Representación de conceptos difusos mediante índices semánticos.

L.J.Linares
Dep. Informática. ESI . UCLM.
13071 Ciudad Real. España
ljimenez@inf-cr.uclm.es

J.A.Olivas
Dep. Informática. ESI . UCLM.
13071 Ciudad Real. España
jaolivas@inf-cr.uclm.es

Resumen

Magnitudes inciertas o imprecisas son representadas habitualmente mediante conjuntos difusos. Estos conjuntos se definen por medio de una función de pertenencia $\mu : X \to [0,1]$ sobre los elementos de un referencial X. Esta representación funcional de los conceptos no nítidos obliga al uso del principio de extensión para transladar operaciones o conceptos definidos en otros dominios. El uso del principio de extensión basado en el concepto de alfa-corte de un conjunto difuso origina en algunos casos la necesidad de recurrir a difíciles cálculos funcionales. En este trabajo se presenta una representación alternativa y compatible con la teoría de conjuntos difusos, definiendo un conjunto de características semánticas propias del carácter impreciso de este tipo de conceptos, que tratadas conjuntamente permiten manipular esos conceptos de una forma alternativa al principio de extensión.

Palabras clave: Índices semánticos, prototipos.

1 Índices asociados a conjuntos difusos.

Varios son los autores que han descrito una serie de índices reales, con los que reflejar la información representada por un conjunto difuso. Ejemplo de estos índices, son los criterios de desborrosificación del máximo, la media del máximo o el centro de masas $CM(\mu)$ de un conjunto difuso (establecido sobre un referencial X) , definido por

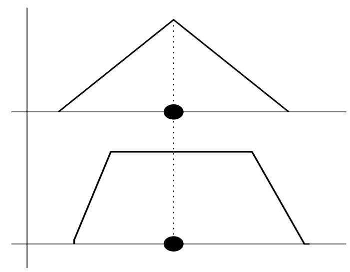

Figura 1: Centro de masas de distintos conjuntos difusos

$$CM(\mu) = \frac{\int \mu(x)x dx}{\int \mu(x) dx} \qquad (1)$$

reflejando un único valor real para la totalidad del conjunto difuso.

Este tipo de simplificaciones son efectivas para algunos tipos de aplicaciones como son el control difuso y el ranking de conjuntos difusos; sin embargo, no son válidas para otros problemas que necesitan más información. Un problema esencial en este tipo de índices únicos, es la imposibilidad de extraer de ellos el conjunto difuso que los originó. En la figura 1 se representan dos conjuntos difusos con características sustancialmente distintas que poseen el mismo valor para cualquiera de los índices señalados.

Otro de los problemas que posee el índice del CM, es que puede desplazar el grueso de la información aportada por el conjunto difuso hacia valores anormales o donde el grado de pertenencia de los elementos del dominio no es el más representativo. Un ejemplo se puede observar en la figura 2 cuando

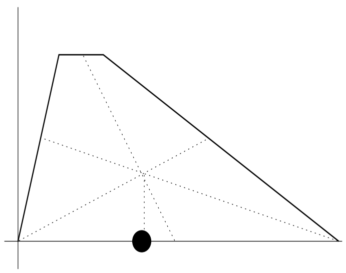

Figura 2: Centro de masas con valor de pertenencia inferior a 1.

el $CM(\mu)$ se sitúa fuera de donde se encuentran los valores de máxima pertenencia .

Este efecto se vería amplificado si al conjunto difuso se le asociase una etiqueta, y a ésta como valor representativo o nominal su CM.

Mediante otro enfoque M. Delgado, M. A. Vila y W. Voxman [3][4] proponen una representación canónica para los números difusos como un par de índices reales $(V_s(\mu), A_s(\mu))$, (Valor, Ambigüedad), definidos como :

$$V_s(\mu) = \int_0^1 s(r)\,[L_\mu(r) + R_\mu(r)]\,dr \qquad (2)$$

$$A_s(\mu) = \int_0^1 s(r)\,[R_\mu(r) - L_\mu(r)]\,dr \qquad (3)$$

donde s(r) es una función de reducción, definida como una función creciente $s : [0,1] \to [0,1]$ con las propiedades s(0)=0 y s(1)=1, que permite ponderar la influencia de los r-cortes que definen a $L_\mu(r)$ y $R_\mu(r)$ como:

$$L_\mu(r) = \begin{cases} inf\,\{x|x \in \mu_r\} & x \in (0,1] \\ inf\,\{x|x \in Supp(\mu)\} & r = 0 \end{cases} \qquad (4)$$

$$R_\mu(r) = \begin{cases} sup\,\{x|x \in \mu_r\} & x \in (0,1] \\ sup\,\{x|x \in Supp(\mu)\} & r = 0 \end{cases} \qquad (5)$$

siendo $\mu_r = \{x|\mu(x) \geq r\}$ y $Supp(\mu) = \{x|\mu(x) > 0\}$.

