

APRENDIZAJE DE HIPERGRAFOS POSIBILISTICOS.*

Luis M. de Campos

Dpto. Ciencias de la Computación e I.A.
Universidad de Granada
18071-Granada, Spain
E-mail: lci@robinson.ugr.es

Juan F. Huete

Dpto. Ciencias de la Computación e I.A.
Universidad de Granada
18071-Granada, Spain
jhg@robinson.ugr.es

Resumen

Un Hipergrafo es una estructura gráfica que nos permite representar distintas componentes de información sobre un determinado problema. En un Hipergrafo posibilístico, a cada componente de información le asociamos una distribución de imposibilidad conjunta. En particular, el trabajo se centra en el estudio de algoritmos que nos permitan recuperar un Hipergrafo utilizando las relaciones de independencia existentes entre las variables.

Palabras Clave: Hipergrafos Posibilísticos, Independencia Posibilística, Aprendizaje.

1 Introducción.

El uso de modelos gráficos es de gran utilidad a la hora de tener una representación eficiente del conocimiento incierto. En general, mediante la estructura del modelo se representan relaciones de dependencia / independencia entre variables, junto con una representación cuantitativa de la información. Un aspecto importante, cuando trabajamos con incertidumbre, es la selección de un formalismo adecuado para representar cuantitativamente la información. En este sentido, podemos encontrar distintos formalismos como la Teoría de la Probabilidad, Teoría de la Evidencia [16] o la Teoría de la Posibilidad [7, 18]. En este trabajo nos centramos en el estudio de sistemas posibilísticos. Estos sistemas permiten diseñar técnicas de razonamiento aproximado, esto es, con un comportamiento poco sensible a pequeños cambios de la información. Basándonos en este hecho, se puede pensar que la teoría de la posibilidad puede representar el mismo papel dentro del

*Este trabajo ha sido financiado por la DGICYT bajo el proyecto n. PB92-0939

razonamiento con incertidumbre que el control difuso en el marco de la ingeniería de control [14].

Cuando trabajamos con incertidumbre, podemos encontrar distintos modelos gráficos. En especial, podemos destacar (1) los modelos causales [3, 14, 15] donde cada variable es un nodo del modelo y donde la ausencia de una conexión directa representa una relación de independencia entre las variables y (2) los modelos basados en Hipergrafos [4, 8, 17] donde las variables se agrupan en nodos que representan relaciones de dependencia. Nuestro interés se centra en el estudio de algoritmos que nos permitan realizar el aprendizaje de estructuras gráficas en entornos posibilísticos. Por ejemplo, Kruse et al. [9] proporciona un algoritmo de aprendizaje para modelos causales posibilísticos. En este trabajo, presentamos un algoritmo de aprendizaje para hipergrafos posibilísticos basado en el estudio de relaciones de independencia entre variables.

En la siguiente sección introducimos brevemente la teoría de la posibilidad. La sección 3 presenta el Hipergrafo posibilístico como una herramienta que permite construir un sistema basado en el conocimiento. En la sección 4 vemos como un Hipergrafo puede verse como la descomposición de una distribución de posibilidad conjunta en términos de conjuntos de variables condicionalmente independientes y, utilizando esta propiedad, presentamos un algoritmo que permite recuperar un Hipergrafo a partir de un conjunto de datos. Para finalizar, se presentan las conclusiones y líneas de trabajo futuras.

2 Teoría de la Posibilidad.

Cuando partimos de un conocimiento impreciso, una medida de posibilidad es una herramienta natural para formalizar la incertidumbre. En este trabajo consideraremos que una distribución de posibilidad establece, esencialmente, un orden entre los valores que una determinada variable puede tomar. Formalmente,

si D es el conjunto de referencia, entonces una medida de posibilidad puede definirse como una función $\Pi : \mathcal{P}(D) \rightarrow [0, 1]$ verificando

$$\forall A, B \subseteq D \quad \Pi(A \cup B) = \max\{\Pi(A), \Pi(B)\} \quad (1)$$

donde $\Pi(\emptyset) = 0$ y $\Pi(D) = 1$. Cuando el conjunto D es finito, una medida de posibilidad Π puede obtenerse a través de los valores en los elementos atómicos de D

$$\forall A, \Pi(A) = \max\{\pi(w) \mid w \in A\}$$

donde $\pi(w) = \Pi(\{w\})$ y π es una función de D en $[0, 1]$ llamada *distribución de posibilidad*. Como condición de normalización, se exige que exista un elemento w tal que $\pi(w) = 1$ indicando que existe al menos un valor en D que es completamente posible. Cuando tenemos una variable que toma valores en D , $\pi(w)$ se interpreta como el grado en que $w \in D$ puede ser un valor posible de la variable y por tanto $\Pi(A)$ puede considerarse como la posibilidad de que la variable tome como valor cualquier elemento de A .

