00365
4
2ej

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS
DIVISION DE ESTUDIOS DE POSGRADO

AGRUPACIONES EN GRAFICAS DIFUSAS

T E S I S

QUE PARA OBTENER EL GRADO ACADEMICO DE

MAESTRO EN CIENCIAS

(MATEMATICAS)

P R E S E N T A

JUAN JOSE MONTELLANO BALLESTEROS

DIRECTOR DE TESIS: DR. JOSE RUIZ SHULCLOPER.

A mis padres, hermanos y amigos.

Quiero expresar mi agradecimiento al Dr. José Ruiz Shulcloper ("Shul") por toda la ayuda, apoyo y asesoramiento recibido en la realización de este trabajo. A Manuel, Nancy y al Grupo de Reconocimiento de Patrones del ICIMAF de Cuba por su gran apoyo y colaboración. Además quiero agradecerles a Luis y Mónica toda la confianza y paciencia que me brindaron. Y a todos aquellos que de una manera u otra hicieron que este trabajo fuera posible, muchas gracias.

INDICE

Introducción

Con la incorporación de la Teoría de los Subconjuntos Difusos para la solución matemática de problemas de Reconocimiento de Patrones en 1966 por Bellman, Kalaba y Zadeh [4], se dió inicio a un gran desarrollo de algoritmos de análisis de cúmulos difusos, siendo claves en este desarrollo el concepto de c-partición difusa definido por E. Ruspini en 1969 [20] y la elaboración del modelo de algoritmos Fuzzy c-means realizada por J. C. Bezdek en 1973 [5]. Es a partir de estas ideas, y sobre todo de este modelo, que se han dirigido la mayor parte de los esfuerzos subsecuentes en la búsqueda de soluciones a los problemas matemáticos de análisis de cúmulos difusos. Sin embargo, los algoritmos desarrollados hasta el momento no han superado ciertas restricciones ya presentes en los primeros algoritmos tales como: la predefinición del número de cúmulos; la necesidad de describir a los objetos en términos cuantitativos; y que la función de semejanza entre objetos sea una norma. Este tipo de restricciones impide la aplicación cabal de estos algoritmos en diferentes problemas de análisis de cúmulos referentes a disciplinas tales como la Biología, Medicina, Sociología y Geociencias entre otras, donde es común encontrar condiciones como las siguientes: no tener ningún indicio de cuántos agrupamientos o cúmulos son; las descripciones de los objetos no están definidas exclusivamente en términos cuantitativos; y la función de semejanza no cumple necesariamente ningún tipo de propiedad más que el hecho de que los valores que toma sean todos comparables entre sí a partir de cierto orden.

En el presente trabajo proponemos un nuevo modelo de algoritmos de análisis de cúmulos difusos, basado en la Teoría de los Subconjuntos Difusos. Este modelo tiene como principal objetivo el ser una herramienta útil para aquellos problemas de análisis de cúmulos en los cuales están presentes las condiciones antes mencionadas.

Con el objetivo de sustentar y exponer lo más claramente posible la definición de este modelo, el trabajo se presenta de la siguiente manera:

En el primer capítulo exponemos los conceptos de la Teoría de Subconjuntos Difusos necesarios para la creación del modelo. En el capítulo dos damos una breve semblanza del problema general de Reconocimiento de Patrones desde la perspectiva de las condiciones antes mencionadas, y ahondamos en el tema de análisis de cúmulos, donde presentamos el modelo de algoritmos Fuzzy c-means y las limitaciones del mismo con respecto a este tipo de problemas. A su vez hacemos un breve repaso de los criterios de agrupación a partir de las cuales se nutrió el nuevo modelo. En el capítulo tres definimos el nuevo modelo y algunas de sus propiedades. Por último presentamos las conclusiones obtenidas a partir de este trabajo.

CAPITULO I

TEORIA DE SUBCONJUNTOS DIFUSOS

Subconjuntos Difusos

Las primeras publicaciones sobre la Teoría de Subconjuntos Difusos, realizadas por Zadeh [24] y Goguen [12,13], muestran la intención de los autores de crear una herramienta para modelar situaciones y fenómenos reales que, dado su comportamiento o complejidad, es imposible describirlos en forma precisa, donde la falta de precisión se refiere a la ausencia de una definición o criterio claro que permita asignar el valor de un parámetro y no al desconocimiento o al comportamiento aleatorio del valor mismo. En otras palabras, se trata de formalizar nociones que en sí mismas no es posible definirlas de forma precisa (como las nociones de "cercanía", "gran altura", "belleza", etcétera). En palabras de Zadeh

Principio de Incertidumbre

La complejidad de un sistema aumenta, así como nuestra habilidad para establecer su comportamiento de modo significativo y preciso disminuye; hasta que se llega a un umbral, más allá del cual, la precisión y el significado o relación, se convierten en características casi mutuamente exclusivas.

Es en este sentido que se explota la idea de multivaluar la pertenencia de un conjunto a otro, considerando entonces que el rango de la función característica de un conjunto no es un conjunto de dos elementos necesariamente sino que puede tener más elementos. En primera instancia el rango de la función de pertenencia se definió como el intervalo [0,1], para después generalizar esta idea a cualquier conjunto. En el presente trabajo primero expondremos los conceptos de la Teoría de Subconjuntos Difusos para los cuales el rango de las funciones características es el intervalo [0,1] y luego los generalizaremos para los tipos de rango concernientes a esta tesis.

Definición 1.1.[24] Sea Ξ un conjunto de puntos (objetos). Un subconjunto difuso $\tilde{A}$ de Ξ está caracterizado por un conjunto de pares ordenados

$$\tilde{A} = \{(x, \upsilon_{\tilde{A}}(x)) \;/\; x \in \Xi\}$$

donde $\upsilon_{\tilde{A}}: \Xi \to [0,1]$. $\upsilon_{\tilde{A}}$ es denominada **función de pertenencia** y para cada $x \in \Xi$ diremos que x pertenece con grado $\upsilon_{\tilde{A}}(x)$ al conjunto $\tilde{A}$. El rango de la función de pertenencia será denominado **el espacio de pertenencia**. De aquí en adelante Ξ será considerado como nuestro universo. ■

Funciones de pertenencia

Dada la definición anterior, vemos que cualquier función $\upsilon : \Xi \to [0,1]$ caracteriza o define un subconjunto difuso en Ξ, sin embargo, es común que en el campo de las aplicaciones la función de pertenencia (el conjunto) pretenda representar una propiedad. Cuando esta propiedad está referida a un concepto que no se define en forma precisa, se hace imposible que exista una única función de pertenencia que lo represente. Así pues, la pregunta ¿qué conjunto difuso representa cierta propiedad? en general carece de sentido. El objetivo entonces es encontrar algún subconjunto difuso que represente la propiedad adecuadamente en los términos de un problema en particular, aunque este subconjunto difuso tampoco sea único en ese sentido.

Veamos algunos ejemplos al respecto. Estos representan la propiedad "números reales cercanos a 10".

$$\tilde{A} = \{(x,\upsilon_{\tilde{A}}(x)) \;/\; [\upsilon_{\tilde{A}}(x) = 1/(1+(x-10)^2)] \wedge x \in \mathbb{R}\}$$
$$\tilde{B} = \{(x,\upsilon_{\tilde{A}}(x)) \;/\; [\upsilon_{\tilde{A}}(x) = (1+\varepsilon)/(1+\varepsilon+(x-10)^2)] \wedge x,\varepsilon \in \mathbb{R}\}$$
$$\tilde{C} = \{(x,\upsilon_{\tilde{A}}(x)) \;/\; [\upsilon_{\tilde{A}}(x) = 10 + e^{-kx}] \wedge k,x \in \mathbb{R}\}$$

Operaciones entre subconjuntos difusos.

Para la Teoría de Subconjuntos Difusos se han definido diversos tipos de operadores de unión, intersección y complementación que responden en general a las características de un problema concreto. Cada familia de estos operadores definen diferentes Teorías de Subconjuntos Difusos. A continuación presentamos un conjunto de operadores, que será la usada en el presente trabajo, que fue propuesta por Zadeh [24], la cual define la Teoría de Subconjuntos Difusos denotada como $(\tilde{P}(\Xi),\ \mathring{\cup},\ \mathring{\cap},\ \text{Ç})$ donde $\tilde{P}(\Xi)$ es el conjunto de todos los subconjuntos difusos de Ξ.

Definición 1.2. Dados dos conjuntos difusos $\tilde{A}$ y $\tilde{B}$ de Ξ, se definen las funciones de pertenencia de los conjuntos $\tilde{A}\mathring{\cup}\tilde{B}$ y $\tilde{A}\mathring{\cap}\tilde{B}$ respectivamente como

$$\forall\ x \in \Xi \quad \upsilon_{\tilde{A}\mathring{\cup}\tilde{B}}(x) = \max\{\upsilon_{\tilde{A}}(x),\ \upsilon_{\tilde{B}}(x)\}$$
$$\forall\ x \in \Xi \quad \upsilon_{\tilde{A}\mathring{\cap}\tilde{B}}(x) = \min\{\upsilon_{\tilde{A}}(x),\ \upsilon_{\tilde{B}}(x)\} \qquad \blacksquare$$

La justificación de esta elección fue dada por Bellman y Giertz [3] desde un punto de vista lógico, interpretando la intersección como un "y" lógico y la unión como un "o" lógico y al subconjunto difuso A como la proposición "x es elemento de A" cuyo valor de verdad no es necesariamente 0 ó 1. Así, dadas dos proposiciones "T" y "S" cuyos valores de verdad están definidos por las funciones $\upsilon_T,\upsilon_S:\Xi \to [0,1]$; el valor de verdad de "T y S" (υ_{TyS}) y "T o S" (υ_{ToS}) son interpretados como la función de pertenencia de la intersección y la unión de estos subconjuntos difusos. A su vez, demuestran que si los operadores de unión (f) e intersección (g) cumplen las siguientes propiedades, entonces estos deben ser **max** y **min**.

i) El valor de pertenencia de un elemento a la unión o intersección de conjuntos debe estar en función únicamente de los valores de pertenencia de ese elemento a los conjuntos, es decir

$$\forall\ x \in \Xi \quad \upsilon_{\tilde{A}\overset{\circ}{\cup}\tilde{B}}(x) = f(\upsilon_{\tilde{A}}(x),\ \upsilon_{\tilde{B}}(x))$$
$$\forall\ x \in \Xi \quad \upsilon_{\tilde{A}\overset{\circ}{\cap}\tilde{B}}(x) = g(\upsilon_{\tilde{A}}(x),\ \upsilon_{\tilde{B}}(x))$$

ii) f y g han de ser operadores conmutativos, asociativos y mutuamente distributivos.

iii) f y g deben ser operadores continuos y no decrecientes con respecto a sus argumentos. Es decir, intuitivamente un pequeño cambio en $\upsilon_{\tilde{A}}(x)$ o en $\upsilon_{\tilde{B}}(x)$ no puede inducir un gran cambio en $\upsilon_{\tilde{A}\overset{\circ}{\cup}\tilde{B}}(x)$ o en $\upsilon_{\tilde{A}\overset{\circ}{\cap}\tilde{B}}(x)$. Similarmente, si $\upsilon_{\tilde{A}}(x)$ o $\upsilon_{\tilde{B}}(x)$ aumentan, los valores de $\upsilon_{\tilde{A}\overset{\circ}{\cup}\tilde{B}}(x)$ y $\upsilon_{\tilde{A}\overset{\circ}{\cap}\tilde{B}}(x)$ no pueden decrecer.

iv) Si $\upsilon_{\tilde{A}}(x_1) = \upsilon_{\tilde{B}}(x_1) > \upsilon_{\tilde{A}}(x_2) = \upsilon_{\tilde{B}}(x_2)$ entonces debemos tener que $\upsilon_{\tilde{A}\overset{\circ}{\cup}\tilde{B}}(x_1) > \upsilon_{\tilde{A}\overset{\circ}{\cup}\tilde{B}}(x_2)$ y $\upsilon_{\tilde{A}\overset{\circ}{\cap}\tilde{B}}(x_1) > \upsilon_{\tilde{A}\overset{\circ}{\cap}\tilde{B}}(x_2)$

v) $\forall\ x \in \Xi \quad \upsilon_{\tilde{A}\overset{\circ}{\cup}\tilde{B}}(x) \geq \max\{\upsilon_{\tilde{A}}(x),\ \upsilon_{\tilde{B}}(x)\}$
$\quad\quad\quad\quad \upsilon_{\tilde{A}\overset{\circ}{\cap}\tilde{B}}(x) \leq \min\{\upsilon_{\tilde{A}}(x),\ \upsilon_{\tilde{B}}(x)\}$

vi) $g(1,1) = 1$ y $f(0,0) = 0$

La complementación definida por Zadeh [24] es la siguiente

Definición 1.3. Dado un subconjunto difuso $\tilde{A}$ de Ξ, la función de pertenencia del complemento de $\tilde{A}$, denotado como $Ç\tilde{A}$, se define como

$$\forall\ x \in \Xi\ \ \upsilon_{Ç\tilde{A}}(x) = 1-\upsilon_{\tilde{A}}(x) \qquad \blacksquare$$

Con respecto a la complementación de un subconjunto difuso no existe una justificación tan natural como en el caso anterior que nos defina un sólo operador, sin embargo, presentamos un conjunto de condiciones que debe cumplir este operador (h) propuestas también por Bellman y Giertz [3].

i) $\upsilon_{Ç\tilde{A}}(x)$ depende únicamente de $\upsilon_{\tilde{A}}(x)$, es decir $\upsilon_{Ç\tilde{A}}(x) = h(\upsilon_{\tilde{A}}(x))$

ii) $h(0)=1.\ h(1)=0.$

iii) h es continua y monótona estrictamente decreciente.

iv) $h(h(\upsilon_{\varsigma\tilde{A}}(x)))=\upsilon_{\tilde{A}}(x)$

Otras definiciones

Definición 1.4. A los subconjuntos cuya imagen de su función de pertenencia esté en el $\{0,1\}$ los denominaremos conjuntos duros. ■

Definición 1.5. Dado $\tilde{A}$ un subconjunto difuso de Ξ y $\alpha \in [0,1]$, el conjunto duro

$$\tilde{A}_{\alpha} = \{x \in \Xi \ / \ \upsilon_{\tilde{A}}(x) \geq \alpha \}$$

es denominado el α-corte de $\tilde{A}$. ■

Definición 1.6. Dado $\tilde{A}$ un subconjunto difuso de Ξ y $\alpha \in [0,1]$, el conjunto duro

$$\tilde{A}_{\alpha f} = \{x \in \Xi \ / \ \upsilon_{\tilde{A}}(x) > \alpha \}$$

es denominado el α-corte fuerte de $\tilde{A}$. ■

Definición 1.7.[17] Si dado un α-corte $\tilde{A}_{\alpha}$ existe $\beta>\alpha$ tal que $\tilde{A}_{\beta} \subsetneq \tilde{A}_{\alpha}$ entonces diremos que $\tilde{A}_{\alpha}$ es un α-corte no máximo. ■

Definición 1.8.[24] Dados $\tilde{A}$, $\tilde{B}$ subconjuntos difusos en Ξ, diremos que $\tilde{A}$ es subconjunto de $\tilde{B}$, denotado como $\tilde{A} \overset{\circ}{\subseteq} \tilde{B}$, si y sólo si

$$\forall \ x \in \Xi \quad \upsilon_{\tilde{A}}(x) \leq \upsilon_{\tilde{B}}(x)$$ ■

Definición 1.9.[24] Dados $\tilde{A}$, $\tilde{B}$ subconjuntos difusos en Ξ, diremos que $\tilde{A}$ y $\tilde{B}$ son iguales, denotado como $\tilde{A} \overset{\circ}{=} \tilde{B}$, si y sólo si

$$\forall \ x \in \Xi \quad \upsilon_{\tilde{A}}(x) = \upsilon_{\tilde{B}}(x)$$ ■

Relaciones difusas

Definición 1.10.[24] Sean Υ, Ψ conjuntos duros, entonces el subconjunto difuso $\tilde{R}$ en $\Upsilon \times \Psi$ definido como

$$\tilde{R} = \{((x,y),\ \upsilon_R(x,y))\ /\ (x,y) \in \Upsilon \times \Psi\}$$

donde $\upsilon_R\colon \Upsilon \times \Psi \to [0,1]$ es denominado relación difusa en $\Upsilon \times \Psi$. ■

Definición 1.11. Sean $\tilde{R}$ y $\tilde{Q}$ dos relaciones difusas sobre $\Upsilon \times \Psi$ y $\Psi \times Z$ respectivamente. La composición max-min de $\tilde{R}\circ\tilde{Q}$ es el conjunto difuso definido como

$$\tilde{R}\circ\tilde{Q} = \{((x,z),\ \upsilon_{\tilde{R}\circ\tilde{Q}}(x,z) = \max_{y\in\Psi}\{\min\{\upsilon_{\tilde{R}}(x,y),\upsilon_{\tilde{Q}}(y,z)\}\})\ /\ x \in \Upsilon, z \in Z\}$$ ■

Definición 1.12. Sea $\tilde{R}$ una relación difusa en $\Xi \times \Xi$.

$\tilde{R}$ es reflexiva si $\forall\ x \in \Xi\ \ \upsilon_{\tilde{R}}(x,x) = 1$ [25]

$\tilde{R}$ es simétrica si $\forall\ x,y \in \Xi\ \ \upsilon_{\tilde{R}}(y,x) = \upsilon_{\tilde{R}}(x,y)$

$\tilde{R}$ es max-min transitiva si y sólo si $\tilde{R}\circ\tilde{R} \overset{\circ}{\subseteq} \tilde{R}$ [25] ■

Definición 1.13.[25] $\tilde{R}$ es una relación difusa de similaridad si es reflexiva, simétrica y max-min transitiva. ■

Proposición 1.1.[25] Dada $\tilde{R}$ una relación de similaridad sobre Ξ, entonces para cualquier $\alpha \in (0,1]$, el α-corte de $\tilde{R}$ definirá una relación de equivalencia R sobre Ξ.