El valor de un número difuso $V_s(\mu)$ representa el valor central de una magnitud definida mediante el número difuso, mientras que la ambigüedad $A_s(\mu)$ representaría "la forma" del mismo.

Esta representación, mediante el par de índices anteriormente descritos, tiene la ventaja que determinados números difusos, aquellos que están centrados, pueden reconstruirse a partir de sus índices. La generalización de este reconstrucción inversa para cualquier número difuso obliga a la sustitución del índice de ambigüedad $A_s(\mu)$ por otros dos índices $A_L(\mu)$ y $A_R(\mu)$, definidos en función de $q = (L_\mu(1) + R_\mu(1))/2$, punto central del intervalo modal del conjunto difuso μ, como

$$A_L(\mu) = \int_0^1 s(r)\,[q - L_\mu(r)]\,dr \qquad (6)$$

$$A_R(\mu) = \int_0^1 s(r)\,[R_\mu(r) - q]\,dr \qquad (7)$$

2 Representación de un número difuso mediante sus índices semánticos.

Desde nuestro punto de vista, son varios los conceptos que quedan reflejados mediante un conjunto difuso. Centrándose en las magnitudes expresadas por medio de un número difuso trapezoidal, se pueden destacar los siguientes:

1. *Generalidad de la magnitud.* Una magnitud representada mediante un conjunto difuso será más o menos general conforme sea mayor o menor el número de elementos del referencial sobre los que dicha magnitud es relevante. El valor que define dicha generalidad $G(\mu)$ será $G(\mu) = (d - a)$ siendo $a = inf\{Supp(\mu)\}$ y $d = sup\{Supp(\mu)\}$.

2. *Imprecisión de una magnitud.* Una magnitud representada mediante un conjunto difuso será más o menos imprecisa conforme sea mayor o menor el número de elementos del referencial sobre los que dicha magnitud está perfectamente definida. El valor que define dicha imprecisión será $I(\mu) = (c - b)$ siendo $b = inf\{Mod(\mu)\}$, $c = sup\{Mod(\mu)\}$ y $Mod(\mu) = \{x|\mu(x) = 1\}$.

3. *Ambigüedad de una magnitud.* Relación entre la imprecisión y generalidad de la magni-

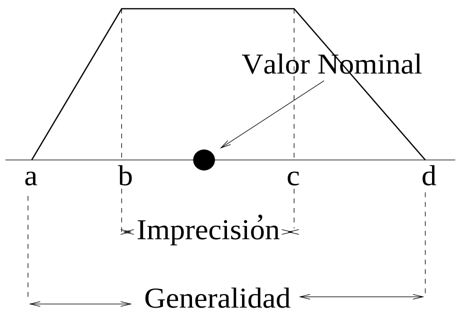

Figura 3: Índices semáticos

tud expresada como $A(\mu) = I(\mu)/G(\mu)$, ó 0 si $G(\mu) = 0$.

4. *Valor nominal de una magnitud.* Valor del conjunto $Mod(\mu)$ elegido como representante de la magnitud en toda su perfección, se definirá como $VN(\mu) = (c - b)/2$.

En la figura 3 se muestran estas características para el caso de un número difuso trapezoidal.

Se define la *representación semántica* de un número difuso como:

$$RS([a, d], VN, A) \tag{8}$$

El intervalo [a,d] proporciona la información concisa sobre la generalidad de la magnitud, que junto a los valores VN y A establecen el resto de valores destacable como el de la imprecisión I=A*(d-a).

Dado un número trapezoidal T(a,b,c,d) su representación semántica es única y viene definida como

$$RS\left([a, d], \frac{b + c}{2}, \frac{c - b}{d - a}\right) \tag{9}$$

Esta representación es reversible, en el sentido de que a partir de un representación semántica dada $RS([a, d], VN, A)$ se puede obtener el número difuso trapezoidal resolviendo el siguiente par de ecuaciones

$$\begin{array}{c} b + c = 2VN \\ c - d = A(d - a) \end{array} \tag{10}$$

Operando en (10) se tiene que

$$\begin{array}{c} c = VN + \frac{A(d-a)}{2} \\ b = VN - \frac{A(d-a)}{2} \end{array} \tag{11}$$