Sean X e Y dos variables y sea π una distribución de posibilidad definida sobre el producto cartesiano de sus conjuntos de referencia, $D_X \times D_Y$. Considerando la semántica asociada a la distribución de posibilidad, utilizaremos como operador de condicionamiento el dado por Hisdal [6, 11]. Esto es, la distribución de posibilidad de X , condicionada al suceso $\{Y = y\}$, y lo notamos por $\pi_h(\cdot \mid y)$, viene dada por

$$\pi_h(x \mid y) = \begin{cases} \pi(x, y) & \text{si } \pi(x, y) < \pi_Y(y). \\ 1 & \text{si } \pi(x, y) = \pi_Y(y). \end{cases}$$

donde π_Y representa la distribución de posibilidad marginal para la variable Y , es decir $\pi_Y(y) = \max_{x \in D_X} \pi(x, y)$.

Otro concepto importante a considerar es el de independencia posibilística. En este caso, utilizamos un criterio de independencia basado en la imprecisión. Es decir, diremos que la variable X es condicionalmente independiente de la variable Y dado que conocemos el valor de Z , y lo notamos por $I(X \mid Z \mid Y)$, cuando al condicionar obtenemos una información coherente con la original, pero menos precisa. Formalmente podemos decir que

$$I(X \mid Z \mid Y) \Leftrightarrow \pi_h(x \mid yz) \geq \pi_h(x \mid z), \quad \forall x, y, z.$$

Notemos que esta definición de independencia es equivalente [1, 12] a la no interactividad posibilística [18], es decir, equivale a $\pi(xyz) = \pi(xz) \wedge \pi(yz)$, $\forall x, y, z$.

3 Hipergrafos Posibilísticos.

En teoría de grafos [10], el concepto de Hipergrafo (ver Figura 1) ha sido ampliamente estudiado. Formalmente, dado un conjunto de variables de referencia

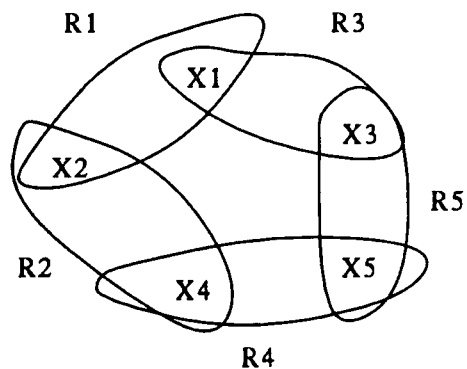


Figura 1: Hipergrafo

$S = \{X_1, \dots, X_n\}$, un Hipergrafo \mathcal{H} sobre S se puede definir como un conjunto no vacío de subconjuntos de S . A los elementos de \mathcal{H} se les llama hiperlados, y a los elementos de S se les denomina vértices. Un Hipergrafo se puede considerar como una estructura gráfica que permite representar distintas componentes de información sobre un determinado problema. En este caso, cada componente (hiperlado) agrupará a un conjunto de variables.

Cuando el conocimiento de partida es impreciso o vago, la teoría de la posibilidad se presenta como un formalismo razonable para representar cuantitativamente la información. En particular, para cada componente de información, R_j , tendremos una distribución de posibilidad (relación difusa) sobre el conjunto de variables que la componen, π_{R_j} . Por tanto, un Hipergrafo Posibilístico [4, 8] se puede considerar como un hipergrafo, donde asociado a cada hiperlado tenemos una distribución de posibilidad marginal.