Demostración.

i) Por reflexividad de $\tilde{R}$, $\forall\ x \in \Xi$, $\upsilon_{\tilde{R}}(x,x) = 1$, por lo que $\forall\ \alpha \in (0,1]$, R es reflexiva.

ii) Por simetría de $\tilde{R}$, $\forall\ x,y \in \Xi$, $\upsilon_{\tilde{R}}(x,y)=\upsilon_{\tilde{R}}(y,x)$ por lo que $\forall\ \alpha \in (0,1]\ [\upsilon_{\tilde{R}}(x,y)\geq\alpha \Leftrightarrow \upsilon_{\tilde{R}}(y,x)\geq \alpha]$

por lo que $\forall\ \alpha \in (0,1]$, R es simétrica.

iii) Sea $\alpha \in (0,1]$, $x,y,z \in \Xi\ /\ \upsilon_{\tilde{R}}(x,y)\geq\alpha \wedge \upsilon_{\tilde{R}}(y,z)\geq\alpha$.

Por max-min transitividad de $\tilde{R}$ tenemos que $\upsilon_{\tilde{R}}(x,z)\geq\alpha$ por lo que R es transitiva. ■

Particiones

Definición 1.14.[20] Sea $\Xi = \{x_1, \ldots, x_n\}$ y $c \in \mathbb{Z}$ tal que $1 < c < n$. Una familia de c subconjuntos difusos $\{\tilde{A}_1, \ldots, \tilde{A}_c\}$ será denominada una c-partición difusa de Ξ si se cumple que

i) $\forall\ i \leq c\ \ \forall\ k \leq n\ \ \upsilon_{\tilde{A}_i}(x_k) \in [0,1]$

ii) $\forall\ k \leq n \qquad \sum_{i=1}^{c} \upsilon_{\tilde{A}_i}(x_k) = 1$

iii) $\forall\ i \leq c \qquad 0 < \sum_{k=1}^{n} \upsilon_{\tilde{A}_i}(x_k) < n$ ■

Notemos que cualquier partición de conjuntos duros cumple las propiedades *i)*, *ii)* y *iii)*. Al conjunto de las c-particiones difusas de Ξ lo denotaremos como M_{fcn}, y las representaremos como matrices reales de dimensión $c \times n$.

Definición 1.15.[17] Una familia de subconjuntos difusos $A = \{\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_c\}$ de un conjunto Ξ que para algún $\beta \in [0,1]$ cumple:

i) $\forall\ x \in \Xi\ \exists!\ \tilde{A}_i \in A\ [\upsilon_{\tilde{A}_i}(x) \geq \beta]$

ii) $\forall\ \tilde{A}_i \in A\ \exists\ x \in \Xi\ [\upsilon_{\tilde{A}_i}(x) \geq \beta]$

será denominada una β-partición difusa de Ξ. ■

Definición 1.16.[17] Una familia de subconjuntos difusos $A = \{\tilde{A}_1, \tilde{A}_2, \ldots, \tilde{A}_c\}$ de un conjunto Ξ que para algún $\beta \in [0,1]$ cumple:

i) $\forall\ x \in \Xi\ \exists\ \tilde{A}_i \in A\ [\upsilon_{\tilde{A}_i}(x) \geq \beta]$

ii) $\forall\ \tilde{A}_i \in A\ \exists\ x \in \Xi\ [\upsilon_{\tilde{A}_i}(x) \geq \beta]$

será denominada un β-cubrimiento difuso de Ξ. ■

Generalización del concepto de subconjunto difuso.

Sean Ξ y L conjuntos duros.

Definición 1.17.[12] Un L-subconjunto difuso $\tilde{A}$ de Ξ está caracterizado por un conjunto de pares ordenados

$$\tilde{A}=\{(x,\upsilon_{\tilde{A}}(x)) \ / \ x \in \Xi\}$$

donde $\upsilon_{\tilde{A}}:\Xi \to L$. L es denominado el espacio de pertenencia. ∎

En el caso en que L tenga asociada una estructura, como la de retícula o de grupo, esta se heredará a $\tilde{P}(\Xi)$.

En lo subsecuente, cuando hablemos de subconjuntos difusos nos estaremos refiriendo a L-subconjuntos difusos para los cuales L tiene como estructura asociada un orden total con un elemento máximo (1) y un mínimo (0).

A partir de esto podemos extender las definiciones dadas hasta ahora excepto la de complementación que involucra una estructura más compleja que el orden y la comparabilidad, por lo que respecto a ésta tan sólo consideraremos que

i) $\forall\ x \in \Xi\ \upsilon_{\text{ÇÇ}\tilde{A}}(x) = \upsilon_{\tilde{A}}(x)$
ii) $\upsilon_{\tilde{A}}(x) = 1$ si y sólo si $\upsilon_{\text{Ç}\tilde{A}}(x) = 0$.

Lógica Multivalente Asociada a la Teoría de Subconjuntos Difusos
$(P(\Xi),\ \mathring{\cup},\ \mathring{\cap},\ \text{Ç})$.
Dada una proposición P, denotaremos como $\upsilon(P) \in L$ al valor de verdad de la misma. Así tenemos que si P y Q son proposiciones, entonces

i) $\upsilon(P \vee Q)= \max\{\upsilon(P),\ \upsilon(Q)\}$
ii) $\upsilon(P \wedge Q)= \min\{\upsilon(P),\ \upsilon(Q)\}$
iii) $\upsilon(\neg\neg P) = \upsilon(P)$
iv) $\upsilon(P) = 1$ si y sólo si $\upsilon(\neg P) = 0$.
v) $\upsilon(\forall\ x \in A\ P(x)) = \min\{\upsilon(P(x))\}$ con $x \in A \subseteq \Xi$[1]

[1] En el presente trabajo consideramos que Ξ es un conjunto finito, por lo que se asegura la existencia de $\upsilon(\forall\ x\ P(x))$.

En el caso en que $L= [0,1]$ entonces tendremos además que

vi) $\upsilon(\neg P)= 1-\upsilon(P)$.

Bajo esta perspectiva, dada una propiedad P podemos definir el subconjunto difuso $\tilde{A}$ como

$$\tilde{A} =\{(x,\upsilon(P(x))) \ / \ x \in \Xi\}$$

donde la pertenencia al subconjunto $\tilde{A}$ será interpretada como el "cumplimiento de la propiedad P por parte de $x \in \Xi$".

Definición 1.18. Dada Q una proposición tal que

$$\upsilon(Q) = 1 \vee \upsilon(Q) = 0$$

diremos que Q es una **proposición dura**. ∎

Proposición 1.2. Sea P una propiedad, A un conjunto tal que $A=\{x \ / \ \upsilon(P(x))\geq\alpha\}$ y $\alpha \neq 0$. Si A es el α-corte no máximo del subconjunto difuso $\tilde{A} =\{(x,\upsilon(P(x))) \ / \ x \in \Xi\}$, entonces existe $x \in \Xi$ tal que la proposición $P(x)$ es una proposición no dura. ∎

"Mayor o igual que" difuso

Presentamos aquí una familia de subconjuntos difusos que será básico para el desarrollo del presente trabajo.

El concepto "mayor que" es una relación que, dado un orden en un conjunto, un elemento cumple o no con respecto a otro. Sin embargo, cuando en la práctica el tamaño, dimensión o diferencia entre los elementos es información que se quiere que esté presente, entonces intuitivamente, diferentes elementos del conjunto se pueden considerar más "mayores que" que otros o más "menores que" que otros con respecto a un elemento dado, en el mismo sentido que existen elementos más cercanos que otros a un elemento dado. Por ejemplo, es verdad que $11\geq10$ y $100\geq10$, sin embargo, podemos pensar que 100 es *más* "mayor que" 10 que lo que puede ser el 11. Tratando de reflejar estas ideas damos la siguiente

Definición 1.19.[17] Sea $\bar{F}$ un subconjunto difuso sobre L definido como

$$\bar{F}=\{(\beta,\upsilon_{\bar{F}}(\beta)) \text{ con } \beta \in L\}$$

Si existe $\alpha \in L$ tal que, dados β, $\delta \in L$, υ_F cumple

i) Si $\delta\geq\alpha>\beta$ entonces $\upsilon_{\bar{F}}(\delta) > \upsilon_{\bar{F}}(\beta)$

ii) Si $\delta>\beta\geq\alpha$ ó $\alpha>\delta>\beta$ entonces $\upsilon_{\bar{F}}(\delta) \geq \upsilon_{\bar{F}}(\beta)$

entonces diremos que $\tilde{F}$ pertenece a la clase de subconjuntos Mayor o igual que α difuso, denotada como $D(\geq\alpha)$. ■

Para mayor claridad, si $\tilde{F} \in D(\geq\alpha)$, entonces para cada $\beta \in L$ denotaremos a $\upsilon_{\bar{F}}(\beta)$ como $\upsilon_{\bar{F}}(\beta\geq\alpha)$.

A la luz de esto último damos la siguiente definición, la cual nos relaciona a cada conjunto duro con una familia de subconjuntos difusos a los cuales denominaremos sus seudodifusiones.

Definición 1.20.[17] Sea $\tilde{F} \in D(\geq\alpha)$ y sea $\bar{A}$ un subconjunto difuso de Ξ definido con base en una propiedad P de la manera siguiente:

$$\bar{A}=\{(x, \upsilon(P(x)) \ / \ x \in \Xi\}$$

donde para cada $x \in \Xi$, $\upsilon(P(x)) \in L$ es el valor de verdad de la proposición "x cumple la propiedad P". Sea C el conjunto duro definido como un α-corte de $\tilde{A}$, es decir

$$C = \tilde{A}_\alpha = \{x \ / \ \upsilon(P(x)) \geq \alpha \in L\}$$

Al subconjunto L-difuso $\bar{D}(C)$ definido como

$$\bar{D}(C)=\{(x, \ \upsilon_{\bar{F}}(\upsilon(P(x))\geq\alpha)) \ / \ x \in \Xi\}$$

lo denominaremos una seudodifusión de C. ■

Notemos que si para cada $\beta \in L$, $\upsilon_{\tilde F}(\beta \geq \alpha) = \beta$ entonces $\tilde D(C) = \tilde A$, por lo que $\tilde A$ es también una seudodifusión de C. La función de pertenencia que define a $\tilde D(C)$ es la representación del concepto "que tanto pertenece x a C, para cada $x \in \Xi$", definido sobre una medida de "que tan mayor es $\upsilon(P(x))$ con respecto a α".

Proposición 1.3. Sea C un α-corte de un subconjunto difuso $\tilde A$. Dada cualquier seudodifusión $\tilde D(C)$ de C definida con algún $\tilde F \in D(\geq\alpha)$ tenemos que

a) $\forall\ y,x \in \Xi$

si $\upsilon_{\tilde A}(x) = \upsilon(P(x)) > \upsilon(P(y)) = \upsilon_{\tilde A}(y)$ entonces

$\upsilon_{\tilde D(C)}(x) = \upsilon_{\tilde F}(\upsilon(P(x)) \geq \alpha) \geq \upsilon_{\tilde F}(\upsilon(P(y)) \geq \alpha) = \upsilon_{\tilde D(C)}(y)$

b) Si $\upsilon_{\tilde F}(\alpha \geq \alpha) = \beta$ entonces

$$x \in C \Leftrightarrow \upsilon_{\tilde D(C)}(x) \geq \beta.$$

Demostración.

a) Por definición de $D(\geq\alpha)$.

b) Si $x \in C$ entonces $\upsilon_{\tilde A}(x) \geq \alpha$ por lo que

$\upsilon_{\tilde D(C)}(x) = \upsilon_{\tilde F}(\upsilon_{\tilde A}(x) \geq \alpha) \geq \beta$

Si $x \notin C$ entonces $\upsilon_{\tilde A}(x) < \alpha$ por lo que

$\upsilon_{\tilde D(C)}(x) = \upsilon_{\tilde F}(\upsilon_{\tilde A}(x) \geq \alpha) < \beta$ ■

Es decir, de *a*) vemos que una seudodifusión $\tilde D(C)$ es un mapa del orden, no necesariamente estricto, de la función de pertenencia del subconjunto difuso $\tilde A$, y por ende, del cumplimiento de la propiedad P por parte de los elementos de Ξ.

En el caso en que para toda $x \in \Xi$, P(x) sea una proposición dura tendremos que

$\forall\ x,y \in C\ \upsilon_{\tilde D(C)}(x) = \upsilon_{\tilde D(C)}(y)$

$\forall\ x,y \notin C\ \upsilon_{\tilde D(C)}(x) = \upsilon_{\tilde D(C)}(y)$

$\forall\ y,x \in \Xi$

si $x \in C \wedge y \notin C \Rightarrow \upsilon_{\tilde D(C)}(x) > \upsilon_{\tilde D(C)}(y)$

Definición 1.21.[17] Sea $A = \{A_1, A_2, \ldots, A_c\}$ una partición (o cubrimiento) de subconjuntos duros de Ξ, donde cada A_i es el α-corte del subconjunto difuso $\tilde{A}_i$. Sea $\tilde{F} \in D(\geq\alpha)$. Entonces a la familia de subconjuntos difusos $B = \{\tilde{B}_1, \tilde{B}_2, \ldots, \tilde{B}_c\}$ tal que cada $\tilde{B}_i$ es la seudodifusión según $\tilde{F}$ de A_i, será denominada una seudodifusión de la partición (cubrimiento) A. ■

Proposición 1.4. Sea $A = \{A_1, A_2, \ldots, A_c\}$ una partición (o cubrimiento) de subconjuntos duros de Ξ donde cada A_i es el α-corte del subconjunto difuso $\tilde{A}_i$. Sea $B = \{\tilde{B}_1, \tilde{B}_2, \ldots, \tilde{B}_c\}$ una seudodifusión de la partición (cubrimiento) A, definida a partir de cierta $\tilde{F} \in D(\geq\alpha)$. Sea $\upsilon_{\tilde{F}}(\alpha\geq\alpha) = \beta$. Entonces B es una β-partición (cubrimiento) difusa de Ξ.

Demostración.

Por definición, cada $\tilde{B}_i$ es una seudodifusión de cada A_i definida a partir de $\tilde{F}$, que por la proposición 1.3 *b*) implica que para cada $x \in \Xi$ y para cada $A_i \in A$ tenemos

$$x \in A_i \Leftrightarrow \upsilon_{\tilde{B}_i}(x) \geq \beta$$

Si A es una partición de Ξ entonces

i) $\forall\ x \in \Xi\ \exists!\ A_i \in A\ [x \in A_i] \Leftrightarrow \forall\ x \in \Xi\ \exists!\ \tilde{B}_i \in B\ [\upsilon_{\tilde{B}_i}(x) \geq \beta]$

ii) $\forall\ A_i \in A\ \exists\ x \in \Xi\ [x \in A_i] \Leftrightarrow \forall\ \tilde{B}_i \in B\ \exists\ x \in \Xi\ [\upsilon_{\tilde{B}_i}(x) \geq \beta]$

Por tanto B es una β-partición difusa de Ξ.

La demostración en el caso de que A sea un cubrimiento es análoga. ■

CAPITULO II

RECONOCIMIENTO DE PATRONES

Introducción

El problema de Reconocimiento de Patrones se puede ubicar, en principio, como un intento de automatizar, formalizar o modelar una actividad característica del ser humano que es, justamente, la de poder reconocer la existencia de patrones de algún tipo en un conjunto de objetos pertenecientes a una misma categoría.

Es decir, dado un conjunto de objetos que pueden ser descritos a partir de un mismo conjunto de rasgos, el ser humano tiene la capacidad de encontrar, a partir de algún criterio, estructuras en el conjunto que responden al comportamiento particular o global de las semejanzas entre los objetos (simetrías, distancias, tendencias, aglomeraciones, coincidencias, etcétera), reveladas estas semejanzas a partir de los rasgos que los describen, y así: crear en el conjunto diferentes subcategorías (agrupaciones de objetos) definidas con base en las estructuras; o bien generalizar una división en subcategorías ya presente en el conjunto de objetos a toda la categoría a la que pertenecen; o bien discriminar entre los diferentes rasgos que describen a los objetos encontrando un subconjunto de estos que caracterizan el comportamiento de los objetos entre sí.

Así pues, el problema de reconocimiento matemático de patrones se puede plantear, de manera muy general, de la forma siguiente: dado un conjunto de objetos del mundo real que son comparables todos entre sí a partir de ciertos rasgos, modelar matemáticamente objetos, rasgos, espacios donde estan ubicados estos objetos, semejanzas y criterios, y buscar a partir del modelo estructuras subyacentes en el conjunto real de objetos y así definir agrupaciones de los mismos (análisis de cúmulos), o seleccionar los rasgos que caracterizan el comportamiento de las semejanzas entre los objetos (selección de rasgos) o bien generalizar una división dada de antemano a todo el espacio al que pertenecen los objetos (problemas con aprendizaje).

Con la incorporación de la Teoría de Subconjuntos Difusos en la aplicación a problemas de reconocimiento de patrones [4], fundamentada ésta sobre todo a partir del Principio de Incompatibilidad, se abrió la posibilidad de ser más certero en la modelación de la actividad humana que genera las semejanzas entre objetos, los criterios con que los agrupa, y en la forma de agrupar a los objetos reales, permitiendo que la etiquetación que se hace de las descripciones de objetos en subcategorías se vuelva, en muchos casos, más real.

Ahora veamos de manera general los pasos del problema de reconocimiento matemático de patrones:

1. Modelación del Espacio de Representación $\Phi(\Xi, \Gamma)$

Dados un conjunto de N objetos del mundo real y un conjunto finito M de rasgos que describe a esos objetos, hemos de definir:

i) Valores Admisibles de los Rasgos.

