WU. "Cooperative Bayesian and Cased-Based Reasoning for Solving Multia-gent Planning Tasks", Novy Center for Applied Research AI, Enero, 1996.

[AND89] ANDERSEN, S.K. OLESEN, K.G., JENSEN, F.V., AND JEUMEN, F. HUGIN - a Shell for bilding Bayesian be-lief universes for expert systems. In proceeding of the ele-venth International Joint Conference on AI, Vol. 2, pages 1080-1085, detroit Michigan. Morgan Kaufman, 1989.

[BRE92] BREESE JHON S., "Construction of belief and Decision networks". Computational Intelligence. Vol. 8, No. 4, 1992.

[BRE95] BREESE JHON S., HECKERMAN D. "Deci-sion-Theoretic Case-Based Reasoning". IEEE, Vol. 26, No. 6, Nov. 1995.

[BUC84] BUCHANAN B. G., SHORTLIFFE E. H. Rule Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Proje-t. Adison-Wesley, Reading, MA. 1984

[CAS89] CASTILLO ENRIQUE, ÁLVAREZ E. "Sistemas Expertos, Aprendizaje e Incertidumbre". 1994.

[CAS96] CASTILLO ENRIQUE, GUTIÉRREZ J. MANUEL, HADI ALI S. "Expert Systems and Probabilistics Network Models", 1996.

[CHA94] "CHAID para SPSS sobre Windows. Técnicas de seg-mentación basadas en razones de verosimilitud Chi-cuadrado", Manual de usuario, SPSS Soft. Inc. 1994.

[DUD80] DUDA R. O., HART P.E. Model Design in the PROSPECTOR Consultant System for Mineral Exploration. In Michie, D., editor, Expert System in the Microelectronic Age. Edinburgh University Press, 153-167. 1980



- [ETX98] ETXEBERRIA, R., LARRANAGA P., PICAZA J. M. "Analysis of the behaviour of genetic algorithms when learning Bayesian network structure from data". *Pattern Recognition Lett.* Vol. 18, 11-13, 1998.
- [EWA91] EDWARDS JOHN S. "Building Knowledge-Base Systems". 1991.
- [FRE87] FREZNEL LOUIS E. "Crash Course in Artificial Intelligence and Expert System". Indianapolis, IN: Howard W. Sams, 1987.
- [FER93] FERTING K. W., BREESE J. S. "Probability Intervals Over Influence Diagrams". *IEEE* Vol. 15, No. 3, Marzo 1993.
- [FUN95] FUNG ROBERT, DEL FAVERO B. "Applying Bayesian Networks to Information Retrieval", *ACM*, Vol. 38, No. 3, Marzo 1995.
- [GAR90] GARCÍA LUCIANO. "Probabilidad e Inteligencia Artificial".
- [HEC91] HECKERMAN DAVID. *Probabilistic Similarity Networks*. MIT Press, Cabridge, Massachusetts, 1991.
- [HEC95] HECKERMAN DAVID, BREESE J., ROMNELSE K. "Decision Theoretic : Troubleshootin". *ACM*, Vol. 38, No. 3, Marzo 1995.
- [JEN90] JENSEN, F.V., LAURITZEN, S.L., AND OLESEN, K.G. *Bayesian updating in causal probabilistic networks by local computations*. Computational Statistics Quartely, 1990.
- [JOB92] JOBSON, J. D. "Applied Multivariate Data Analysis". Vol. II: Categorical and Multivariate Methods, Springer, New York, 11-54, 1992.
- [KIM83] KIM, J.H. *CONVINCE: A conversational Inference Consolidation Engine*. PhD Thesis Department of Computer Science, University of California at Los Angeles, 1983.
- [LAM98] LAM W. "Bayesian network refinement via machine learning approach". *IEEE Trans Patt Anal Mach Int.* Vol. 20, No. 3, 240-251, 1998.
- [MAR98] MARDIA K.V., BACZKOWSKI A.J., FENG X., Hainsworth T.J. "Statistical methods for automatic interpretation or digitally scanned finger prints". *Pattern Recognition Lett.* Vol. 18, 11-13, 1998.
- [MYL93] MYLLIMAKI P., TIRRI H. " Massively parallel case-based reasoning with probabilistic similarity metrics ". *Proceedings of the First European Workshop on Case-Based Reasoning.* 145-154, 1993.
- [PEA88] PEARL J. "Probabilistic Reasoning in Intelligent System". San Mateo, CA: Morgan Kaufman, 1988.
- [REG93] REGAZZONI C.S., MURINO V., VERNAZZA G. "Distributed propagation of a priori constraints in Bayesian Network of Markov random fields". *IEEE* Vol. 40, No. 1, Febrero 1993.
- [SHO86] SHORTLIFFE, E. H. *Computer-Based Medical Consultations: MYCIN*. Elsevier North-Holland, amsterdam, London, New York, 1986.
- [SMY98] SMYTH P. "Belief networks, hidden Markov models, and Markov random fields: A unifying view". *Pattern Recognition Lett.* Vol. 18, 11-13, 1998.
- [SPI86] SPIGELHALTER, D., DAVID,P., LAURITZEN, S., AND COWEEL, R. *Bayesian analysis in expert systems*. Statistical Science, 1993.
- [STU96] STUART RUSELL, NAVING P. "Inteligencia Artificial: Un enfoque Moderno". 1996.
- [WIL98] WILLIAMS W. L., WILSON, R. C., HANCOCKE. R. "Multiple graph matching with Bayesian inference". *Pattern Recognition Lett.* Vol 38, 11-13, 1998.
- [ZAM97] ZAMORA L., "Tesis de Maestria", 1997.