Este tipo de estructuras permiten realizar razonamiento posibilístico. Esto es, a la luz de nueva información, obtener un conjunto de conclusiones lo más precisas posibles. A la hora de realizar inferencias, son dos las operaciones básicas [8]: la primera es la combinación de información (mediante el operador mínimo), que permite obtener la distribución de posibilidad menos restrictiva compatible con las restricciones y la segunda es una proyección (mediante el operador máximo) de los resultados sobre las variables de interés.

Hemos de notar que en un entorno posibilístico, se satisfacen los axiomas de Shafer y Shenoy [17] y por tanto es posible utilizar técnicas locales para realizar una propagación de la información en árboles de Markov: Un árbol de variables donde dos variables están separadas si existe una relación de independencia condicional, dado el conjunto de variables que las separan. Dubois y Prade [8] proporcionan un conjunto de técnicas que permiten transformar un Hipergrafo posibilístico en

un Hiperárbol, cuyos hiperlados se pueden organizar como un árbol de Markov.

4 Aprendizaje de Hipergrafos Posibilísticos a partir de Datos.

A la hora de realizar tareas de razonamiento en un hipergrafo, es necesario tener como punto de partida la estructura gráfica ya construida. Sin embargo, para muchos problemas, el proceso de elicitación de la estructura a partir de un experto puede ser una tarea bastante costosa. En esta sección nos centraremos en el estudio de algoritmos que permitan, a partir de una base de casos, obtener este tipo de estructuras, junto con los valores cuantitativos asociados a las mismas. Este tipo de algoritmos puede ser de gran ayuda en el proceso de diseño de un hipergrafo.

En concreto, partimos de una base de casos sobre un conjunto de variables $S = \{X_1, X_2, \dots, X_n\}$, de la que obtenemos la distribución de posibilidad que siguen los datos, π . Nuestro objetivo será construir el Hipergrafo que mejor represente esta distribución. Para ello consideramos que en un hipergrafo, además de representarse relaciones de dependencia entre variables, también se representan relaciones de independencia condicionales.

En este proceso, tratamos de factorizar la distribución de posibilidad original en componentes condicionalmente independientes. Así, supongamos que π es una distribución de posibilidad sobre S y que existe un conjunto de variables $C = \{X_i, \dots, X_j\}$ verificando la relación $I(X_1, \dots, X_{i-1} | C | X_{j+1}, \dots, X_n)$. Entonces, podemos factorizar la distribución original en dos distribuciones, π_1 sobre $S_1 = \{X_1, \dots, X_{i-1}, X_i, \dots, X_j\}$ y π_2 sobre $S_2 = \{X_i, \dots, X_j, X_{j+1}, \dots, X_n\}$. Basándonos en el criterio de independencia utilizado, tenemos que la distribución original se puede recuperar mediante $\pi = \pi_1 \wedge \pi_2$. Aplicando este esquema de forma recursiva para los conjuntos S_1 y S_2 , obtenemos una estructura arbórea, que denominaremos árbol posibilístico.

La Figura 2 muestra un árbol posibilístico, donde en cada nivel se presentan junto con el conjunto de nodos, las relaciones de independencia que permiten construirlo. En estas estructuras, la información se localiza únicamente en las hojas del árbol, $L_j, j = 1 \dots m$, en forma de distribuciones de posibilidad marginales. La distribución de posibilidad original se puede recuperar [2] mediante

$$\pi(x_1, x_2, \dots, x_n) = \pi_{L_1} \wedge \pi_{L_2} \wedge \dots \wedge \pi_{L_m} = \bigwedge_{j=1}^m \pi_{L_j}$$

con π_{L_j} siendo la distribución de posibilidad marginal almacenada en la hoja L_j