Para el i-ésimo rasgo, el conjunto $<M_i,R_i>$ el cual será un conjunto totalmente ordenado que representará al conjunto de valores admisibles del i-ésimo rasgo (M_i) y la relación de orden entre estos valores (R_i).

ii) Elementos del espacio de representación.

Definimos $\Delta=\prod_{i=1}^{m} M_i$, donde cada M_i es el conjunto de valores admisibles del i-ésimo rasgo y $\prod$ el producto cartesiano.

Así, a cada objeto real lo representaremos en Δ con el elemento $O_i = <x_1^i, \ldots, x_m^i>$ siendo $x_r^i \in M_r$ el valor que toma el i-ésimo objeto en el r-ésimo rasgo.

iii) Modelación de las Relaciones entre Objetos.

Este problema consiste en modelar matemáticamente la forma en que se hacen las comparaciones reales entre los objetos reales, es decir, tiene como objetivo el plantear una función

$$\Gamma:\Delta\times\Delta \rightarrow L$$

donde L es un conjunto cuyos elementos son comparables entre sí a partir de un orden total, siendo estos elementos los que interpretaremos como la medida de la semejanza entre cada par de objetos. De esta manera dotamos de estructura (no necesariamente normada) al espacio de representación.

Notemos que la estructura del espacio de representación no tiene por qué ser definida a partir de la descripción de los elementos del mismo, es decir, aunque las descripciones de los objetos de Δ pudieran ser ubicadas en $\mathbb{R}^n$, esto no implicaría que tiene sentido el pensar que el espacio de representación es un espacio métrico. Si existe algún tipo de propiedad, ésta tendrá que surgir a partir de la modelación que hagamos de las relaciones entre los objetos reales.

En este sentido, el problema medular de esta modelación es el hecho de que el concepto o idea de semejanza entre objetos del mundo real se nutre en general de muchas formas y referencias tales como cercanía, analogía, parecido, similitudes parciales, distancia, etecétera.

Como ejemplos tenemos definiciones de semejanzas a partir de:

a) Una métrica o norma cuando las descripciones lo permiten [5,8].

b) Coincidencias entre los valores de los rasgos que describen a los objetos [19].

c) Coincidencias entre ciertos subconjuntos de rasgos de las descripciones de los objetos [19].

Etcétera.

2. Definición de los Criterios y Estructuras

Este problema consiste en la definición de las propiedades o reglas que deben cumplir las estructuras que se generarán en el espacio de representación. En otras palabras, se debe definir qué y cómo es un comportamiento o propiedad entre los objetos que sea significativa para el problema en cuestión, es decir, se debe plantear un criterio agrupacional. Como ejemplos de este tipo de reglas o propiedades que resultan significativas nos encontramos con:

a) Máxima semejanza de un objeto a otro creando agrupaciones conexas [19,21].

b) Cierto grado de semejanza entre objetos creando agrupaciones conexas [19,23].

c) Cierto grado de semejanza entre todos los objetos de cada agrupación [1,15].

d) Cierto grado de semejanza entre cierta cantidad de objetos en cada agrupación [16].

e) Máxima semejanza posible entre los objetos de una agrupación y máxima diferencia posible con los objetos de otras agrupaciones [5].

Etcétera.

Tipos de Problemas de Reconocimiento de Patrones

Dentro del problema general de reconocimiento de patrones, nos encontramos con cuatro grandes clases de ellos: reducción del espacio de representación, reconocimiento de patrones con aprendizaje, reconocimiento de patrones con aprendizaje parcial y reconocimiento de patrones sin aprendizaje. Este último es parte medular de este trabajo por lo que le dedicaremos mayor atención. En cuanto a los otros tres, daremos tan sólo una breve descripción de los mismos.

1. Reducción del Espacio de Representación.

La reducción del espacio de representación, donde $\Delta=\prod_{i=1}^{m}M_i$, es la búsqueda de proyecciones $\prod_{j=1}^{s}M_{i_j}$ con $s\leq m$ de dimensión mínima del espacio que contengan toda o la máxima información posible al respecto de Γ que se tenía en el espacio inicial.

Esta búsqueda puede tener distintos fines. Por un lado, el lograr una aplicación eficiente de la modelación al tener que manipular dimensiones lo más pequeñas posibles; por otro lado, como se dijo antes, se busca un conjunto mínimo de rasgos que *definen o caracterizan el fenómeno en cuestión.* El llegar a tener una buena aproximación a la solución de este problema hace que las estructuras en el conjunto de objetos sean caracterizadas directamente a partir de este conjunto de rasgos.

Sin embargo, el problema particular de la reducción del espacio de representación está realmente resuelto sólo en el caso en que tenemos de antemano las agrupaciones de los objetos del conjunto inicial.

2. Reconocimiento de Patrones con Aprendizaje

El problema de reconocimiento de patrones con aprendizaje se puede plantear como el dar una familia de agrupaciones (duras o difusas) que involucren a todo el espacio de representación $\Phi(\Delta,\Gamma)$, es decir, hemos de etiquetar a todos los elementos del espacio. En general se tiene de antemano una clasificación en subcategorías (agrupaciones) de un subconjunto de objetos del espacio de representación (matriz de aprendizaje), y entonces se define un criterio a partir del cual se generaliza esa clasificación a todo el espacio. Es decir, podemos plantear que el problema de reconocimiento de patrones con aprendizaje consiste en un proceso de generalización de los agrupaciones ya existentes de los objetos de la matriz de aprendizaje a nuevos objetos que no pertenecían a la matriz, habiendo sido creados estas agrupaciones por un proceso de abstracción a partir del conjunto de objetos de la matriz de aprendizaje. La resolución de este problema puede tener diferentes formas. Como ejemplos tenemos la reducción de dimensión del espacio de representación de tal forma que cada subcategoría quede caracterizada por ciertos valores de algunos rasgos y entonces cotejar estos rasgos con los de los objetos del espacio de representación [19]. Otra forma es considerar para cada subcategoría uno o varios objetos prototipos o más representativos de cada agrupación, y medir la semejanza entre ellos y los objetos [8], etcétera.

3. Reconocimiento de Patrones con Aprendizaje Parcial

El problema de reconocimiento de patrones con aprendizaje parcial se plantea cuando en la clasificación dada de antemano de un subconjunto de objetos del espacio de representación $\Phi(\Delta,\Gamma)$, sucede que para una de las agrupaciones la información con respecto a ésta es mucho menor a la información acerca de las otras, y la caracterización de la misma se da más en función de la no pertenencia de sus objetos a las otras agrupaciones que a propiedades independientes de las otras agrupaciones que compartan los objetos de esta agrupación. Como ejemplo de esto tenemos los casos de búsqueda de minerales en diferentes zonas geológicas. Los estudios y descripciones de las zonas donde no existan los minerales buscados será mucho más pobre que en los lugares en que sí se encuentran [19].

En este caso es cuando para el problema de ubicar o relacionar un nuevo objeto en las clases, no es posible usar directamente los métodos de reconocimiento de patrones con aprendizaje ya que muchas veces no es clara la manera de verificar a que clase pertenece más ese objeto, pues para una de las clases no existe mucha información propia. Es por ello que este tipo de problemas se plantea separado del anterior.

4. Reconocimiento de Patrones sin Aprendizaje

El problema de reconocimiento de patrones sin aprendizaje, también conocido como análisis de cúmulos, se puede plantear, de manera muy general, como sigue:

Dado un conjunto de descripciones de objetos, subconjunto de un espacio de representación $\Phi(\Delta,\Gamma)$, hemos de identificar a las descripciones en diferentes subgrupos, donde los subgrupos responden o se generan de manera "natural" según el comportamiento global o particular de las semejanzas entre las descripciones de los objetos atendiendo a una cierta propiedad. En otras palabras, podemos plantear que el análisis de cúmulos es la búsqueda de estructuras en un subconjunto del espacio de representación[2]. Partiendo del hecho de que no se tiene *a priori* ninguna información al respecto de los subgrupos o estructuras, la solución al problema del análisis de cúmulos inicia con ciertas suposiciones o criterios que nos describan cómo o qué es un subgrupo que responde de manera natural al comportamiento de las semejanzas entre los objetos, o en otras palabras, cómo son las estructuras que estamos buscando en el subconjunto de las descripciones. Visto de esta manera, el criterio es la *razón* por la cual una descripción va a pertenecer a un agrupamiento o el *por qué* dos descripciones pertenecerán a un mismo cúmulo. Todo esto nos lleva a que la selección del criterio a usar es crucial en la calidad de la solución del problema de análisis de cúmulos. Con todo esto observamos que la definición del criterio debe estar basada en el conocimiento que se tenga al respecto del problema en concreto que se está tratando, para poder definir así el tipo de comportamiento entre las descripciones a partir de sus semejanzas que resulte, según el problema en particular, significativo. Nótese que al seleccionar algún criterio, dado un conjunto de descripciones, hemos ya definido indirectamente a la familia de agrupaciones.

Tras la definición del criterio a usar, se presenta el problema computacional del análisis de cúmulos, es decir, la realización algorítmica eficiente de estos criterios para la aplicación de los mismos en problemas concretos, pero este será un tópico que en el presente trabajo no tocaremos.

Modelos de algoritmos de análisis de cúmulos

Tras todo lo anterior, podemos pensar que un modelo de algoritmos de análisis de cúmulos ideal para resolver el problema planteado

[2] En palabras de J. C. Bezdek "The search for structure in data sets, or samples $S \subsetneq \Xi$".

en la introducción debería de poder adecuarse al tipo de descripciones de objetos; a la manera en que se haya definido la semejanza entre las descripciones; al tipo de estructura o criterio que se defina según el problema en particular; no requiera definir de antemano el número de agrupaciones; y que genere un conjunto de agrupamientos que revelen de la manera más informativa el comportamiento de los objetos. Es decir, debe de poder manipular

i) Cualquier tipo de descripción de objetos.
ii) Cualquier tipo de medida de semejanza entre descripciones.
iii) La definición de cualquier tipo de estructuras o criterio.
iv) No requerir la predefinición del número de agrupaciones.
v) Generar familias de agrupaciones que describan lo más finamente posible el comportamiento de las descripciones con respecto al criterio usado.

Desde esta perspectiva haremos una revisión del modelo de algoritmos Fuzzy c-means, que es el modelo más ampliamente usado y alrededor del cual se han producido una gran cantidad de generalizaciones, mejoras y nuevos algoritmos. Este modelo es una generalización de la extensión al caso difuso del algoritmo c-means [2,8].

Modelo Fuzzy c-means[5]

Sea $\Delta = \mathbb{R}^p$, $\Xi = \{O_1, \ldots O_n\} \subseteq \mathbb{R}^p$ el conjunto de descripciones, y $\Gamma = \|\cdot\|$ una norma inducida por un producto interno para $\mathbb{R}^p$.

El modelo de algoritmos Fuzzy c-means plantea como solución al problema de análisis de cúmulos la minimización del funcional

$$J_m(U,v) = \sum_{k=1}^{n} \sum_{i=1}^{c} (\upsilon_{ik})^m (d_{ik})^2$$

donde $U = \{\upsilon_{ik}\} \in M_{fcn}$[3], $v = (v_1, \ldots, v_c) \in \mathbb{R}^{pc}$; $d_{ik} = \|O_k - v_i\|$ y

[3] M_{fcn} es el conjunto de c-particiones difusas de Ξ con $|\Xi| = n$ (Definición 1.14. página 7).

$m \in [1, \infty)$. En otras palabras, busca minimizar la norma entre los centros y los objetos ponderada ésta por la m-ésima potencia de la pertenecia de los objetos a los agrupamientos. Un algoritmo de este modelo se plantea de la manera siguiente:

1.- Se fija c el número de agrupaciones a formar, se escoge $\|\cdot\|$ una norma definida por un producto interno para $\mathbb{R}^p$, se fija $m \geq 1$. Se toma $U^0 = \{u^0_{ik}\} \in M_{fcn}$, una c-partición difusa como partición inicial.

Para el paso l, con l= 0, 1,..., tenemos que:

2.- Se calculan los c centros $v^l = \{v^l_1, \ldots, v^l_c\} \subseteq \mathbb{R}^p$ a partir de U^l y

$$v^l_i = \sum_{k=1}^{n} (u^l_{ik})^m \, O_k \,/ \sum_{k=1}^{n} (u^l_{ik})^m$$

3.- Se define U^{l+1} como sigue:

Definimos para cada $k \leq n$ los conjuntos

$$I_k = \{\, i \,/\, 1 \leq i \leq c;\ d_{ik} = \|O_k - v_i\| = 0\}$$

$$I_k' = \{\, 1, 2, \ldots, c \,\} \setminus I_k$$

y así, si $I_k = \emptyset$ entonces

$$u^{l+1}_{ik} = 1/[\sum_{j=1}^{c} (d_{ik}/d_{jk})^{2/(m-1)}]$$

y si $I_k \neq \emptyset$ entonces

$$\forall\ i \in I_k' \quad u^{l+1}_{ik} = 0 \quad y \quad \sum_{i \in I_k} u^{l+1}_{ik} = 1$$

4.- Se compara U^l y U^{l+1} a partir de alguna norma matricial y umbral convenientes. Si $\|U^l - U^{l+1}\| \leq \varepsilon$ termina. En otro caso se regresa a 2.- con l=l+1.

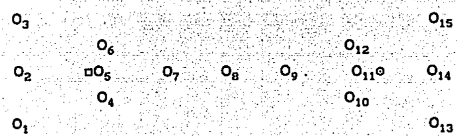
El parámetro m controla la difusión de la c-partición resultante en el sentido de que mientras $m \to \infty$ entonces los valores u_{ik} tenderán a 1/c, es decir, la c-partición será más "difusa"; y si $m \to 1$ entonces los u_{ik} tenderán a 1 ó a 0. Por otro lado, dadas las definiciones antes mencionadas para la obtención de v^l y U^l se tiene que la sucesión $\{(U^l,v^l) \ / \ l = 0,1, \ldots\}$ dado $m>1$ converge o contiene una subsucesión que converge a un mínimo local del funcional J_m [6,26].

Existen diversas dificultades para la implementación de este modelo en el tipo de problemas con las condiciones planteadas en la introducción. En primer lugar se requiere que el espacio de representación sea normado lo cual implica que las descripciones de los objetos han de ser en términos cuantitativos y además se necesita que la función de semejanza sea una norma. También es necesaria la predefinición del número de agrupamientos y una ubicación inicial de los centros, además de que el resultado final (la c-partición difusa resultante) dependerá de esta ubicación inicial. Por otro lado, el criterio agrupacional (minimizar la norma entre los centros y los objetos ponderada por la m-ésima potencia de la pertenecia) define agrupaciones de forma hiperesférica, y la interpretación del parámetro *m*, que controla la "difusión" del resultado, no es muy clara.

Diversas mejoras y generalizaciones se han implementado para este modelo de algoritmos, tales como: la utilización de distintas normas para cada agrupamiento logrando que cada agrupación tenga distinto corte hiperelipsoidal [14], el cálculo del número de agrupamientos *c* y la ubicación de los centros de los mismos a partir de criterios estadísticos [11], o la extensión de la noción de hiperelipse para el reconocimiento de agrupaciones de forma lineal [11]. Sin embargo el uso de estas mejoras también implica que el espacio de representación sea normado.

Por otro lado, al igual que el modelo Fuzzy c-means, casi todos los algoritmos para el análisis de cúmulos difusos existentes presentan como solución al problema de análisis de cúmulos un elemento de M_{fcn}, sin embargo, al respecto es importante notar lo siguiente. Veamos un ejemplo.que es resultado de la aplicación del algoritmo Fuzzy c-means basado en la norma euclidiana buscando 2 agrupaciones del siguiente conjunto de elementos del plano [7]. Sea $\Xi = \{O_1, \ldots O_{15}\} \subseteq \mathbb{R}^2$ donde la ubicación de los objetos esta dada como

Objeto	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x	0	0	0	1	1	1	2	3	4	5	5	5	6	6	6
y	0	2	4	1	2	3	2	2	2	1	2	3	0	2	4



La partición difusa resultante es la siguiente:

$$\begin{array}{ccccccccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ \end{array}$$
$$\begin{bmatrix} .99 & 1 & .99 & 1 & 1 & 1 & .99 & .\underline{47} & .01 & 0 & 0 & 0 & .01 & 0 & .01 \\ .01 & 0 & .01 & 0 & 0 & 0 & .01 & .\underline{53} & .99 & 1 & 1 & 1 & .99 & 1 & .\underline{99} \end{bmatrix}$$

donde la primera fila son las pertenencias de los objetos al agrupamiento con centro $\square$ =(.85, 2) y la segunda al agrupamiento con centro $\circ$ = (5.14, 2). Vemos que con respecto al punto O_8 la partición difusa nos dice que la pertenencia de este punto a ambos agrupamientos es muy parecida, lo cual es natural, sin embargo, notemos que la distancia de O_8 al centro $\circ$ es menor que la de O_{15}, lo cual debería implicar, si la partición representa la pertenencia de los objetos a los agrupamientos, que la pertenencia (.53) de O_8 al agrupamiento fuera mayor o igual a la de O_{15} (.99). Por esto podemos ver que la c-partición difusa no nos da ninguna información al respecto de la pertenencia de los diferentes objetos a un mismo agrupamiento. En otras palabras, los valores de pertenencia de los diferentes objetos no son comparables y por ello no podemos extraer de estos agrupamientos ninguna otra información referente al problema en particular. Lo que tenemos en una c-partición difusa es, para cada objeto, las relaciones existentes entre sus pertenencias a los diferentes agrupamientos.