Resultando el número trapezoidal

$$T(a, VN - \frac{A(d - a)}{2}, VN + \frac{A(d - a)}{2}, d)$$

cuya función de pertenecia $\mu(x)$ se define como

$$\mu(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{VN - \frac{A(d-a)}{2} - a} & a \leq x < VN - \frac{A(d-a)}{2} \\ 1 & \begin{array}{l} VN - \frac{A(d-a)}{2} \leq x \\ x \leq VN + \frac{A(d-a)}{2} \end{array} \\ \frac{d-x}{d - \left(VN + \frac{A(d-a)}{2}\right)} & VN + \frac{A(d-a)}{2} < x \leq d \\ 0 & d < x \end{cases} \tag{12}$$

3 Prototipo semántico de una colección de valores no nítidos.

Cuando se realiza un análisis de algún tipo de información, parece de especial relevancia la capacidad de determinar un objeto típico o prototipo de una colección de objetos. De igual forma, la idea de la tipicalidad de un valor o valor típico juega un papel importante para el razonamiento por sentido común, es más, estos valores son muchas veces utilizados como valores por defecto.

Es habitual la necesidad de determinar un valor típico en situaciones en las que se dispone de una colección de observaciones que queremos categorizar o extraer las características que las identifican. Yager[5] expresa el concepto de prototipo o valor típico como " *el valor que es él mismo o muy similar para muchas de las observaciones que estamos intentando tipificar.*"

Sea el concepto de prototipo de una colección de elementos, aquel elemento que posee unas características que reflejan a la generalidad de la colección.

De manera más formal, considérese una colección $L = \{e_1, \ldots, e_m\}$ de m elementos, cada uno de ellos $e_i = (x_1^i, \ldots, x_n^i)$ definidos mediante n características $X_1 \times \ldots \times X_n$. Sea d_i una medida de distancia de los valores de la característica i-esima. Se definirá el prototipo $Prot(L) = (p_1, \ldots, p_n)$ como el conjunto de valores, de las características que minimizan las distancias entre los valores posibles de cada característica:

$$d_i(x, p_i) \leq d(x, y) \, \forall x, y \in X_i \tag{13}$$

Si se toma una colección de elementos no nítidos, definidos mediante una representación semántica

$$e_i = RS_i([a_i, d_i], VN_i, A_i) \tag{14}$$

Se define el prototipo Prot(L), ecuación (15), como

$$RS\left(\left[\frac{1}{m}\sum_{i=1}^{m} a_i, \frac{1}{m}\sum_{i=1}^{m} d_i\right], \frac{1}{m}\sum_{i=1}^{m} VN_i, \frac{1}{m}\sum_{i=1}^{m} A_i\right) \tag{15}$$

donde cada índice semántico es la media de los valores de los índices semánticos de la colección de elementos.

Este enfoque se puede utilizar para determinar el prototipo de una colección de número difusos.

En primer lugar se obtendrá la representación semántica de cada número difuso, luego se calculará el prototipo mediante la ecuación (15); para finalmente reconstruir el número difuso prototipo de la colección.

4 Conclusiones y trabajos futuros

Lo que se ha pretendido es definir y mostrar la utilidad de índices semánticos para la caracterización reversible de conjuntos borrosos representados por números borrosos trapezoidales.

En base a esta definición, se introduce la definición de prototipo de un conjunto de números borrosos, que se intuye más operativa y significativa que las que se usan habitualmente.

La proyección de este trabajo pasa por extender el concepto de índice semántico a otras caracterizaciones de conjuntos difusos, así como hacia una operatividad para manipular sistemas de reglas y sistemas de control con esta concepción.

Finalmente se comprobará el funcionamiento en sistemas reales, aplicaciones en las que ya se han usado técnicas clásicas, tanto de manipulación de conceptos borrosos como de uso de la noción de prototipo derivada de estos conceptos.

Referencias

[1] Kóczy, L. T., Hirota, K.: "Ordering, distance and closeness of fuzzy sets", Fuzzy Sets and Systems 59 (1993) 281-293.

[2] Kóczy, L. T., Hirota, K.: "Approximate Reasoning by Linear Rule Interpolation and General Approximation", International Journal of Approximate Reasoning (1993); 9:197-225.

[3] M Delgado, M. A. Vila, W. Voxman "On canonical representation of fuzzy numbers." Fuzzy Sets and Systems 93 (1998) 125-135.

[4] M Delgado, M. A. Vila, W. Voxman "A fuzzy measure for fuzzy numbers: Applications." Fuzzy Sets and Systems 94 (1998) 205-216.

[5] Ronald R. Yager. “A note on fuzzy measure of typicality.” International Journal of intelligent systems. VOL 12, (1997) 233-249

[6] Zhu, Q., Lee, E. S.: "Comparison and Ranking of Fuzzy Numbers", en Fuzzy Regression Analysys, J. Kacprzyk y M. Fedrizzi (eds.), (1992), Physica-Verlag.