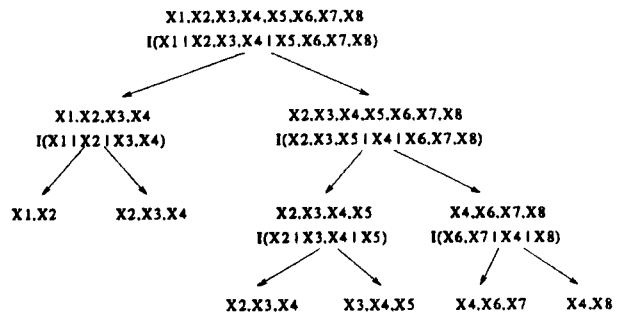


Figura 2: Arbol Posibilístico

Asociado a un árbol posibilístico, es posible construir un Hipergrafo posibilístico (Figura 3). Para ello, basta con considerar como hiperlados las hojas del árbol posibilístico. En este caso, en un Hipergrafo se representan además de las relaciones entre variables, un conjunto de relaciones de independencia condicional entre ellas. Este hecho hace que un Hipergrafo represente un papel similar a las redes bayesianas [15] en un formalismo probabilístico, permitiéndonos utilizar técnicas de propagación local, de forma que se obtengan los mismos resultados que se obtienen cuando utilizamos la distribución de posibilidad inicial [2]

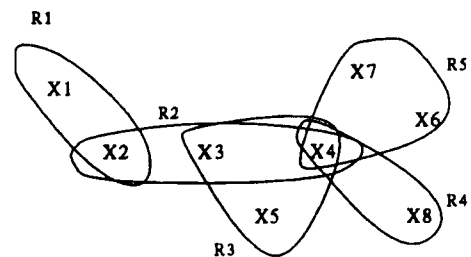


Figura 3: Hipergrafo asociado a T

Por tanto, para obtener un hipergrafo, nos bastará con encontrar el conjunto de nodos hoja en un árbol posibilístico. El algoritmo seguirá el mismo esquema utilizado para construir el árbol posibilístico, esto es:

```

BuscarHojas( S:Conjunto de Nodos )
{
  Buscar subconjuntos disjuntos L, R, C en S con
  I(L | C | R) y L ∪ R ∪ C = S.
  Si no es posible, Entonces
    - Incluir S en la lista de nodos hoja.
    - Asociar a S la distribución marginal πS.
  En caso contrario
    BuscarHojas( R ∪ C )
    BuscarHojas( L ∪ C )
}
  
```

El siguiente ejemplo nos muestra el comportamiento del anterior algoritmo cuando partimos de una base de casos. En el ejemplo, al partir de una base de casos, hemos considerado que estos no tienen porque ser una muestra exacta de la distribución de posibilidad que sigue el modelo. Por tanto, a la hora de realizar los tests, relajaremos el criterio de independencia. En particular, consideraremos que se satisface la relación $I(X|Z|Y)$ cuando, conocido el valor de Z , el conocer el valor de la variable Y proporciona una **información similar** (sobre los valores de X) a la que teníamos antes de conocer el valor de Y [1]. Más formalmente, diremos que

$$I(X|Z|Y) \Leftrightarrow \pi(xyz) \simeq [\pi(xz) \wedge \pi(zy)]$$

donde \simeq viene definida por:

$$\pi \simeq \pi' \Leftrightarrow \forall x, |\pi(x) - \pi'(x)| \leq \epsilon$$

Ejemplo: Sean X, Y, Z, W variables bivaluadas, y $B1, B2$ dos bases de casos sobre $X \times Y \times Z \times W$, con tamaño 1.000 y 10.000 respectivamente. En la Tabla 1 representamos la frecuencia con la que aparecen cada uno de los distintos casos en $B1$ y $B2$ ¹. Partiendo

Tabla 1: Bases de Casos: $B1$ y $B2$.

YZW	B1		B2	
	$X = x_1$	$X = x_2$	$X = x_1$	$X = x_2$
$y_1 z_1 w_1$	12	11	135	129
$y_1 z_1 w_2$	13	13	139	134
$y_1 z_2 w_1$	53	41	533	401
$y_1 z_2 w_2$	66	42	665	394
$y_2 z_1 w_1$	93	93	938	931
$y_2 z_1 w_2$	108	131	1059	1337
$y_2 z_2 w_1$	81	80	801	791
$y_2 z_2 w_2$	80	82	809	804

de estos datos, utilizando una normalización por el máximo [13] obtenemos las distribuciones de posibilidad asociadas, π_1 y π_2 , que siguen los datos (Tabla 2). Considerando un umbral $\epsilon = 0.02$, obtenemos que $I(X|Y|ZW)$ y $I(Z|Y|W)$. Tras aplicar el algoritmo