Algunos criterios agrupacionales duros y funciones de semejanza
Ahora presentaremos una serie de conceptos en torno a la función de semejanza que permitirá al nuevo modelo generar en particular algoritmos de análisis de cúmulos que cumplen con las condiciones de los problemas planteados en la introducción del presente trabajo. También comentaremos las características de los criterios agrupacionales duros que fueron el punto de partida para la definición del nuevo modelo.

Funciones de semejanza
Consideremos $M = \{r_1, \ldots, r_m\}$ el conjunto de rasgos; $(O_1, \ldots, O_n)$ el conjunto de descripciones de objetos de Ξ donde cada descripción O_i es la m-upla $<x_i^1, \ldots, x_i^m>$ siendo x_i^r el valor que toma el i-ésimo objeto en el r-ésimo rasgo.

Definición 2.1.[19] Sea $\Omega = \prod_{j=1}^{s} M_{t_j}$ una proyección de Δ. Por **ω-parte** de una descripción O_i entenderemos la subdescripción $\omega(O_i) = <x_i^{t_1}, \ldots, x_i^{t_s}>$. A la proyección Ω la denominaremos **ω-parte**. ■

Definición 2.2.[19] Por **sistema de conjuntos de apoyo** Ω_A entenderemos un conjunto de ω-partes. ■

El sistema de conjuntos de apoyo permite introducir la idea de precedencia parcial entre descripciones, es decir, poner de relieve la importancia o preponderancia que pueden tener ciertos subconjuntos de rasgos en cuanto a la semejanza entre las descripciones de los objetos.

Definición 2.3.[19] Un criterio de comparación de valores de una variable (rasgo) r_i es una función $C_i : M_i \times M_i \longrightarrow T_i$, siendo M_i el conjunto de valores admisibles de dicha variable, tal que

$$\forall\, x,y \in M_i\ [C_i(x,y) = C_i(y,x)]$$

La imagen de esta función es algún $T_i \subseteq [0,1]$, para cada $i \leq m$, en dependencia de la naturaleza del criterio de comparación, estos criterios se denominarán booleanos, k-valentes, o reales[4]. ■

Otras características que pueda tener la función C_i dependerán de la modelación particular del problema planteado. La familia $\{C_1, \ldots, C_m\}$ de criterios de comparación permiten el tratamiento diferenciado, no uniforme de los distintos rasgos que describen a los objetos, y por tanto facilita el tratamiento simultáneo de variables cuantitativas y cualitativas en las descripciones de los objetos, además de admitir la ausencia de información en las descripciones de algunos de los mismos.

Sea $\Omega_A = \{\Omega_1, \ldots, \Omega_q\}$ un sistema de conjuntos de apoyo. Sea para cada $s \leq q$, $\Omega_s = M_{s_1} \times \ldots \times M_{s_p}$ donde la subsucesión $\{s_1, \ldots, s_p\}$ depende de s. En lo sucesivo nos referiremos indistintamente a r_i ó a M_i.

Definición 2.4.[19] Sea $\Omega_s = M_{s_1} \times \ldots \times M_{s_p} \in \Omega_A$. Una **función de semejanza parcial** es un operador

$$\delta_s : \prod_{i \in \{s_1, \ldots, s_p\}} T_i \longrightarrow Q_s$$

siendo $\prod$ el producto cartesiano donde para cada $i \leq s_p$, T_i es la imagen del criterio de comparación C_i; donde tanto Q_s como T_i pueden ser dominios booleanos, k -valentes o reales. ■

[4] Se pueden considerar, aunque no es frecuente, criterios de comparación de valores de una variable cuya imagen sea un subconjunto de algún conjunto totalmente ordenado no numérico.

El papel de la función de semejanza parcial es el valuar la semejanza entre las subdescripciones de los objetos definidas por las ω-partes.

Definición 2.5.[19] Sea $\Omega_A = \{\Omega_1, \ldots, \Omega_q\}$. Una **función de semejanza (total)** es un operador

$$\Gamma : \prod_{s=1}^{q} Q_s \longrightarrow L$$

sobre los valores asignados para cada proyección Ω_s por parte de la función de semejanza parcial. ■

Aunque de hecho la función de semejanza total se define sobre el producto cartesiano de las imágenes de las funciones de semejanza parciales, en lo sucesivo por comodidad sólo haremos referencia al espacio de representación de los objetos Ξ, es decir, dados $\{C_1, \ldots C_m\}$ una familia de criterios de comparación, $\Omega_A = \{\Omega_1, \ldots, \Omega_q\}$ un sistema de conjuntos de apoyo; $\delta_{\Omega_A} = \{\delta_1, \ldots, \delta_q\}$ un conjunto de funciones de semejanza parcial y $O_j=\{x_j^1, \ldots, x_j^m\}$, $O_r=\{x_r^1, \ldots, x_r^m\} \in \Xi$, denotaremos como $\Gamma(O_j,O_r) \in L$ al valor

$$\Gamma(\prod_{s=1}^{q}(\ \delta_s(\prod_{1\in(s_1,\ldots,s_p)} C_1(x_j^1, x_r^1))))$$

de donde tenemos que

$$\forall\ O_r, O_j \in \Xi\ \Gamma(O_r,O_j)=\Gamma(O_j,O_r)$$

A partir de Ξ, δ_{Ω_A} y Ω_A se puede construir una matriz simétrica $\{\Gamma_{rj}\}_{nxn}$, donde, por ejemplo,[19]

$$\Gamma_{rj} = \Gamma(O_r,O_j) = \frac{\gamma(O_{rj})}{|\ \Omega_A\ |}\sum_{s=1}^{q} P(\Omega_s)\ \omega_s\delta_{rj}$$

siendo $\omega_s\delta_{rj} = \delta_s(\omega_s(O_r),\omega_s(O_j))$; $|\Omega_A|$ la cantidad de conjuntos de apoyo; $\gamma(O_{rj})$ un parámetro de ponderación resultante de la comparación de los objetos O_r, O_j y que puede poner de relieve sendas ponderaciones de estos objetos; análogamente $P(\Omega_s)$ es un parámetro de ponderación asociado a cada conjunto de apoyo Ω_s y que pudiera estar en función de la importancia informacional $P(r)$ de cada uno de los rasgos, por ejemplo [19]

$$P(\Omega_s) = \frac{\sum_{r \in \Omega_s} P(r)}{|\Omega_s|}$$

Definición 2.6.[19] Dados $O_i, O_j \in \Xi$ y $\beta_0 \in L$, diremos que O_i es β_0 -semejante a O_j si y sólo si $\Gamma(O_i,O_j) \geq \beta_0$.
Dado $O_i \in \Xi$, diremos que O_i es β_0-aislado si y sólo si
$\forall\ O_j \neq O_i\ O_j \in \Xi\ [\ \Gamma(O_i,O_j) < \beta_0]$ ∎

La magnitud $\beta_0 \in L$, que es denominada umbral de semejanza, puede ser definida, por ejemplo, en el caso en que $L = [0,1]$, como sigue

a) $\beta_0 = \dfrac{2}{n\ (n-1)} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \Gamma(O_i,O_j)$

b) $\beta_0 = \dfrac{1}{n} \sum_{i=1}^{n} \max_{\substack{j=1,\ldots,n \\ i \neq j}} \{\Gamma(O_i,O_j)\}$

c) $\beta_0 = \min_{i=1,\ldots,n-1} \{\min_{j=i+1,\ldots,n} \{\Gamma(O_i,O_j)\}\}$

Criterios de agrupación gráficos de análisis de cúmulos
Sea $\Gamma: \Delta \times \Delta \rightarrow L$ una función de semejanza y sea $\Xi = \{O_1, \ldots, O_n\}$ el conjunto de descripciones de Δ.
Este tipo de criterios manejan en general como parámetros una matriz simétrica $\{\Gamma_{ij}\}$ denominada matriz de semejanza para la cual cada $\Gamma_{ij} = \Gamma(O_i,O_j) \in L$; y un umbral $\beta_0 \in L$ a partir del cual se define el concepto de β_0-semejanza. La definición de los

agrupamientos se da a partir del cumplimiento de ciertas propiedades predefinidas (criterio agrupacional) entre las semejanzas de las descripciones que generan particiones duras en el conjunto total de éstas sin que sea necesario definir de antemano el número de agrupaciones. Existe una identificación natural entre el espacio de representación y una gráfica cuyos vértices son las descripciones y las aristas tendrán como peso el valor de la semejanza entre sus vértices adyacentes.
Veamos algunos ejemplos:

Criterio β_0-conexo
Sea $\beta_0 \in L$. Un subconjunto $C \neq \emptyset$ de Ξ es una *componente* β_0-*conexa* si y sólo si

a) $\forall\ O_i, O_j \in C\ \exists\ O_{i_1}, \ldots, O_{i_s} \in C$

$$[O_i = O_{i_1} \wedge O_j = O_{i_s} \wedge \forall\ p<s\ \Gamma(O_{i_p}, O_{i_{p+1}}) \geq \beta_0].$$

b) $\forall\ O_i \in \Xi\ [(O_j \in C \wedge \Gamma(O_i, O_j) \geq \beta_0) \Rightarrow O_i \in C]$.

Por definición, todo punto aislado constituye una componente β_0-conexa (*degenerada*). ∎

Criterio β_0-compacto
Diremos que un subconjunto $B \neq \emptyset$ de Ξ es β_0-*compacto* si y sólo si

a) $\forall\ O_j \in \Xi\ [O_i \in B \wedge \max\limits_{O_t \in \Xi} \{\Gamma(O_i, O_t)\} = \Gamma(O_i, O_j) \geq \beta_0] \Rightarrow O_j \in B$

b) $\forall\ O_i, O_j \in B\ \exists\ O_{i_1}, \ldots, O_{i_s} \in B\ [O_i = O_{i_1} \wedge O_j = O_{i_s} \wedge$
$\forall\ p<s\ [\max\limits_{O_t \in \Xi}\{\Gamma(O_{i_p}, O_t)\} = \Gamma(O_{i_p}, O_{i_{p+1}}) \geq \beta_0\ \vee$
$\max\limits_{O_t \in \Xi}\{\Gamma(O_{i_{p+1}}, O_t)\} = \Gamma(O_{i_p}, O_{i_{p+1}}) \geq \beta_0]]$.

Por definición todo punto aislado constituye un conjunto β_0 -compacto (*degenerado*). ∎

Dado el hecho de que estos criterios no manejan los téminos en que están definidas las descripciones de los objetos, entonces permiten cualquier tipo de descripción de los mismos (cuantitativa, cualitativa, mixta, etcétera), y para la imagen de la función de semejanza es posible exigir tan sólo que sea un orden lineal.

Criterios de cúmulos traslapados

Este tipo de criterios no plantea la condición de que el resultado del análisis de cúmulos sea una partición en conjuntos duros y presenta como solución al problema un cubrimiento de conjuntos duros del conjunto de descripciones. Es decir, existen descripciones que pueden pertenecer a varios agrupamientos. Al igual que los criterios gráficos, dado $\Xi = \{O_1, \ldots, O_n\}$ y $\Gamma:\Delta\times\Delta \rightarrow L$ la función de semejanza, estos criterios utilizan como parámetros la matriz de semejanza $\{\Gamma_{ij}\}$ y un umbral $\beta_0 \in L$ que define cuándo un objeto es semejante a otro. Además utilizan una función entre subconjuntos de Ξ, $D:P(\Xi)\times P(\Xi) \rightarrow Q$ para algún Q (donde $P(\Xi)$ es la potencia de Ξ), la cual es una medida del encimamiento entre agrupaciones, y un umbral $\alpha \in Q$. En general, el método a seguir para obtener las agrupaciones es el siguiente:
Se define $\{K_1, \ldots, K_c\}$ la familia de agrupaciones que cumple
$\forall\ r \leq c\ \forall\ O_i,\ O_j \in K_r\ \ [\Gamma_{ij} \geq \beta_0]$

Después, para cada par de agrupaciones K_i, K_r se medirá por medio de D el encimamiento entre ambas.
Si $D(K_i, K_r) \geq \alpha$ entonces K_i y K_r se fusionarán en un sólo agrupamiento. Como ejemplos de D tenemos $|K_i \cap K_r|$[15] y $|K_i \cap K_r| / |K_i \cup K_r|$[22].

Este tipo de criterios nos resulta interesante por lo siguiente: sugieren interpretar a las agrupaciones como representantes de conceptos o propiedades diferentes, y el hecho de que existan agrupamientos encimados no es más que una consecuencia de la similitud entre los conceptos que cada agrupación representa. Esto nos permite ver a estas agrupaciones duras como representaciones de "centros conceptuales" de cúmulos.

CAPITULO III

UN NUEVO MODELO DE ALGORITMOS PARA EL ANALISIS DE CUMULOS DIFUSOS

Presentamos aquí un nuevo modelo de algoritmos para el análisis de cúmulos difusos el cual retoma ciertas propiedades de los criterios gráficos y traslapados duros. Se aprovecha la flexibilidad por parte de los criterios gráficos en el manejo del tipo de descripciones de objetos, de la función de semejanza, en la definición de diferentes tipos de criterios agrupacionales y el que no requiera definir de antemano el número de agrupaciones; a la vez que interpretaremos a las agrupaciones como "centros conceptuales" de agrupamientos difusos.

Criterio Agrupacional Difuso

Sea un espacio de representación $\Phi = \{\Delta, \Gamma\}$ cuyo conjunto de descripciones de objetos denotamos como $\Xi = \{O_1, \ldots, O_n\}$, una función de semejanza $\Gamma: \Delta\times\Delta \to L$, que define la estructura de Φ, y un umbral $\beta_0 \in L$ que define la β_0-semejanza entre los elementos de Ξ. A la función de semejanza Γ la consideraremos la función de pertenencia $\upsilon_\Gamma: \Xi\times\Xi \to L$ de la relación difusa $\tilde{\Gamma} = \{((O_i, O_j), \Gamma(O_i, O_j)) \,/\, (O_i, O_j) \in \Xi\times\Xi\}$ que representa la semejanza entre las descripciones de los objetos de Ξ. De aquí en adelante,por comodidad, no haremos distinción entre *objetos* y *descripciones de objetos*.

Definición 3.1.[18] **Como criterio agrupacional** $\Pi(\Xi, \Gamma, \beta_0)$ **sobre** Φ entenderemos un conjunto de proposiciones con parámetros Ξ, Γ y β_0 que:

a) Genera una familia $\tau = \{NU_1, \ldots, NU_c\}$ de subconjuntos de Ξ (agrupaciones duras) que cumple

i) $\forall\, NU \in \tau\ [NU \neq \emptyset]$

ii) $\bigcup_{NU\in\tau} NU = \Xi$.

iii) $\neg\exists\ NU_r, NU_{j_1}, \ldots, NU_{j_k} \in \tau\ [NU_r \subseteq \bigcup_{t=1}^{k} NU_{j_t}\ /\ j_t \neq r\ \forall\ t=1,..,k]$

b) Define una relación $R_\Pi \subseteq \Xi \times \Xi \times P(\Xi)$ (donde $P(\Xi)$ es la potencia de Ξ) que cumple

iv) $\forall\ O_i, O_j \in \Xi\ [\exists\ NU \in \tau\ /\ O_i, O_j \in NU \leftrightarrow \exists\ S \subseteq \Xi\ /\ R_\Pi(O_i, O_j, S)]$ ■

Definición 3.2.[18] A cada NU_r elemento de la familia τ lo denominaremos núcleo. ■

La relación R_Π

La relación R_Π es en sí la representación matemática de las proposiciones de Π, y por ello R_Π genera a la familia τ en el mismo sentido que una relación de equivalencia genera una partición, es decir, no define cuándo un objeto O_i pertenece a un núcleo NU_r, sino que nos dice cuándo dos objetos O_i y O_j están en un mismo núcleo. Los argumentos de R_Π son tales ya que la razón por la que dos objetos pertenecen o no a un mismo núcleo puede ser dependiente del comportamiento global de las semejanzas entre otros objetos además de ellos mismos. Ese conjunto de otros objetos es S, y por esto, R_Π representa que de acuerdo a: el comportamiento entre las semejanzas de los objetos de $S \subseteq \Xi$ entre sí, la semejanza entre O_i y O_j, y las semejanzas de éstos con los elementos de S, se dará el que O_i y O_j pertenezcan a un mismo núcleo de la familia τ o no.

Ejemplo 3.1.

Sea $\Phi = \{\Delta, \Gamma\}$ un espacio de representación, $\Xi = \{O_1, \ldots, O_n\}$ el conjunto de descripciones de objetos, $\Gamma: \Delta \times \Delta \to L$ una función de semejanza y $\beta_0 \in L$ un umbral.

Sea $\Pi(\Xi,\ \Gamma,\ \beta_0)$ el siguiente conjunto de proposiciones que generan a τ:

$NU \in \tau$ si y sólo si

i) $\forall\ O_i, O_j \in NU\ \exists\ O_{i_1}, \ldots, O_{i_q} \in NU$

$[\Gamma(O_i, O_{i_1}) \geq \beta_0 \wedge \Gamma(O_{i_q}, O_j) \geq \beta_0 \wedge \forall\ p<q\ \Gamma(O_{i_p}, O_{i_{p+1}}) \geq \beta_0]$.

ii) $\forall\ O_i \in \Xi\ [(O_j \in NU \wedge \Gamma(O_i,O_j) \geq \beta_0) \Rightarrow O_i \in NU]$.
iii) $\forall\ O_i \in \Xi\ [NU=\{O_i\} \Leftrightarrow \forall\ O_j \neq O_i\ O_j \in \Xi\ [\Gamma(O_i,O_j)<\beta_0]]$

Notemos que en el caso de que $\Gamma(O_i,O_j) \geq \beta_0$, el hecho de que O_i y O_j pertenezcan a un mismo núcleo no depende de ningún otro objeto. Y si $\Gamma(O_i,O_j) < \beta_0$ entonces el que ambos pertenezcan a un mismo núcleo dependerá de que haya un subconjunto de objetos que los "encadenen" con β_0-semejanzas. Así pues:
Para cada $S \subseteq \Xi$, denotamos como $Su(S)$ al conjunto de todos los ordenamientos sin repetición de todos los elementos de S, es decir, siendo $|S| \in \mathbb{N}$ el cardinal de S

$$Su(S)=\{F \subseteq |S| \times S\ /\ \forall x \in |S|\ \exists!\ O \in S\ \langle x,O\rangle \in F,\ \forall\ O \in S\ \exists!\ x \in |S|\ \langle x,O\rangle \in F\}$$

Entonces tenemos que
$\forall\ O_i,O_j \in \Xi\ \forall\ S \subseteq \Xi,\ R_\Pi(O_i,O_j,S)$ si y sólo si
$[\Gamma(O_i,O_j) \geq \beta_0 \wedge S=\emptyset]\ \vee$
$[\exists\ \{O_{i_1},\ldots,O_{i_q}\} \in Su(S) \wedge \Gamma(O_i,O_{i_1}) \geq \beta_0 \wedge \Gamma(O_{i_q},O_j) \geq \beta_0$
$\wedge\ \forall\ p<q\ \Gamma(O_{i_p},O_{i_{p+1}}) \geq \beta_0]$. ∎

El hecho de que la relación R_Π y las familias de núcleos τ fueran definidas con base en las β_0-semejanzas entre las descripciones de los objetos y cierto comportamiento entre estas, implica intuitivamente que cada núcleo NU_r (cada agrupación dura) también esté definido por una propiedad P_r referente a las descripciones (que no es "la descripción de O pertenece a NU_r"), propiedad que tienen en común cumplir los objetos pertenecientes al mismo núcleo NU_r. Así pues, consideraremos que para cada $r \leq c$

$$NU_r = \{\ O \in \Xi\ /\ P_r(O)\ \}$$

Es de hacerse notar que en función de Π, las propiedades P_r que definen a los núcleos NU_r serán o no relevantes dado un problema concreto en el campo de las aplicaciones. Lo único que se va a asegurar es que las propiedades que definen a los núcleos estarán íntimamente relacionadas con Π, siendo por tanto el conjunto de proposiciones de Π el factor que desempeña el papel más relevante en la búsqueda de un análisis de cúmulos con sentido en el marco de un problema dado.

Los núcleos y los α-cortes

En esta sección plantearemos que tiene sentido el medir de forma difusa la pertenencia de los objetos a los núcleos.

Sean $\beta_0, \alpha \in L$, $\alpha > \beta_0$. Sea $\Pi(\Xi, \Gamma, \beta_0)$ un criterio agrupacional que genera una familia de núcleos $\tau = \{NU_1, \ldots, NU_c\}$ y una relación R_Π, y sea $\Pi'(\Xi, \Gamma, \alpha)$ el criterio agrupacional cuyas proposiciones son las proposiciones de Π a las cuales se les sustituyeron todas las ocurrencias de β_0 por α, y que genera una familia de núcleos $\tau' = \{NU'_1, \ldots, NU'_{c'}\}$ y una relación R'_Π. Sea $P_1, \ldots, P_c$ el conjunto de propiedades que respectivamente definen a cada $NU_r \in \tau$ y sea $P'_1, \ldots, P'_{c'}$ el conjunto de propiedades que respectivamente definen a cada $NU'_r \in \tau'$. Notemos que el conjunto de proposiciones a partir del cual se generaron τ y τ' es el mismo salvo que Π' es más restrictivo en cuanto a la semejanza entre los objetos que Π, y entonces ocurre que $R'_\Pi \leq R_\Pi$. Si α y β_0 son tales que $R'_\Pi \lneq R_\Pi$, entonces tenemos lo siguiente:

$$\forall\ NU_r \in \tau\ \exists\ NU'_{r_1}, \ldots, NU'_{r_{s_r}} \in \tau' \ /\ NU_r = \bigcup_{t=1}^{s_r} NU'_{r_t}$$

y existen $NU_j \in \tau$, $NU'_{j_1}, \ldots, NU'_{j_{s_j}} \in \tau'$ tal que

$$NU_j = \bigcup_{t=1}^{s_j} NU'_{j_t} \text{ con } s_j > 1 \tag{1}$$

y $\forall\ r \leq s_j$

$$NU_j \nsubseteq \bigcup_{\substack{t=1 \\ t\neq r}}^{s_j} NU'_{j_t} \tag{2}$$

Por *iii*) de la definición de Π, (1) y (2), tenemos que para cada NU'_{j_q} y NU'_{j_k} existen $O_i, O_t \in NU_j$ tales que

$$O_i \in NU'_{j_k},\ O_t \notin NU'_{j_k},\ O_t \in NU'_{j_q},\ O_i \notin NU'_{j_q}$$

y por tanto

$\forall\ q,k \le s_j,\ \exists\ O_i, O_t\ /$

$$P_j(O_i),\ P_j(O_t),\ P'_{j_q}(O_t),\ \neg P'_{j_q}(O_i),\ P'_{j_k}(O_i),\ \neg P'_{j_k}(O_t)$$

Como el conjunto de proposiciones de Π y Π' que generan a τ y τ' son las mismas salvo por el parámetro de la semejanza, las propiedades P'_{j_q} que definen a los núcleos NU'_{j_q} deben estar estrechamente relacionadas con la propiedad P_j que define al núcleo NU_j y su diferencia sólo debe recaer en las *semejanzas entre los objetos*.

Una interpretación de este comportamiento puede ser la siguiente. Para cada núcleo $NU_r \in \tau$

$$P_r(O_i) \cong \bigwedge_{k=1}^{s_r} P_{r_k}(O_i)$$

$$P_{r_k}(O_i) \cong "\upsilon_{T_{r_k}}(O_i) \ge \varphi_{\beta_0}"$$

donde φ_{β_0} es un umbral en función de β_0, $\upsilon_{T_{r_k}}$ es la función de pertenencia de un subconjunto difuso, donde cada $\upsilon_{T_{r_k}}(O_i)$ es la medida de verdad de "O_i cumple la propiedad T_{r_k}" y

$$P'_{r_k}(O_i) \cong "\upsilon_{T_{r_k}}(O_i) \ge \varphi_{\alpha}"$$

con φ_α un umbral en función de α tal que $\varphi_\alpha > \varphi_{\beta_0}$.

Es decir, cada núcleo NU_r está definido por una propiedad P_r que es una conjunción de propiedades $\bigwedge_{k=1}^{s_r} P_{r_k}$ y cada uno de los elementos de NU_r cumple cada $P_{r_k}(O_i) \cong "\upsilon_{T_{r_k}}(O_i) \ge \varphi_{\beta_0}"$, o lo que es lo mismo

$$P_r(O_i) \cong \bigwedge_{k=1}^{s_r} "\upsilon_{T_{r_k}}(O_i) \ge \varphi_{\beta_0}" \cong "\upsilon_{T_r}(O_i) \ge \varphi_{\beta_0}"$$

donde, como propiedad, $T_r \cong \bigwedge_{k=1}^{s_r} T_{r_k}$ y así

$$P_r(O_i) \cong "\upsilon_{T_r}(O_i) \ge \varphi_{\beta_0}" \cong "\min_{k \le s_r}\{\upsilon_{T_{r_k}}(O_i)\} \ge \varphi_{\beta_0}"$$

En el caso de NU_j, al elevar el umbral de semejanza en R'_π no se genera el núcleo NU_j ya que no todos los elementos que pertenecen a NU_j cumplen todas las propiedades T_{j_k} con grado φ_α, no se asemejan tanto entre sí como para cumplir todos las mismas propiedades en ese grado, sino que se crean otros núcleos, a saber, los NU'_{j_k} definidos cada uno de ellos por los $P'_{j_k}(O_i) \cong "\upsilon_{T_{j_k}}(O_i)\geq\varphi_\alpha"$.

Nótese que consideramos a φ_{β_0} como fijo para todos los núcleos pertenecientes a una misma familia τ. Esto en cierto modo es natural ya que los núcleos han sido definidos a partir de las mismas proposiciones de Π y por el mismo umbral β_0.

De todo esto podemos concluir que:

Conclusión 3.1. Sea $\Pi(\Xi,\ \Gamma,\ \beta_0)$ un criterio agrupacional y sea R_π la relación definida por Π. Sea $\tau = \{NU_1,\ \ldots, NU_c\}$ la familia de núcleos generada por R_π. Entonces, para cada núcleo $NU_r \in \tau$

$$P_r(O_i) \cong \bigwedge_{k=1}^{s_r} P_{r_k}(O_i) \cong \bigwedge_{k=1}^{s_r} "\upsilon_{T_{r_k}}(O_i)\geq\varphi_{\beta_0}" \cong "\upsilon_{T_r}(O_i)\geq\varphi_{\beta_0}" \quad \blacksquare$$

Conclusión 3.2. Si existe $\alpha \in L$ tal que el criterio agrupacional $\Pi'(\Xi,\ \Gamma,\ \alpha)$, cuyas proposiciones son las proposiciones de Π a las cuales se les sustituyeron todas las ocurrencias de β_0 por α, define una familia $\tau' = \{NU'_1,\ \ldots, NU'_{c'}\}$ de núcleos en la cual existen, para algún $NU_j \in \tau$

$$NU'_{j_1}, \ldots,\ NU'_{j_{s_j}} \in \tau' \ /\ NU_j = \bigcup_{t=1}^{s_j} NU'_{j_t} \text{ con } s_j > 1$$

y $\forall\ r \leq s_j$

$$NU_j \not\subseteq \bigcup_{\substack{t=1 \\ t\neq r}}^{s_j} NU'_{j_t}$$

entonces NU_j es un φ_{β_0}-corte no máximo de un conjunto difuso. Es decir

$$\exists\ O_i \in NU_j \ /\ [1 \geq \varphi_\alpha > \upsilon_{T_j}(O_i) \geq \varphi_{\beta_0}] \quad \blacksquare$$

Esto último es significativo ya que en general, para cualquier núcleo NU tal que $|NU|>1$, existe un umbral para el cual este núcleo se "rompe".

Lo visto hasta ahora plantea uno de los aspectos básicos que conforman el modelo. A partir de un criterio agrupacional Π, es generado un conjunto de agrupaciones (núcleos NU_r) caracterizadas éstas por distintas propiedades P_r las cuales dicen "cumplir la propiedad T_r en cierto grado". Es decir, "alrededor" de las distintas propiedades T_r se "aglomeran" los objetos del espacio de representación y los núcleos NU_r se conforman de los objetos que están suficientemente "cerca" (según el umbral φ_{β_0}) de este "centro conceptual" de las "aglomeraciones".

A continuación buscaremos la manera de medir esa "cercanía" por parte de todos los objetos del espacio de representación a esos "centros".

Criterio Agrupacional Difuso

Como ya se ha mencionado, en la definición de un criterio agrupacional interviene el concepto de la β_0-semejanza, el cual es un límite a partir del cual consideramos cuándo un objeto es semejante a otro o no. Sin embargo, en estos términos se pierde información en cuanto a la semejanza misma entre los objetos (de hecho, el valor de la semejanza entre ellos). En general, esto que sucede con respecto a la β_0-semejanza es aplicable a cualquier propiedad P que intervenga en Π cuya bivaluación a partir de un umbral $\alpha_P \in L$ implique la pérdida de información relevante en cuanto a P que debiera reflejarse en la relación R_π. A este tipo de propiedades las denominaremos propiedades no duras según Π. Esta pérdida general de información repercute de la misma manera en R_π, ya que R_π sólo nos dice si existe la relación o no, cuando intuitivamente debería suceder, por ejemplo, que si dos objetos son más semejantes que otros, los primeros deberían estar más relacionados que los segundos. Si en el marco de un problema dado podemos definir para cada propiedad P no dura según

Π una medida que:
1.- Nos diga si se cumple P por arriba del umbral α_p o no; y
2.- Nos dé una medida del cumplimiento de P;

entonces, a partir de las proposiciones de Π, considerando en lugar del cumplimiento o no de P a la medida que hayamos definido del cumplimiento de ésta propiedad (una seudodifusión del conjunto $\{x \in \Xi_p / \upsilon(P(x)) \geq \alpha_p\}$ con Ξ_p el dominio de P, por ejemplo), podemos definir una relación difusa $\tilde{R}_\pi$ que:
1.- Nos diga si R_π se cumple o no; y
2.- Nos dé una medida de que tanto se cumple la relación en función de las propiedades no duras según Π.

Ejemplo 3.2.
Sea dada la relación R_π tal que
$\forall\ O_i, O_j \in \Xi\ \forall\ S \subseteq \Xi,\ R_\pi(O_i, O_j, S)$ si y sólo si
$[\Gamma(O_i, O_j) \geq \beta_0 \wedge S = \emptyset] \vee$
$[\exists\ \{O_{i_1}, \ldots, O_{i_q}\} \in Su(S) \wedge \Gamma(O_i, O_{i_1}) \geq \beta_0 \wedge \Gamma(O_{i_q}, O_j) \geq \beta_0$
$\wedge\ \forall\ p<q\ \Gamma(O_{i_p}, O_{i_{p+1}}) \geq \beta_0]$.

donde $Su(S)$ es el conjunto de todos los ordenamientos sin repetición de todos los elementos de S.
Sea $\tilde{F} \in D(\geq\beta_0)$[5] a partir del cual se define la seudodifusión de la β_0-semejanza. Entonces definimos la función de pertenencia de la relación $\tilde{R}_\pi$ como
$\forall\ O_i, O_j \in \Xi\ \forall\ S \subseteq \Xi,$

$$\upsilon_{\tilde{R}_\pi}(O_i, O_j, S) = \begin{cases} \upsilon_{\tilde{F}}(\Gamma(O_i, O_j) \geq \beta_0) & \text{si } S = \emptyset \\ \max\limits_{(O_{i_1}, \ldots, O_{i_q}) \in Su(S)} \{\min\{\upsilon_{\tilde{F}}(\Gamma(O_i, O_{i_1}) \geq \beta_0), & \\ \upsilon_{\tilde{F}}(\Gamma(O_{i_q}, O_j) \geq \beta_0), \min\limits_{p \leq q-1}\{\upsilon_{\tilde{F}}(\Gamma(O_{i_p}, O_{i_{p+1}}) \geq \beta_0)\}\}\} & \text{si } S \neq \emptyset \end{cases}$$

y así

[5] $D(\geq\alpha)$ es la clase de subconjuntos Mayor o Igual que α Difuso.

$\forall\ O_i, O_j \in \Xi\ \forall\ S \subseteq \Xi,$

$$R_\Pi(O_i,O_j,S) \iff \upsilon_{\tilde{R}_\Pi}(O_i,O_j,S) \geq \upsilon_{\tilde{F}}(\beta_0 \geq \beta_0) \qquad \blacksquare$$

En general, la definición de la relación difusa $\tilde{R}_\Pi$ asociada a una relación R_Π se realiza de la siguiente manera:

Sea R_Π una relación definida a partir de un criterio agrupacional $\Pi(\Xi,\ \Gamma,\ \beta_0)$ la cual está caracterizada como:

$\forall\ O_i, O_j \in \Xi\ \forall\ S \subseteq \Xi,\ R_\Pi(O_i,O_j,S)$ si y sólo si

$P_1(O_i,O_j,S) \circledast \ldots \circledast P_r(O_i,O_j,S)$

donde $P_1(O_i,O_j,S) \circledast \ldots \circledast P_r(O_i,O_j,S)$ denota las conjunciones y/o disyunciones ($\circledast$ representa el "$\wedge$" ó el "$\vee$" lógicos) de las propiedades $P_1, \ldots, P_r$ referentes a los elementos y subconjuntos de Ξ, la función Γ y el umbral β_0, a partir de las cuales se caracteriza la relación R_Π.

Sea $ND = \{P_{i_1}, \ldots, P_{i_q}\} \subseteq \{P_1, \ldots, P_r\}$ el conjunto de propiedades tales que

$$\forall\ k \leq q\ \ P_{i_k}(O_i,O_j,S) \cong \text{"}\upsilon(T_k(O_i,O_j,S)) \geq \alpha_k\text{"}$$

donde cada T_k es una propiedad no dura según Π y $\alpha_k \in L$ un umbral.

Sean $\{\tilde{F}_k\}$ una familia de q seudodifusiones tal que

a) $\forall\ k \leq q\ \ \tilde{F}_k \in D(\geq\alpha_k)$ y

b) $\forall\ m,k \leq q\ \ \upsilon_{\tilde{F}_k}(\alpha_k \geq \alpha_k) = \upsilon_{\tilde{F}_m}(\alpha_m \geq \alpha_m)$

Definimos la relación difusa $\tilde{R}_\Pi$ como

$\forall\ O_i, O_j \in \Xi\ \forall\ S \subseteq \Xi,$

$$\upsilon_{\tilde{R}_\Pi}(O_i,O_j,S) = \upsilon(Q_1(O_i,O_j,S) \circledast \ldots \circledast Q_r(O_i,O_j,S))$$

donde para cada $m \leq r$

a) $Q_m(O_i,O_j,S) \cong \text{"}\upsilon(T_k(O_i,O_j,S)) \in \tilde{F}_k\text{"}$ si $P_r \cong P_{i_k} \in ND$ siendo entonces $\upsilon(Q_m(O_i,O_j,S)) = \upsilon_{\tilde{F}_k}(\upsilon(T_k(O_i,O_j,S)) \geq \alpha_k)$;

b) $Q_m(O_i,O_j,S) \cong P_m(O_i,O_j,S)$ si $P_m \notin ND$ siendo $\upsilon(Q_m(O_i,O_j,S))$ el valor de verdad (0 ó 1) de la proposición dura según Π $P_m(O_i,O_j,S)$.