¹Nota: El conjunto de datos en $B1$ y $B2$ se han obtenido mediante el siguiente proceso: (1) Partiendo del Hipergrafo de la Figura 4, calculamos $\pi(xyzw)$ mediante

$$\pi(xyzw) = \pi(xy) \wedge \pi(yz) \wedge \pi(yw)$$

con $\pi_{XY}(x_1 y_1, x_1 y_2, x_2 y_1, x_2 y_2) = (0.5, 0.8, 0.3, 1)$, $\pi_{YZ}(y_1 z_1, y_1 z_2, y_2 z_1, y_2 z_2) = (0.1, 0.5, 1, 0.6)$ y $\pi_{YW}(y_1 w_1, y_1 w_2, y_2 w_1, y_2 w_2) = (0.4, 0.5, 0.7, 1)$; (2) La distribución de posibilidad se transforma en una distribución de probabilidad, P , utilizando una normalización por el máximo [13]; (3) A partir de P se generan aleatoriamente las bases de casos.

Tabla 2: Distribuciones de Posibilidad π_1 y π_2 .

YZW	π_1		π_2	
	$X = x_1$	$X = x_2$	$X = x_1$	$X = x_2$
$y_1 z_1 w_1$	0.099	0.083	0.100	0.096
$y_1 z_1 w_2$	0.115	0.099	0.103	0.100
$y_1 z_2 w_1$	0.404	0.312	0.398	0.299
$y_1 z_2 w_2$	0.503	0.320	0.497	0.294
$y_2 z_1 w_1$	0.709	0.709	0.701	0.696
$y_2 z_1 w_2$	0.824	1	0.792	1
$y_2 z_2 w_1$	0.618	0.610	0.599	0.591
$y_2 z_2 w_2$	0.610	0.625	0.605	0.601

BuscarHojas, obtenemos el Hipergrafo de la Figura 4 (esto es, recuperamos el Hipergrafo original), donde las distribuciones de posibilidad (utilizando π_2) para cada hiperlado vienen dadas por:

$$\pi_{XY}(x_1 y_1, x_1 y_2, x_2 y_1, x_2 y_2) = (0.497, 0.792, 0.299, 1),$$

$$\pi_{YZ}(y_1 z_1, y_1 z_2, y_2 z_1, y_2 z_2) = (0.103, 0.497, 1, 0.605)$$

$$\text{y } \pi_{YW}(y_1 w_1, y_1 w_2, y_2 w_1, y_2 w_2) = (0.398, 0.497, 0.701, 1)$$

□

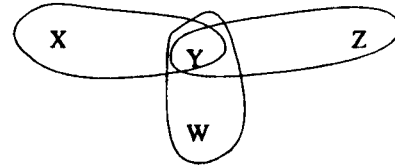


Figura 4: Hipergrafo aprendido

Cuando utilizamos un umbral ϵ igual a cero, obtenemos que el Hipergrafo refleja fielmente la distribución de posibilidad que siguen los datos. Sin embargo, cuando el umbral es estrictamente positivo, el Hipergrafo representará una aproximación de la distribución original. La siguiente proposición nos permite relacionar ambas distribuciones:

Proposición: Sean U un conjunto de variables, B una base de casos sobre U , π_B la distribución de posibilidad que se obtiene a partir de B , \mathcal{H} un Hipergrafo obtenido al aplicar el algoritmo *BuscarHojas* con un umbral $\epsilon > 0$ y $\pi_{\mathcal{H}}$ la distribución de posibilidad representada por el Hipergrafo. Entonces, $\pi_{\mathcal{H}}$ es menos precisa que π_B

$$\forall u \in U, \pi_B(u) \leq \pi_{\mathcal{H}}(u)$$

Demostración: Es suficiente con ver que, para toda instanciación $u \in U$, asignamos a $\pi_{\mathcal{H}}(u)$ un valor mayor o igual que $\pi_B(u)$. En particular, tenemos que $\pi_{\mathcal{H}}(u) = \bigwedge_j \pi_{L_j}(u^{\downarrow L_j})$, con $u^{\downarrow L_j}$ la proyección de la instanciación u sobre las variables en L_j . Considerando que cada $\pi_{L_j}(u^{\downarrow L_j})$ es una marginal de $\pi_B(u)$, esto es, $\pi_{L_j}(u^{\downarrow L_j}) = \max_{U \setminus L_j} \pi_B(u)$, tenemos $\pi_{L_j}(u^{\downarrow L_j}) \geq$