A su vez vemos que
$\forall\ O_i,O_j \in \Xi\ \forall\ S\subseteq\Xi,$
$R_\pi(O_i,O_j,S) \Leftrightarrow \upsilon_{\tilde{R}}(O_i,O_j,S) \geq \upsilon(Q_1'(O_i,O_j,S)) \circledast \ldots \circledast Q_r'(O_i,O_j,S)))$

donde
a) $Q_m'(O_i,O_j,S) \cong$ "$\alpha_k \in \bar{F}_k$" si $P_m \cong P_{l_k} \in ND$ siendo entonces $\upsilon(Q_m'(O_i,O_j,S)) = \upsilon_{\bar{F}_k}(\alpha_k \geq \alpha_k)$;
b) $Q_m'(O_i,O_j,S) \cong P_m(O_i,O_j,S)$ si $P_m \notin ND$ siendo $\upsilon(Q_m'(O_i,O_j,S))$ el valor de verdad (0 ó 1) de la proposición dura según Π $P_m(O_i,O_j,S)$.

Con todo esto, tenemos la siguiente

Definición 3.3.[18] Como criterio agrupacional difuso $\Pi_d(\Xi,\ \Gamma,\ \beta_0)$ sobre Φ entenderemos

a) Un criterio agrupacional $\Pi(\Xi,\ \Gamma,\ \beta_0)$ que define a la familia τ y la relación $R_\pi \subseteq \Xi\times\Xi\times P(\Xi)$; y

b) Un conjunto de seudodifusiones $\bar{F}_1,\ldots,\bar{F}_q$ que respectivamente describen a las propiedades no duras según Π y que generan una relación difusa $\tilde{R}_\pi$ sobre $\Xi\times\Xi\times P(\Xi)$ que cumple

$$R_\pi(O_i,O_j,S) \Leftrightarrow \upsilon_{\tilde{R}_\pi}(O_i,O_j,S) \geq f(\beta_0)$$

con $f(\beta_0)$ un umbral en función de β_0, Π y de $\bar{F}_1,\ \ldots,\ \bar{F}_q$. ■

Obsérvese que en la definición seguimos considerando tanto a τ como a R_π. Esto no es más que para recalcar la función de cada uno de ellos. R_π define las propiedades $P_1,\ \ldots,\ P_c$ y $T_1,\ \ldots,\ T_c$ a partir de un umbral, es decir, define los "centros conceptuales"; y la relación difusa $\tilde{R}_\pi$ mide qué tan relacionados están los objetos entre sí en función de las propiedades no duras según Π. Como veremos a continuación, $\bar{R}_{\Pi}$ nos ayudará a definir una medida de cuánto cumple cada objeto las propiedades definidas por R_{Π} o que tan "cerca" está de esos "centros".

Medida de pertenencia de los objetos a un núcleo

Sea Π_d un criterio agrupacional difuso sobre Φ que define una familia de núcleos $\tau = \{NU_1, \ldots, NU_c\}$, una relación $R_{\Pi} \subseteq \Xi \times \Xi \times P(\Xi)$ y una relación difusa $\bar{R}_{\Pi}$ a partir de la familia de seudodifusiones $\bar{F}_1, \ldots, \bar{F}_q$.

Definamos para cada $r \leq c$ y $O_i \in \Xi$, el conjunto

$$C^r_{O_i} = \{\langle O_j, S\rangle \in (\Xi \backslash O_i) \times P(\Xi) \;/\; P_r(O_j) \wedge (R_{\Pi}(O_i, O_j, S) \Leftrightarrow P_r(O_i))\}$$

que es el conjunto de parejas $\langle O_j, S\rangle$ donde $O_j \neq O_i$, O_j pertenece a NU_r y S es un subconjunto de Ξ que junto con O_j es "responsable" de que O_i esté o no en el mismo núcleo NU_r en el que está O_j.

A partir de la definición de $C^r_{O_i}$ tenemos lo siguiente:

a) $\forall\, r \leq c\ \forall\, O_i \in \Xi\ \{C^r_{O_i} \neq \emptyset \Rightarrow [P_r(O_i) \Leftrightarrow \exists\, \langle O_j, S\rangle \in C^r_{O_i}\ [R_{\Pi}(O_i, O_j, S)]]\}$

lo cual es equivalente en términos de $\bar{R}_{\Pi}$ y T_r a que

$$\forall\, r \leq c\ \forall\, O_i \in \Xi\ \{C^r_{O_i} \neq \emptyset \Rightarrow$$
$$[\upsilon_{T_r}(O_i) \geq \varphi_{\beta_0}] \Leftrightarrow \exists\, \langle O_j, S\rangle \in C^r_{O_i}\ [\upsilon_{\bar{R}_{\Pi}}(O_i, O_j, S)) \geq f(\beta_0)]\}$$

Ahora consideremos lo siguiente. En la teoría de conjuntos si tenemos que una proposición $\forall x \in A(a(x) \Rightarrow [b(x) \Leftrightarrow c(x)])$ es verdadera, entonces podemos concluir metaconjuntistamente que cuando se cumple $a(x)$, entonces el valor de verdad de $b(x)$ es igual al de $c(x)$. Siguiendo este curso de ideas en términos de subconjuntos difusos es congruente concluir "metasubconjuntistamente difuso" que si el valor de verdad de $\forall x \in A(a(x) \Rightarrow [b(x) \Leftrightarrow c(x)])$ y el de $a(x)$ es mayor que el de sus respectivas negaciones, entonces el que "el valor de verdad de $b(x)$ es igual al de $c(x)$" será más verdadero que el que "el valor de verdad de b(x) no es igual al de c(x)", por lo que proponemos que

$\forall\ r\leq c\ \forall\ O_i \in \Xi$, si $C^r_{O_i}\neq\emptyset$ entonces

$$\upsilon(\upsilon_{T_r}(O_i)\geq\varphi_{\beta_0}) = \max_{\langle O_j,S\rangle\in C^r_{O_i}} \{\upsilon(\upsilon_{\bar{R}_\pi}(O_i,O_j,S)\geq f(\beta_0))\}$$

de donde se puede interpretar que en la misma medida en que $\upsilon_{T_r}(O_i)$ sea mayor o igual que φ_{β_0}, lo será $\max\{\upsilon_{R_\pi}(O_i,O_j,S)\}$ con respecto a $f(\beta_0)$, lo que en otros términos implica, recordando que φ_{β_0} y $f(\beta_0)$ son fijos, que

$\forall\ O_i,O_j \in \Xi\ \forall\ r,k\leq c$, si $C^r_{O_i}\neq\emptyset$ y $C^k_{O_j}\neq\emptyset$ entonces

$$\upsilon_{T_r}(O_i)>\upsilon_{T_k}(O_j) \Rightarrow \max_{\langle O_t,S\rangle\in C^r_{O_i}} \{\upsilon_{\bar{R}_\pi}(O_i,O_t,S)\} \geq \max_{\langle O_l,N\rangle\in C^k_{O_j}} \{\upsilon_{\bar{R}_\pi}(O_j,O_l,N)\} \quad (3)$$

siendo todo esto bastante "natural" si notamos que $\upsilon_{\bar{R}_\pi}(O_i,O_t,S)$ con $\langle O_t,S\rangle \in C^r_{O_i}$ es la representación de "según R_π, O_t y S, que tanto pertenece O_i a NU_r", es decir, "O_i que tanto cumple T_r", lo cual además nos dice que

$\forall\ O_i \in \Xi\ \forall\ r\leq c$, si $C^r_{O_i}\neq\emptyset$ entonces

$$\upsilon_{T_r}(O_i) = 0 \Leftrightarrow \max_{\langle O_t,S\rangle\in C^r_{O_i}} \{\upsilon_{\bar{R}_\pi}(O_i,O_t,S)\} = 0$$

y

$$\upsilon_{T_r}(O_i) = 1 \Leftrightarrow \max_{\langle O_t,S\rangle\in C^r_{O_i}} \{\upsilon_{\bar{R}_\pi}(O_i,O_t,S)\} = 1$$

b) Por *i*) y *iii*) de la definición de Π tenemos que

$\forall\ r\leq c\ \forall\ O_i \in \Xi$

$$C^r_{O_i}=\emptyset \Rightarrow [NU_r = \{O_i\} \wedge \forall\ k\neq r\ [O_i\notin NU_k]]$$

y además

$\forall\ r\leq c\ \forall\ O_i \in \Xi$

$$[NU_r = \{O_i\}] \Rightarrow C^r_{O_i}=\emptyset$$

Asi que si existen $k \le c$ y $O_i \in \Xi$ tales que

$$NU_k=\{O \in \Xi \ / \ P_k(O)\}=\{O \in \Xi \ / \ \upsilon_{T_k}(O_i)\ge\varphi_{\beta_0}\}=\{O_i\}$$

entonces tenemos, considerando la conclusión 3.2., que

$$\forall \ \alpha>\varphi_{\beta_0} \ [NU_k=\{O \in \Xi \ / \ \upsilon_{T_k}(O_i)\ge\alpha\}=\{O_i\}]$$

es decir, que

$\forall \ r\le c \ \forall \ O_i \in \Xi$

$$C^r_{O_i}= \emptyset \Rightarrow [\upsilon_{T_r}(O_i)=1]$$

Definición 3.4.[18] Sea $\Pi_d(\Xi, \ \Gamma, \ \beta_0)$ un criterio agrupacional difuso sobre Φ que define una relación $R_\Pi \subseteq \Xi\times\Xi\times P(\Xi)$, una familia de núcleos $\tau = \{NU_1, \ ...,NU_c\}$ y una relación difusa $\tilde{R}_\Pi$ a partir de una familia de seudodifusiones $\tilde{F}_1,...,\tilde{F}_q$. Para cada núcleo $NU_r \in \tau$ definimos el subconjunto difuso NB_r de Ξ a partir de la función de pertenencia

$$\upsilon_{NB_r}(O_i) = \begin{cases} 1 & \text{si } C^r_{O_i}=\emptyset \\ \max\limits_{\langle O_j, S\rangle\in C^r_{O_i}} \{\upsilon_{\tilde{R}_\Pi}(O_i,O_j,S)\} & \text{en otro caso.} \end{cases}$$

el cual será denominado la **nube** NB_r del núcleo NU_r. ■

Proposición 3.1. Sea $\Pi_d(\Xi, \ \Gamma, \ \beta_0)$ un criterio agrupacional difuso sobre Φ que define una relación $R_\Pi \subseteq \Xi\times\Xi\times P(\Xi)$, una familia de núcleos $\tau = \{NU_1, \ ...,NU_c\}$ y una relación difusa $\tilde{R}_\Pi$ a partir de una familia de seudodifusiones $\tilde{F}_1,...,\tilde{F}_q$ la cual cumple que

$$R_\Pi(O_i,O_j,S) \Leftrightarrow \upsilon_{\tilde{R}_\Pi}(O_i,O_j,S)\ge f(\beta_0)$$

donde $f(\beta_0)$ un umbral en función de β_0, Π y de $\tilde{F}_1,...,\tilde{F}_q$. La familia de nubes $NB = \{NB_1, \ ...,NB_c\}$ donde cada NB_r es la nube correspondiente al núcleo NU_r perteneciente a la partición (cubrimiento) $\tau = \{NU_1, \ldots, NU_c\}$ generada por Π_d, es una seudodifusión de la partición (cubrimiento) τ.

Demostración

a) Si $C^r_{O_i} = \emptyset$ entonces $\upsilon_{T_r}(O_i)=1$ y tenemos que

$$\upsilon_{NB_r}(O_i) = \upsilon_{T_r}(O_i) = 1.$$

b) $\forall\ r\le c\ \forall\ O_i \in \Xi$, si $C^r_{O_i}\ne\emptyset$ tenemos que

1.- $\upsilon_{NB_r}(O_i)\ge f(\beta_0) \iff \max\limits_{<O_j,S>\in C^r_{O_i}} \{\upsilon_{\bar{R}_\pi}(O_i,O_j,S)\}\ge f(\beta_0) \iff$

$$\exists <O_j,S>\in C^r_{O_i}[\upsilon_{\bar{R}_\pi}(O_i,O_j,S)\ge f(\beta_0)] \iff$$

$$\exists <O_j,S>\in C^r_{O_i}[R_\pi(O_i,O_j,S)] \iff P_r(O_i) \iff \upsilon_{T_r}(O_i)\ge\varphi_{\beta_0}$$

es decir,

$\forall\ r\le c\ \forall\ O_i \in \Xi$

$$C^r_{O_i}\ne\emptyset \Rightarrow [\upsilon_{T_r}(O_i)\ge\varphi_{\beta_0} \iff \upsilon_{NB_r}(O_i)\ge f(\beta_0)]$$

y por tanto si $C^r_{O_i}\ne\emptyset$ y $C^k_{O_j}\ne\emptyset$ entonces

$\forall\ O_i, O_j \in \Xi\ \forall\ r,k\le c$

$$\upsilon_{T_r}(O_i) \ge \varphi_{\beta_0} > \upsilon_{T_k}(O_j) \Rightarrow \upsilon_{NB_r}(O_i) \ge f(\beta_0) > \upsilon_{NB_k}(O_j)$$

2.- Recordando (3) tenemos que

$\forall\ O_i, O_j \in \Xi\ \forall\ r,k\le c$, si $C^r_{O_i}\ne\emptyset$ y $C^k_{O_j}\ne\emptyset$ entonces

$$\upsilon_{T_r}(O_i)>\upsilon_{T_k}(O_j) \Rightarrow \max\limits_{<O_t,S>\in C^r_{O_i}} \{\upsilon_{\bar{R}_\pi}(O_i,O_t,S)\} \ge \max\limits_{<O_l,N>\in C^k_{O_j}} \{\upsilon_{\bar{R}_\pi}(O_j,O_l,N)\}$$

o lo que es lo mismo

$$\upsilon_{T_r}(O_i) > \upsilon_{T_k}(O_j) \Rightarrow \upsilon_{NB_r}(O_i) \ge \upsilon_{NB_k}(O_j)$$

Por *a*) y *b*) tenemos que si definimos el conjunto difuso $\tilde{F}$ como:

$\forall\ O_i \in \Xi,\ \forall\ r\le c$

$$\upsilon_{\tilde{F}}(\upsilon_{T_r}(O_i)\ge\varphi_{\beta_0}) = \upsilon_{NB_r}(O_i)$$

entonces $\tilde{F} \in D(\ge\varphi_{\beta_0})$.

Así pues, cada nube $NB_r \in NB$ es la seudodifusión según $\bar{F} \in D(\geq\varphi_{\beta_0})$ del núcleo $NU_r \in \tau$ y NB es una seudodifusión de la partición (cubrimiento) τ. ■

De todo esto podemos concluir que dado un criterio agrupacional difuso, el conjunto de nubes definido a partir de la familia de núcleos es un mapa del comportamiento de los objetos ("distancias") con respecto a los "centros conceptuales".

Ejemplos de criterios agrupacionales

Criterio β_0-conexo difuso[18]

Un subconjunto $NU_r \neq \emptyset$ de Ξ es una *componente β_0-conexa* si y sólo si

a) $\forall\, O_i, O_j \in NU_r\ \exists\ O_{i_1}, \ldots, O_{i_s} \in NU_r$

$[O_i = O_{i_1} \wedge O_j = O_{i_s} \wedge \forall\ p<s\ \Gamma(O_{i_p}, O_{i_{p+1}}) \geq \beta_0]$.

b) $\forall\ O_i \in \Xi\ [(O_j \in NU_r \wedge \Gamma(O_i, O_j) \geq \beta_0) \Rightarrow O_i \in NU_r]$.

c) Todo punto aislado constituye una componente β_0-conexa (*degenerada*).

Este conjunto de proposiciones definen la siguiente relación R_π:

$\forall\ O_i, O_j \in \Xi\ \forall\ S \subseteq \Xi,\ R_\pi(O_i, O_j, S)$ si y sólo si

$[\Gamma(O_i, O_j) \geq \beta_0 \wedge S = \emptyset]\ \vee$

$[\exists\ \{O_{i_1}, \ldots, O_{i_q}\} \in Su(S) \wedge \Gamma(O_i, O_{i_1}) \geq \beta_0 \wedge \Gamma(O_{i_q}, O_j) \geq \beta_0$

$\wedge\ \forall\ p<q\ \Gamma(O_{i_p}, O_{i_{p+1}}) \geq \beta_0]$.

donde $Su(S)$ es el conjunto de todos los ordenamientos sin repetición de todos los elementos de S.