$\pi_B(u)$, luego $\pi_B(u) \leq \pi_H(u)$, quedando demostrada la proposición. \square

Esta proposición nos permite asegurar que en el Hipergrafo tenemos representada una distribución de posibilidad menos informativa que la original, pero satisfaciendo las relaciones de independencia testeadas. El principal problema que plantea este algoritmo *BuscarHojas* es su alto coste computacional, pudiendo ser necesario realizar un número exponencial de tests de independencia para encontrar el conjunto C .

Con el fin de obtener algoritmos más eficientes de aprendizaje, nos planteamos una segunda aproximación al problema. En particular, consiste en utilizar, como paso intermedio, una estructura causal, para posteriormente transformar esta estructura en un árbol posibilístico. En este sentido, podemos considerar el concepto de independencia como un concepto abstracto, independiente del formalismo que se utilice para representar la información, y por tanto utilizar algoritmos de aprendizaje de estructuras causales, particularizados para el formalismo posibilístico [12].

Una vez que tenemos la estructura causal G , esto es, tenemos un grafo dirigido acíclico, el siguiente algoritmo nos permite trasformarla en un árbol posibilístico. En el algoritmo utilizamos la siguiente relación de orden sobre el conjunto de variables en G : Si existe el arco $X_i \rightarrow X_j$ en G , entonces $X_i < X_j$ en el orden d . Notaremos como $\text{Pred}_d(X_i)$ al conjunto de predecesores de X_i en el orden d , esto es, $\text{Pred}_d(X_i) = \{X_1, X_2, \dots, X_{i-1}\}$ y notaremos como $\text{Padres}(X_i)$ al conjunto de variables, X_p , tales que existe un arco directo $X_p \rightarrow X_i$ en G . Para realizar la transformación utilizaremos la siguiente propiedad [15]: Sea G una estructura causal y d un orden sobre las variables (obtenido mediante la relación anterior). Entonces, se satisface que

$$I(X_i | \text{Padres}(X_i) | \text{Pred}_d(X_i) \setminus \text{Padres}(X_i)).$$

Por tanto, considerando el conjunto de variables en orden inverso, esto es, $S = \{X_n, X_{n-1}, \dots, X_1\}$, podemos contruir un árbol posibilístico T siguiendo los siguientes pasos:

- Construir un árbol T con raíz $I(X_n | \text{Padres}(X_n) | \text{Pred}_d(X_n) \setminus \text{Padres}(X_n))$.
- Insertar como hijo izquierda de T el conjunto de variables $L = \{X_n\} \cup \text{Padres}(X_n)$, asociado con la distribución de posibilidad marginal $\pi_{\{X_n\} \cup \text{Padres}(X_n)}$.
- Recursivamente, si $\text{Pred}_d(X_n) \setminus \text{Padres}(X_n)$ no es el conjunto vacío, insertar como hijo derecha de T el árbol que se obtiene al repetir el proceso para el conjunto de variables $\{X_{n-1}, X_{n-2}, \dots, X_1\}$.

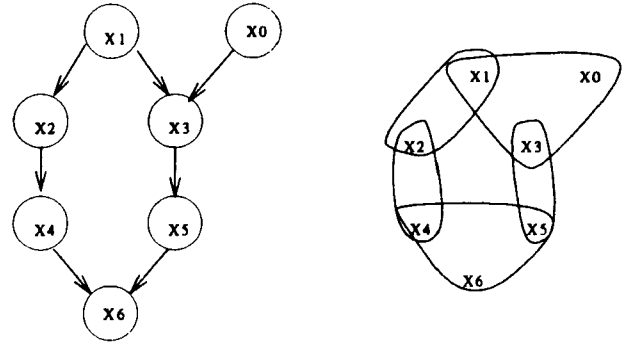


Figura 5: Red Causal e Hipergrafo asociado

Es posible que, como resultado de aplicar el esquema anterior, aparezcan hiperlados anidados en el Hipergrafo salida. En este caso, puesto que operamos con el mínimo, podemos eliminar los hiperlados más internos (con menos variables) sin modificar la distribución de posibilidad representada en el Hipergrafo. Por ejemplo, dada la red causal de la Figura 5 (parte izquierda), el algoritmo recupera el Hipergrafo representado en la parte derecha de la misma figura, incluyendo dos hiperlados más, el formado por la variable x_0 y el formado por la variable x_1 .