Calculamos $\upsilon_{\bar{R}_\pi}(O_i, O_j, S)$ definiendo la seudodifusión de la β_0-semejanza como

$\forall\ O_i, O_j \in \Xi$

$$\upsilon_{\bar{F}}(\Gamma(O_i, O_j) \geq \beta_0) = \Gamma(O_i, O_j) \quad .$$

y así tenemos que

$\forall\ O_i, O_j \in \Xi\ \forall\ S\subseteq\Xi$

$$\upsilon_{\tilde{R}_\pi}(O_i,O_j,S)=\begin{cases}\Gamma(O_i,O_j) & \text{si } S=\varnothing\\ \max\limits_{(O_{i_1},\dots,O_{i_q})\in Su(S)}\{\min\{\Gamma(O_i,O_{i_1}),\\ \Gamma(O_{i_q},O_j),\min\limits_{p<q}\{\Gamma(O_{i_p},O_{i_{p+1}})\}\}\} & \text{si } S\neq\varnothing\end{cases}$$

Recordando que

$$C^r_{O_i}=\{\langle O_j,S\rangle \in (\Xi\backslash O_i)\times P(\Xi)\ /\ P_r(O_j)\wedge(R_\pi(O_i,O_j,S)\Leftrightarrow P_r(O_i))\}$$

tenemos entonces que

$$\langle O_j,S\rangle \in C^r_{O_i} \Rightarrow O_j \in NU_r \wedge S \subseteq NU_r\backslash\{O_j,O_i\}$$

Es decir, que los "responsables" de que un objeto O_i esté en el mismo núcleo NU_r (en la misma componente β_0-conexa) que un O_j son justamente elementos de NU_r. Todo esto implica que

$$\max_{\langle O_j,S\rangle\in C^k_{O_i}}\{\upsilon_{\tilde{R}_\pi}(O_i,O_j,S)\} = \max_{O_j\in NU_r\backslash O_i}\{\Gamma(O_i,O_j)\}$$

pues

$\forall\ S \subseteq NU_r\backslash\{O_i\}\ \forall\ \{O_{i_1},\dots,O_{i_q}\} \in Su(S)$

$$\max_{O_j\in NU_r\backslash O_i}\{\Gamma(O_i,O_j)\} \geq \max_{O_{i_1}\in S\subseteq NU_r\backslash O_i}\{\Gamma(O_i,O_{i_1})\}$$

asi que

$$\upsilon_{NB_r}(O_i) = \begin{cases}1 & \text{si } C^r_{O_i}=\varnothing\\ \max\limits_{O_j\in NU_r\backslash O_i}\{\Gamma(O_i,O_j)\} & \text{en otro caso.}\end{cases}$$

■

Criterio β_0-compacto difuso [18]

Para este criterio consideraremos que L=[0,1].

Diremos que un subconjunto $NU_r\neq\varnothing$ de Ξ es *β_0-compacto* si y sólo si

a) $\forall\, O_j \in \Xi\ [O_i \in B \wedge \max\limits_{O_t \in \Xi} \{\Gamma(O_i,O_t)\} = \Gamma(O_i,O_j) \geq \beta_0] \Rightarrow O_j \in B$

b) $\forall\, O_i,O_j \in B\ \exists\, O_{i_1},\ldots,O_{i_s} \in B\ [O_i = O_{i_1} \wedge O_j = O_{i_j} \wedge$

$\forall\, p<s\ [\max\limits_{O_t \in \Xi}\{\Gamma(O_{i_p},O_t)\} = \Gamma(O_{i_p},O_{i_{p+1}}) \geq \beta_0\ \vee$

$\max\limits_{O_t \in \Xi}\{\Gamma(O_{i_{p+1}},O_t)\} = \Gamma(O_{i_p},O_{i_{p+1}}) \geq \beta_0]]$.

c) Todo punto aislado constituye un conjunto β_0-compacto (*degenerado*).

Este conjunto de proposiciones definen la siguiente relación R_π:

$\forall\, O_i,O_j \in \Xi\ \forall\, S \subseteq \Xi,\ R_\pi(O_i,O_j,S)$ si y sólo si

$[\theta(O_i,O_j) \wedge S=\emptyset]\ \vee$

$[\exists\{O_{i_1},\ldots,O_{i_q}\} \in Su(S) \wedge \theta(O_i,O_{i_1}) \wedge \theta(O_{i_q},O_j)\ \wedge\ \forall\, p \leq q-1\ \theta(O_{i_p},O_{i_{p+1}})]$

donde Su(S) es el conjunto de los ordenamientos de los elementos de S y donde $\theta(O_i,O_j)$ denota

$\Gamma(O_i,O_j) \geq \beta_0 \wedge [\Gamma(O_i,O_j) = \max\limits_{O \in \Xi}\{\Gamma(O_i,O)\} \vee \Gamma(O_j,O_i) = \max\limits_{O \in \Xi}\{\Gamma(O_j,O)\}]$

Calculamos $\upsilon_{\tilde{R}_\pi}(O_i,O_j,S)$ definiendo las seudodifusiones de la β_0-semejanza y maximalidad como

$\forall\, O_i,O_j \in \Xi$

$$\upsilon_{\tilde{F}}(\Gamma(O_i,O_j) \geq \beta_0) = \frac{\Gamma(O_i,O_j) \cdot \beta_0}{\beta_0} = \Gamma(O_i O_j)$$

y

$$\upsilon_{\tilde{F}'}(\upsilon(\Gamma(O_i,O_j) = \max_{O \in \Xi}\{\Gamma(O_i,O)\}) \geq 1) = \frac{\Gamma(O_i,O_j) \cdot \beta_0}{\max\limits_{O \in \Xi}\{\Gamma(O_i,O)\}}$$

y así tenemos que

$\forall\, O_i,O_j \in \Xi\ \forall\, S \subseteq \Xi$

$$\upsilon_{\tilde{R}_\pi}(O_i,O_j,S) = \begin{cases} \upsilon(\theta(O_i,O_j)) & \text{si } S=\emptyset \\ \max\limits_{(O_{i_1},\ldots,O_{i_q}) \in Su(S)} \{\min\{\upsilon(\theta(O_i,O_{i_1})), & \\ \upsilon(\theta(O_{i_q},O_j)), \min\limits_{p<q}\{\upsilon(\theta(O_{i_p},O_{i_{p+1}}))\}\}\} & \text{si } S \neq \emptyset \end{cases}$$

donde

$$\upsilon(\Theta(O_i,O_j))=\min\{\Gamma(O_i,O_j),\ \max\{\frac{\Gamma(O_i,O_j)\cdot\beta_0}{\max\limits_{O\in\Xi}\{\Gamma(O_i,O)\}},\ \frac{\Gamma(O_j,O_i)\cdot\beta_0}{\max\limits_{O\in\Xi}\{\Gamma(O_j,O)\}}\}\}$$

Por otro lado, al igual que en el caso anterior

$$<O_j,S>\ \in C_0^r \Rightarrow O_j \in NU_r \wedge S \subseteq NU_r\setminus\{O_j,O\}$$

y entonces tenemos que

$$\max_{<O_j,S>\in C_{O_i}^k}\{\upsilon_{\tilde{R}_\pi}(O_i,O_j,S)\} = \max_{O_j\in NU_r\setminus O_i}\{\upsilon(\Theta(O_i,O_j))\}$$

es decir

$$\max_{<O_j,S>\in C_{O_i}^k}\{\upsilon_{\tilde{R}_\pi}(O_i,O_j,S)\} =$$

$$\max_{O_j\in NU_r\setminus O_i}\{\min\{\Gamma(O_i,O_j),\ \max\{\frac{\Gamma(O_i,O_j)\cdot\beta_0}{\max\limits_{O\in\Xi}\{\Gamma(O_i,O)\}},\ \frac{\Gamma(O_j,O_i)\cdot\beta_0}{\max\limits_{O\in\Xi}\{\Gamma(O_j,O)\}}\}\}$$

por lo que

$$\upsilon_{NB_r}(O_i)=\begin{cases} 1 & \text{si } C_{O_i}^r=\emptyset \\ \max\limits_{O_j\in NU_r\setminus O_i}\{\min\{\Gamma(O_i,O_j), & \\ \max\{\frac{\Gamma(O_i,O_j)\cdot\beta_0}{\max\limits_{O\in\Xi}\{\Gamma(O_i,O)\}},\ \frac{\Gamma(O_j,O_i)\cdot\beta_0}{\max\limits_{O\in\Xi}\{\Gamma(O_j,O)\}}\}\}\} & \text{en otro caso.} \end{cases}$$

■

Criterio β_0-fuertemente compacto difuso[18]

También para este criterio consideraremos que $L=[0,1]$.

Diremos que un subconjunto $NU_r \neq \emptyset$ de Ξ es *β_0-fuertemente compacto* si y sólo si

a) $\forall\ O_j \in \Xi\ [O_i \in NU_r \wedge \max\limits_{O\in\Xi}\{\Gamma(O_i,O)\} = \Gamma(O_i,O_j) \geq \beta_0] \Rightarrow O_j \in NU_r$

b) $\exists\ O_i \in NU_r\ \forall\ O_j \in NU_r,\ \exists\ O_{i_1},\ldots,O_{i_q} \in NU_r$

$[O_i=O_{i_1} \wedge O_j=O_{i_q} \wedge \forall\ p < q\ [\max\limits_{O\in\Xi}\{\Gamma(O_{i_p},O)\} = \Gamma(O_{i_p},O_{i_{p+1}}) \geq \beta_0]]$.

c) No existe NU_s subconjunto β_0-fuertemente compacto tal que $NU_r \subsetneq NU_s$.
d) Todo punto aislado constituye un conjunto β_0-fuertemente compacto *(degenerado)*.

Este conjunto de proposiciones definen la siguiente relación R_π:
$\forall\ O_i, O_j \in \Xi\ \forall\ S \subseteq \Xi,\ R_\pi(O_i, O_j, S)$ si y sólo si
$[S=\varnothing \wedge [\theta(O_i,O_j) \vee \theta(O_j,O_i)]] \vee$
$[\exists\ \{O_{i_1},\ldots,O_{i_q}\},\ \{O'_{i_1},\ldots,O'_{i_{q'}}\} \in Su(S) \wedge O_{i_1} = O'_{i_1} \wedge$
$\forall\ p \le q\ \theta(O_{i_p},O_{i_{p+1}}) \wedge \forall\ p \le q'\ \theta(O'_{i_p},O'_{i_{p+1}})$
$\wedge [[\theta(O_{i_q},O_i) \wedge \theta(O_{i_{q'}},O_j)] \vee [\theta(O_{i_q},O_j) \wedge \theta(O_{i_{q'}},O_i)]]]$

donde $\theta(O_i,O_j)$ denota $\Gamma(O_i,O_j) \ge \beta_0 \wedge \Gamma(O_i,O_j) = \max\limits_{O \in \Xi}\{\Gamma(O_i,O)\}$.

Calculamos $\upsilon_{\tilde{R}_\pi}(O_i,O_j,S)$ definiendo las seudodifusiones de la β_0-semejanza y maximalidad como en el caso β_0-compacto. Asi tenemos

$$\upsilon_{\tilde{R}_\pi}(O_i,O_j,S)=\begin{cases} \max\{\upsilon(\theta(O_i,O_j)),\upsilon(\theta(O_j,O_i))\} & \text{si } S=\varnothing \\ \max\{\max\limits_{\{O_{i_1},\ldots,O_{i_q}\}\in Su(S)}\{\min\{\min\limits_{p\le q-1}\{\upsilon(\theta(O_{i_p},O_{i_{p+1}}))\}, & \\ \max\{\min\{\upsilon(\theta(O_i,O_{i_1})),\upsilon(\theta(O_{i_q},O_j))\}, & \\ \min\{\upsilon(\theta(O_j,O_{i_1})),\upsilon(\theta(O_{i_q},O_i))\}\}\}\}, & \\ \max\limits_{\{O_{i_1}\ldots O_{i_q}\}\in Su(S)}\ \max\limits_{0<t<q}\{\max\{\min\{\min\limits_{p\le t-1}\{\upsilon(\theta(O_{i_{p+1}},O_{i_p}))\}, & \\ \min\limits_{t\le p\le q-1}\{\upsilon(\theta(O_{i_p},O_{i_{p+1}}))\},\max\{\min\{\upsilon(\theta(O_{i_1},O_i)), & \\ \upsilon(\theta(O_{i_q},O_j))\},\min\{\upsilon(\theta(O_{i_1},O_j)),\upsilon(\theta(O_{i_q},O_i))\}\}\}\} & \text{si } S\neq\varnothing \end{cases}$$

donde

$$\upsilon(\theta(O_i,O_j)) = \min\{\Gamma(O_i,O_j),\ \frac{\Gamma(O_i,O_j)\cdot\beta_0}{\max\limits_{O\in\Xi}\{\Gamma(O_i,O)\}}\}$$

Observemos que en este caso R_π no define una partición si no que define un cubrimiento. Recordando nuevamente que

$$C^r_{O_i}=\{<O_j,S> \in (\Xi\backslash O_i)\times P(\Xi) \ / \ P_r(O_j) \wedge (R_\pi(O_i,O_j,S) \Leftrightarrow P_r(O_i))\}$$

vemos que en este caso los "responsables" de que un objeto O_i esté en el mismo núcleo NU_r que un objeto O_j son elementos de NU_r si O_i pertenece a NU_r; y son elementos que pertenecen a NU_r y que no pertenecen a ningún otro núcleo al que también pertenezca O_i si O_i no pertenece a NU_r. Es decir

$$<O_j,S> \in C^r_{O_i} \Rightarrow O_j \in NU_r \wedge S \subseteq B^r_{O_i}$$

y

$$<O_j,\varnothing> \in C^r_{O_i} \Rightarrow O_j \in B^r_{O_i}$$

donde

$$B^r_{O_i}= \{O_t\in NU_r/O_t\neq O_i \wedge [O_i\notin NU_r \Rightarrow \forall\, NU \in \tau\backslash NU_r\ [O_t\in NU \Rightarrow O_i\notin NU]]\}$$

Vemos entonces que

$$\max_{<O_j,S>\in C^k_{O_i}} \{\upsilon_{\bar{R}_\pi}(O_i,O_j,S)\}= \max_{O_j\in B^r_{O_i}} \{\max\{\upsilon(\Theta(O_i,O_j)),\upsilon(\Theta(O_j,O_i))\}$$

es decir

$$\max_{<O_j,S>\in C^k_{O_i}} \{\upsilon_{R_\pi}(O_i,O_j,S)\} =$$

$$\max_{O_j\in B^r_{O_i}}\{\max\{\min\{\Gamma(O_i,O_j),\frac{\Gamma(O_i,O_j)\cdot\beta_0}{\max\limits_{O\in\Xi}\{\Gamma(O_i,O)\}}\},\min\{\Gamma(O_i,O_j),\frac{\Gamma(O_j,O_i)\cdot\beta_0}{\max\limits_{O\in\Xi}\{\Gamma(O_j,O)\}}\}\}\}$$

por lo que

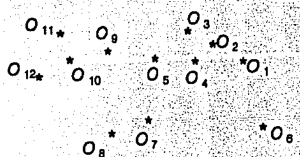
$$\upsilon_{NB_r}(O_i)=\begin{cases} 1 & \text{si } C^r_{O_i}=\varnothing \\ \max\limits_{O_j\in B^r_{O_i}}\{\max\{\min\{\Gamma(O_i,O_j),\ \dfrac{\Gamma(O_i,O_j)\cdot\beta_0}{\max\limits_{O\in\Xi}\{\Gamma(O_i,O)\}}\}, & \\ \min\{\Gamma(O_i,O_j),\ \dfrac{\Gamma(O_j,O_i)\cdot\beta_0}{\max\limits_{O\in\Xi}\{\Gamma(O_j,O)\}}\}\}\} & \text{en otro caso.} \end{cases}$$

■

Ejemplos

A continuación presentamos unos ejemplos de la aplicación de estos criterios difusos a un conjunto de datos de una matriz de semejanza $\{\Gamma_{ij}\}$, donde $\Gamma_{ij} = \Gamma(O_i, O_j)$, $L = [0,1]$ y $\beta_0 = 0.80$.

$$
\begin{array}{c}
 \\ O_1 \\ O_2 \\ O_3 \\ O_4 \\ O_5 \\ O_6 \\ O_7 \\ O_8 \\ O_9 \\ O_{10} \\ O_{11} \\ O_{12}
\end{array}
\begin{array}{c}
\begin{array}{cccccccccccc}
O_1 & O_2 & O_3 & O_4 & O_5 & O_6 & O_7 & O_8 & O_9 & O_{10} & O_{11} & O_{12}
\end{array} \\
\begin{bmatrix}
1.00 & .85 & .75 & .78 & .63 & .70 & .51 & .38 & .43 & .27 & .22 & .13 \\
.85 & 1.00 & .90 & .90 & .76 & .62 & .58 & .45 & .56 & .38 & .36 & .27 \\
.75 & .90 & 1.00 & .88 & .83 & .50 & .60 & .47 & .65 & .47 & .47 & .33 \\
.78 & .90 & .88 & 1.00 & .85 & .58 & .68 & .55 & .63 & .45 & .42 & .30 \\
.63 & .76 & .83 & .85 & 1.00 & .48 & .73 & .65 & .80 & .65 & .58 & .50 \\
.70 & .62 & .50 & .58 & .48 & 1.00 & .51 & .36 & .27 & .11 & .05 & 0.00 \\
.51 & .58 & .60 & .68 & .73 & .51 & 1.00 & .83 & .66 & .58 & .48 & .51 \\
.38 & .45 & .47 & .55 & .65 & .36 & .83 & 1.00 & .66 & .63 & .53 & .61 \\
.43 & .56 & .65 & .63 & .80 & .27 & .66 & .66 & 1.00 & .83 & .80 & .70 \\
.27 & .38 & .47 & .45 & .65 & .11 & .58 & .63 & .83 & 1.00 & .88 & .85 \\
.22 & .36 & .47 & .42 & .58 & .05 & .48 & .53 & .80 & .88 & 1.00 & .80 \\
.13 & .27 & .33 & .30 & .50 & 0.00 & .51 & .61 & .70 & .85 & .80 & 1.00
\end{bmatrix}
\end{array}
$$

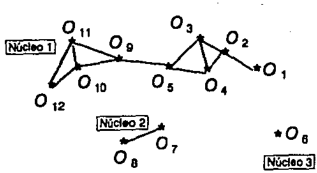


$X = \{O_1, O_2, O_3, O_4, O_5, O_6, O_7, O_8, O_9, O_{10}, O_{11}, O_{12}\}$

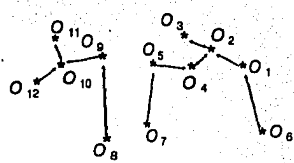
β_0-conexo difuso

Familia de núcleos:

$\tau = \{\{O_1, O_2, O_3, O_4, O_5, O_9, O_{10}, O_{11}, O_{12}\}, \{O_7, O_8\}, \{O_6\}\}$