5 Conclusiones y Trabajos Futuros.

En este trabajo hemos visto como un Hipergrafo se puede considerar como una representación gráfica de una distribución de posibilidad. Para ello, vemos como el concepto de independencia condicional representa un importante papel. Sin embargo, a la hora de diseñar un algoritmo de aprendizaje que intente recuperar la distribución de forma exacta, el principal problema que se plantea es el alto coste computacional necesario. Por tanto, como posibles trabajos futuros nos plantearemos el estudio de algoritmos aproximados de aprendizaje para este tipo de estructuras. Además, es interesante realizar un estudio axiomático sobre las relaciones de independencia que se representan mediante un Hipergrafo y su relación con la axiomática de independencias dada para modelos causales [15]. Otro punto de interés es el estudio de los hipergrafos cuando consideramos que la medida de posibilidad es un caso particular de las medidas de evidencia [5], en concreto es una medida de plausibilidad consonante.

Referencias

- [1] L.M.de Campos and J.F. Huete. Possibilistic Independence. In *Proceeding of third European Congress on Intelligent Techniques and Soft Computing. EUFIT'95*, pages 69-74, 1995.

- [2] L.M.de Campos and J.F. Huete. How to construct possibility measures from a set of marginals. In *International Conference on Information Processing and Management of Uncertainty in Knowledge Based Systems, IPMU'96*, Julio, 1996.
- [3] J.E. Cano. *Propagación de probabilidades inferiores y superiores en grafos*. Tesis Doctoral. Universidad de Granada, 1992.
- [4] P. Chatalic, D. Dubois, and H. Prade. A system for handling relational dependencies in approximate reasoning. In *Proc. 3 Inter. Expert Systems Conference, London*, pages 495–502, 1987.
- [5] A.P. Dempster. Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematics and Statistics*, 38:325–339, 1967.
- [6] D. Dubois. Belief structures, possibility theory, decomposable confidence measures on finite sets. *Computer and Artificial Intelligence*, 5(5):403–417, 1986.
- [7] D. Dubois and H. Prade. *Possibility Theory: An approach to computerized processing of uncertainty*. Plenum Press, 1988.
- [8] D. Dubois and H. Prade. Inference in Possibilistic Hypergraphs. In *International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, IPMU'90*, pages 250–259. Lecture Notes in Computer Science 521. Eds B. Bouchon-Meunier and R.R. Yager and L.A. Zadeh, 1990.
- [9] J. Gebhardt and R. Kruse. Learning Possibilistic Graphical Models. In *Proceeding of third European Congress on Intelligent Techniques and Soft Computing. EUFIT'95*, pages 57–64. Lecture Notes in Computer Science 747. Eds M. Clarke and R. Kruse and S. Moral, 1993.
- [10] M.C. Golumbic. *Algorithmic Graph Theory and Perfect Graphs*. Academic Press., 1980.
- [11] E. Hisdal. Conditional possibilities, independence and noninteraction. *Fuzzy Sets and Systems*, 1:283–297, 1978.
- [12] J.F. Huete. *Aprendizaje de Redes de Creencia mediante la detección de independencias: Modelos no probabilísticos*. Ph.D. Thesis, University of Granada, 1995.
- [13] G. Klir and B. Parviz. Probability-Possibility transformations: A comparison. *International Journal of General Systems*, 21:291–310, 1992.
- [14] R. Kruse, J. Gebhardt, and F. Klawonn. *Foundations of fuzzy systems*. Wiley, 1994.
- [15] J. Pearl. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan and Kaufmann, San Mateo, 1988.
- [16] G. Shafer. *A mathematical theory of evidence*. Princeton University Press, Princeton N.J., 1976.
- [17] G. Shafer and P.P. Shenoy. Local computation in hypertrees. In *Working Paper 201*. School of Business, Univ. of Kansas, 1988.
- [18] L.A. Zadeh. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, (1):3–28, 1978.