Núcleos del criterio β_0-conexo difuso



Nube 1 del criterio β_0-conexo difuso

La familia de nubes resultante la representaremos con una matriz $\{\upsilon_{ij}\}$, donde $\upsilon_{ij} = \upsilon_{NB_i}(O_j)$:

$$
\begin{array}{c}
\\ NB_1 \\ NB_2 \\ NB_3
\end{array}
\begin{array}{c}
\begin{array}{cccccccccccc} O_1 & O_2 & O_3 & O_4 & O_5 & O_6 & O_7 & O_8 & O_9 & O_{10} & O_{11} & O_{12} \end{array} \\
\left[\begin{array}{cccccccccccc}
.85 & .90 & .90 & .90 & .85 & .70 & .73 & .66 & .83 & .88 & .88 & .85 \\
.51 & .58 & .60 & .68 & .73 & .51 & .83 & .83 & .66 & .63 & .53 & .61 \\
.70 & .62 & .50 & .58 & .48 & 1.00 & .51 & .36 & .27 & .11 & .05 & 0.00
\end{array}\right]
\end{array}
$$

■

β_0-compacto difuso

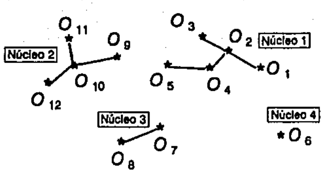
Familia de núcleos:

$\tau = \{\{O_1, O_2, O_3, O_4, O_5\}, \{O_9, O_{10}, O_{11}, O_{12}\}, \{O_7, O_8\}, \{O_6\}\}$

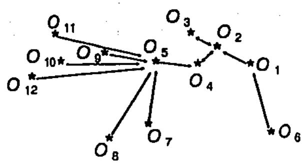
La familia de nubes resultante nos define la matriz $\{\upsilon_{ij}\}$

$$
\begin{array}{c}
\\ NB_1 \\ NB_2 \\ NB_3 \\ NB_4
\end{array}
\begin{array}{c}
\begin{array}{cccccccccccc} O_1 & O_2 & O_3 & O_4 & O_5 & O_6 & O_7 & O_8 & O_9 & O_{10} & O_{11} & O_{12} \end{array} \\
\left[\begin{array}{cccccccccccc}
.80 & .80 & .80 & .80 & .80 & .70 & .70 & .62 & .77 & .61 & .55 & .47 \\
.41 & .54 & .63 & .60 & .77 & .21 & .64 & .64 & .80 & .80 & .80 & .80 \\
.49 & .56 & .58 & .65 & .70 & .51 & .80 & .80 & .64 & .61 & .51 & .59 \\
.70 & .62 & .50 & .58 & .48 & 1.00 & .51 & .36 & .27 & .11 & .05 & 0.00
\end{array}\right]
\end{array}
$$

con $\upsilon_{ij} = \upsilon_{NB_i}(O_j)$. ■



Núcleos del criterio β_0-compacto difuso



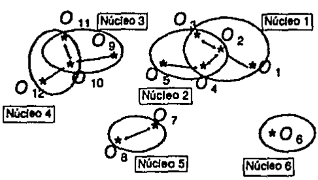
Nube 1 del criterio β_0-compacto difuso

β_0-fuertemente compacto difuso.
Familia de núcleos:
$\tau = \{ \{O_1,O_2,O_3,O_4\},\ \{O_2,O_3,O_4,O_5\},\ \{O_9, O_{10},O_{11}\},\ \{O_{10},O_{11},O_{12}\},$
$\{O_7,O_8\},\ \{O_6\} \}$

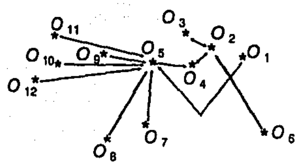
La familia de nubes resultante nos define la matriz $\{\upsilon_{ij}\}$ siguiente:

$$
\begin{array}{c}
\\ NB_1 \\ NB_2 \\ NB_3 \\ NB_4 \\ NB_5 \\ NB_6
\end{array}
\begin{array}{c}
\begin{array}{cccccccccccc}
O_1 & O_2 & O_3 & O_4 & O_5 & O_6 & O_7 & O_8 & O_9 & O_{10} & O_{11} & O_{12}
\end{array} \\
\begin{bmatrix}
.80 & .80 & .80 & .80 & .59 & .70 & .65 & .53 & .62 & .42 & .42 & .31 \\
.59 & .80 & .80 & .80 & .80 & .62 & .70 & .62 & .77 & .61 & .54 & .47 \\
.41 & .54 & .62 & .60 & .77 & .21 & .64 & .64 & .80 & .80 & .80 & .67 \\
.25 & .34 & .43 & .41 & .61 & .11 & .56 & .60 & .67 & .80 & .80 & .80 \\
.49 & .56 & .58 & .65 & .70 & .51 & .80 & .80 & .64 & .61 & .51 & .59 \\
.70 & .62 & .50 & .58 & .48 & 1.00 & .51 & .36 & .27 & .11 & .05 & 0.00
\end{bmatrix}
\end{array}
$$

donde $\upsilon_{ij} = \upsilon_{NB_i}(O_j)$. ■



Núcleos del criterio β_*-fuertemente compacto difuso



Nube 2 del criterio β_*-fuertemente compacto difuso

El Nuevo Modelo de Algoritmos Difusos de Análisis de Cúmulos

Ahora presentamos la estructura paramétrica del nuevo modelo de algoritmos:

Sea dado un espacio $\Phi = \{\Delta, \Gamma\}$ de representación cuyo conjunto de descripciones de objetos en términos del conjunto de rasgos

$M = \{r_1, \ldots, r_m\}$ denotamos como $\Xi = \{O_1, \ldots, O_n\}$. Un algoritmo A del modelo para la formación de cúmulos se puede expresar como:

$$A(\Gamma,\ \beta_0,\ \Pi,\ \bar{F})$$

donde Γ es una función de semejanza cuya definición está en dependencia de: C_R un conjunto de criterios de comparación de valores de cada una de los rasgos, Ω_A un sistema de conjuntos de apoyo, y δ_{Ω_A} una familia de funciones de semejanza parcial de las ω-partes de Ω_A; β_0 es un umbral de semejanza a partir del cual se define el concepto de semejanza o analogía entre objetos con base en los valores de Γ; Π un criterio agrupacional con parámetros Γ, Ξ y β_0; y $\bar{F}$ una familia de seudodifusiones $\{\bar{F}_1, \ldots, \bar{F}_q\}$ de Π. Los algoritmos de esta familia pueden ser obtenidos de la siguiente manera:

Paso 1.- Definición de la función de semejanza Γ.

Paso 1.1.- Definir C_R el conjunto de criterios de comparación de valores de cada una de los rasgos.

Paso 1.2.- Definir Ω_A un sistema de conjuntos de apoyo.

Paso 1.3.- Definir δ_{Ω_A} una familia de funciones de semejanza parcial de las ω-partes de Ω_A.

Paso 1.4.- Definir una función de semejanza Γ.

Paso 2.- Determinar el umbral β_0 de semejanza.

Paso 3.- Definición del criterio agrupacional difuso Π_d.

Paso 3.1.- Definir un criterio agrupacional Π con parámetros Γ, Ξ y β_0.

Paso 3.2.- Definir la relación R_π a partir de Π.

Paso 3.3.- Definir la familia de seudodifusiones $\{\bar{F}_1, \ldots, \bar{F}_q\}$ y a partir de ella y R_π definir la relación difusa $\bar{R}_\pi$.

Paso 4.- Generar la familia τ de núcleos.

Paso 5.- Definir la familia de nubes correspondiente a τ.

Dada la flexibilidad en el manejo de distintas funciones de semejanza por parte del modelo contamos con la posibilidad, como se mencionó en el capítulo 2, de manipular objetos cuyas descripciones no estan definidas exclusivamente en términos cuantitativos. Por otro lado, tenemos la opción de utilizar criterios agrupacionales difusos que definen familias de agrupaciones sin que se requiera definir de antemano el número de éstas y para los cuales la función de semejanza sólo tiene que cumplir que su imagen sea un orden lineal (β_0-conexo difuso) o que sea el intervalo [0,1] (β_0-compacto difuso y β_0-fuertemente compacto difuso).

CONCLUSION

En el presente trabajo se ha propuesto un nuevo modelo de algoritmos para el análisis de cúmulos difusos para la solución de problemas de reconocimiento de patrones sin aprendizaje que, dadas sus características, parece adecuado para su aplicación en problemas con condiciones como las planteadas en la introducción. Por un lado, no se limita al manejo de descripciones de objetos en términos exclusivamente cuantitativos, no es necesario definir de antemano el número de agrupaciones, y se presentan condiciones muy relajadas para la definición de la función de semejanza. Por otro lado, el parámetro correspondiente al criterio agrupacional está condicionado a la definición de un conjunto de proposiciones, de un umbral de semejanza entre descripciones y a un conjunto de seudodifusiones, siendo esto último el punto más delicado ya que los valores que tomen diferentes seudodifusiones en un mismo criterio han de ser comparables (como en la β_0-compacidad, las seudodifusiones de la β_0-semejanza y la maximalidad). Sin embargo es de suponer que la definición o determinación de este tipo de conceptos no debe de resultar muy ajena a un especialista de alguna disciplina que se propone plantear un problema de análisis de cúmulos. Por otro lado es de hacerse notar que la naturaleza del conjunto de nubes resultante (difusa, dura, partición, cubrimiento, etcétera) depende exclusivamente del criterio y las seudodifusiones, es decir, este modelo de algoritmos no "obliga" la naturaleza de las agrupaciones resultantes. Por último, al respecto del criterio agrupacional, este tipo de algoritmos utiliza de hecho un par de relaciones, una dura y la otra una "difusión" de la primera. En ese sentido este modelo es susceptible a ser ampliado a muy diversos tipos de criterios agrupacionales con diversos tipos de parámetros.

En cuanto a la familia NB de nubes resultante de aplicar estos algoritmos podemos comentar lo siguiente: de la familia NB se obtiene la matriz M

$$\begin{bmatrix} \nu_{NB_1}(O_1) & \nu_{NB_1}(O_2) & \cdots & \nu_{NB_1}(O_n) \\ \nu_{NB_2}(O_1) & \nu_{NB_2}(O_2) & \cdots & \nu_{NB_2}(O_n) \\ \vdots & \vdots & \vdots & \vdots \\ \nu_{NB_c}(O_1) & \nu_{NB_c}(O_2) & \cdots & \nu_{NB_c}(O_n) \end{bmatrix}$$

la cual brinda una valiosa información desde el punto de vista de las aplicaciones, a saber: las filas, que son el grado de pertenencia de los objetos de Ξ a cada nube NB_r, mapea el orden entre los valores de cumplimiento de los objetos de Ξ a la propiedad T_r; las columnas, que son el grado con que cada objeto pertenece a las diferentes nubes, nos describe el orden entre los valores de cumplimiento de cada objeto de Ξ a las diferentes propiedades $T_1, \ldots, T_c$[6]; y la matriz M en sí nos describe no sólo el comportamiento de los objetos con respecto a las propiedades sino también el comportamiento de las mismas propiedades con respecto al conjunto de objetos.

Se debe hacer notar que si L=[0,1], la familia NB de nubes no cumple necesariamente la condición de ser una partición difusa en el sentido de Ruspini, es decir, que puede ocurrir que la suma

$$\sum_{r=1}^{c} \nu_{NB_r}(O)$$

con $O \in \Xi$ sea distinta de 1. En ese caso bastaría con normalizar la función de pertenencia mediante

$$\nu_{NB_r}'(O) = (\nu_{NB_r}(O) \ / \ \sum_{r=1}^{c} \nu_{NB_r}(O)) \qquad (1)$$

6 Aquí adoptamos para los grados de pertenencia de los objetos a las diferentes nubes la intepretación de Dubois y Prade [9]: "The grades of membership reflect an ordering of the objects in the universe, induced by the predicate associated with (the fuzzy set) $\tilde{X}$; this ordering when its exits, is more important than the membership values themselves".

para tener una partición difusa en ese sentido, ya que de las condiciones de la definición de β_0-partición (o cubrimiento) difusa y (1) se obtiene la condición de que

$$0 < \sum_{i=1}^{n} \nu_{NB_r}(O_i) < n, \text{ para } r = 1,\ldots,c.$$

Sin embargo, como ya se ha comentado, la matriz M sin normalizar guarda información con respecto a los objetos de Ξ y sus propiedades que al normalizar se pierden, a saber, la información que brindan las filas y la totalidad de la matriz misma. Se puede plantear que el hecho de que la suma de las pertenencias de los objetos a una nube no sea un mismo número, más que una extrañeza debe de considerarse como un hecho informativo más dentro del problema concreto.

Finalmente es importante subrayar que con este trabajo se abre la posibilidad de continuar un conjunto de investigaciones en el área de la Teoría de Análisis de Cúmulos bajo una nueva óptica basada en el principio de difusificación de los núcleos obtenidos sobre la base de criterios de agrupamientos que respondan a las exigencias de los modelos de los especialistas de las áreas de aplicación en cuestión.

Referencias

[1] G. Augustson y J. Minker. "An analysis of some graph theorical cluster techniques". J. Ass. Comput Mach. 17, pp. 571-588, 1971.

[2] G. H. Ball y D. J. Hall. "A clustering technique for summarizing multivariate data". Behav. Sci., Vol. 12, pp. 153-155, 1967.

[3] R. Bellman y M. Giertz. "On the analytic formalism of the Theory of Fuzzy Sets". Inf. Sci. 5, pp. 149-157, 1973.

[4] R. Bellman, R. Kalaba y L. A. Zadeh. "Abstraction and Pattern Classification". J. Math. Anal. Applic. 13, pp. 1-7, 1966.

[5] J. C. Bezdek. "Fuzzy mathematics in Pattern Classification". Ph.D. Thesis, Applied Math. Center, Cornell University, Ithaca, 1973.

[6] J.C. Bezdek y J.C. Dunn. "Optimal fuzzy partitions: A heuristic for estimating the parameters in a mixture of normal distributions". IEEE Trans. Comp., Vol. C-24, pp. 835-838. 1975.

[7] J. C. Bezdek. "Pattern Recognition with fuzzy objetive function algorithms". Plenum Press, 1981.

[8] R. Duda. P. Hart. "Pattern Classification and Scene Analysis". Wiley, New York, 1973.

[9] D. Dubois, H. Prade. "Fuzzy set and systems: theory and aplications". Academic Press, 1980.

[10] J. C. Dunn. "A fuzzy relative of the ISODATA process and its use in detecting compact, well separated clusters". J. Cybern., Vol. 3, pp. 32-57, 1974.

[11] I. Gath y A. Geva. "Unsupervised optimal fuzzy clustering". IEEE Trans. Pattern Anal. Mach. Intell. Vol. PAMI-11. No. 7, pp. 773-781. 1989.

[12] J. Goguen. "L-fuzzy sets". J. Math. Anal. Appl. 18, pp. 145-174, 1967.

[13] J. Goguen. "Concept representation in natural and artificial languages: axioms, extensions and applications for fuzzy sets". Int. J. Man-Mach. Stud. 6, pp. 513-561, 1969.

[14] D. Gustafson y W. Kessel. "Fuzzy clustering with a fuzzy covariance matrix". Proc. IEEE-CDC Vol. 12. pp.761-766, IEEE Press, Piscataway, New Jersey, 1979.

[15] N. Jardine y R. Sibson. "Mathematical taxonomy". New York, Wiley, 1971.

[16] R. Ling. "On the theory and construction of K-clusters". Comput. J. 15, pp. 326-332, 1972.

[17] J. Montellano y J. Ruiz Shulcloper. "Un nuevo modelo de técnicas difusas para problemas de análisis de cúmulos", (En proceso).

[18] J. Ruiz Shulcloper y J. Montellano. "A new aproach to fuzzy clustering", (Preprint).

[19] J. Ruiz Shulcloper y M. Lazo Cortés. "Modelos matemáticos para el Reconocimiento de Patrones". UNAM, 1992. (En prensa).

[20] E. Ruspini. "A new approach to clustering". Inf. Cont. Vol. 15, pp. 22-32, 1969.

[21] S. Sirotinskaia. "Metodología de análisis de la información geológica". Nedra, 1986.

[22] T. Tanimoto. "An elementary mathematical theory of classsification and prediction". Report, IBM. Corp. 1958.

[23] L. Voronin, G. Karataeva y E. Epshtein. "Programas 'Holotipo' para la solución de problemas de Reconocimiento de Patrones". Alma Ata, 1968.

[24] L. Zadeh. " Fuzzy sets". Inf. Control 8, pp. 338-353, 1965.

[25] L. Zadeh. "Simlarity relations and fuzzy orderings". Inform. Sci., vol. 3, pp. 177-200, 1971.

[26] W. Zangwill. "Non-Linear programming: A unified approach". Prentice-Hall, Englewood Cliffs, New Jersey, 1969.

Bibliografía

-Brian Everitt. "Cluster analysis". Heinemann Educational Books, 1977.

-D. Dubois, H. Prade. "Fuzzy sets and systems: theory and applications". Academic Press, 1980.

-H. J. Zimmerman. "Fuzzy Set Theory and its applications". Kluwer Academic Publishers, 1991.

-A. Kaufmann. "Introducction to the Theory of Fuzzy Subsets". Academic Press, 1975.

-J. Ruiz Shulcloper y M. Lazo Cortés. "Modelos matemáticos para el Reconocimiento de Patrones". UNAM, 1992. (En prensa).

-J. Bezdek y S. Pal. "Fuzzy models for Pattern Recognition. Methods that search for structures in data". IEEE Press, 1992.

-J. Bezdek. "Pattern Recognition with fuzzy objetive function algorithms". Plenum Press, 1981.

-L.Escudero. "Reconocimiento de Patrones". Paraninfo. Madrid, 1977.