Fuzzy set theory by H.-J. Zimmermann.

Néstor Rodríguez Vico - nrv23@correo.ugr.es

3 de marzo de 2019

1. Introducción.

La mayoría de las herramientas tradicionales para el modelado formal, el razonamiento y la computación son *crisp* y deterministas. Por ejemplo, en la lógica dual tradicional una afirmación puede ser verdadera o falsa; en la teoría de conjuntos, un elemento puede pertenecer a un conjunto o no; en problemas de optimización una solución puede ser factible o no. Desafortunadamente esto no se justifica si es importante que el modelo describa bien la realidad.

Durante mucho tiempo, la teoría de la probabilidad y la estadística han sido las teorías predominantes para modelar incertidumbre en los sistemas, ambas se basadas en axiomas. Sin embargo, en paralelo, se han desarrollado otras teorías. La teoría de conjuntos difusos es una de estas teorías, originalmente concebida como una extensión de la lógica dual y la teoría de conjuntos clásica.

2. Historia.

En las primeras publicaciones de la teoría de conjuntos difusos de Zadeh y Goguen muestran la intención de los autores de generalizar la noción clásica de un conjunto y una propuesta para acomodar la confusión. Zadeh escribió: "... un conjunto difuso proporciona un punto de partida para la construcción de un marco conceptual similar en muchos aspectos al marco utilizado en el caso de los conjuntos ordinarios, pero más general y, potencialmente, con un espectro de aplicaciones más amplio ... Dicho marco proporciona una forma natural de abordar problemas en los que la fuente de la imprecisión es la ausencia de criterios claramente definidos de pertenencia a una clase y no la presencia de variables aleatorias". La imprecisión aquí se entiende en el sentido de la vaguedad y no de la falta de conocimiento sobre el valor de un parámetro. La teoría de conjuntos difusos proporciona un marco matemático estricto en el que los fenómenos conceptuales difusos pueden ser estudiados de forma precisa y rigurosa.

La aceptación de esta teoría creció lentamente en los años 60-70 del siglo XX. En la segunda mitad de los años 70, surgieron las primeras aplicaciones prácticas exitosas en el control de procesos tecnológicos usando sistemas basados en reglas difusas, llamados control difuso, incrementando así el interés en esta área. En 1984 ya existían unas 4000 publicaciones y en el año 2000 más de 30.000. En 1992, en tres conferencias simultáneas en Europa, Japón y Estados Unidos, las tres áreas de la teoría de conjuntos difusos, las redes neuronales y la computación evolutiva (algoritmos genéticos) unieron fuerzas y en adelante se conocen como "inteligencia computacional". De manera similar, el término "soft computing" se utiliza para una serie de enfoques que tratan esencialmente de la incertidumbre y la imprecisión.

3. Teoría matemática y evidencia empírica.

En esta sección podemos ver de forma detallada la formulación matemática detrás de los conjuntos difusos.

3.1. Definiciones.

Me gustaría destacar las definiciones que, desde mi punto de vista, me han parecido más relevantes e interesantes:

Definición 1 "Si X es una colección de objetos denotados genéricamente por x, entonces un conjunto difuso en X es un conjunto de pares ordenados:

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)|x \in X)\} \tag{1}$$

donde $\mu_{\tilde{()}A}(x)$ es la función de pertenencia, la cual mapea X al espacio de pertenencia."

Definición 2 "Un conjunto difuso tipo m (con $m > 1$) es un conjunto difuso cuyos valores de la función de pertenencia son tipo $m - 1$ conjuntos difusos en $[0, 1]$." Esta definición parece rebuscada, pero nos permite definir conjuntos difusos cuyas funciones de pertenencia son, a su vez, conjuntos difusos, lo cual es bastante interesante.

Definición 4 Esta definición nos permite definir variables lingüísticas. Una variable lingüística es una variable cuyos posibles valores son palabras y pueden ser representados mediante conjuntos difusos. Estás variables permiten describir el estado de un objeto o suceso. Una variable lingüística permite que sus valores sean *etiquetas lingüísticas.* Estas son términos lingüísticos definidos como conjuntos difusos sobre un dominio subyacente.

Definición 7 El soporte de un conjunto difuso $\tilde{A}$, $S(\tilde{A})$, es el conjunto de los elementos *crisp* de todos los $x \in X$ tal que $\mu_{\tilde{A}}(x) > 0$. También podemos definir un α-corte como:

$$A_\alpha = \{x \in X|\mu_{\tilde{A}}(x) \geq \alpha\} \tag{2}$$

Es decir, el conjunto *crisp* de los elementos que pertenencen a un conjunto difuso $\tilde{A}$ con un grado mayor que α.

Definición 9 La cardinalidad de un conjunto difuso $|\tilde{A}|$ se define como:

$$|\tilde{A}| = \sum_{x \in X} \mu_{\tilde{A}}(x) \tag{3}$$

3.2. Operaciones.

Al igual que podemos hacer operaciones con conjuntos *crisp* podemos hacerlo con conjuntos difusos:

Definición 10 Intersección: La función de pertenencia de la intersección de dos conjuntos difusos $|\tilde{A}|$ y $|\tilde{B}|$ se define como:

$$\mu_{\tilde{A} \cap \tilde{B}}(X) = Min(\mu_{\tilde{A}}(X), \mu_{\tilde{B}}(X)) \forall x \in X \tag{4}$$

Definición 11 Unión: La función de pertenencia de la unión de dos conjuntos difusos $|\tilde{A}|$ y $|\tilde{B}|$ se define como:

$$\mu_{\tilde{A} \cap \tilde{B}}(X) = Min(\mu_{\tilde{A}}(X), \mu_{\tilde{B}}(X)) \forall x \in X \tag{5}$$

Definición 12 Complemento: La función de pertenencia de la unión de dos conjuntos difusos $|\tilde{A}|$ y $|\tilde{B}|$ se define como:

$$\mu_{\tilde{A}}(X) = 1 - \mu_{\tilde{A}}(X) \forall x \in X \tag{6}$$

Aparte de las opreaciones aquí descritas, podemos más operaciones relacionadas con las *t-normas* y las *t-conormas*.

3.3. El principio de extensión.

Supongamos X como el producto cartesiando de los universos $X = X_1 \times ... \times X_r$ y $\tilde{A}_1, ..., \tilde{A}_r$ conjuntos difusos definidos en $X_1, ..., X_r$ respectivamente. f es una función que mapea elementos del universo X al universo Y, $y = f(x_1, ..., x_r)$. Entonces, el principio de extensión nos permite definir un conjunto difuso $\tilde{B}$ en el universo Y mediante:

$$\tilde{B} = \{(y, \mu_{\tilde{B}}(y)) | y = f(x_1, ..., x_r), (x_1, ..., x_r) \in X\} \tag{7}$$

donde

$$\mu_{\tilde{B}}(y)) = \begin{cases} sup_{(x_1,...,x_r) \in f^{-1}(y)} \; min\{\mu_{\tilde{A}_1}(x_1), ..., \mu_{\tilde{A}_r}(x_r)\} & si \; f^{-1}(y) \neq \Theta \\ 0 & en \; caso \; contrario \end{cases} \tag{8}$$

y f^{-1} es la función inversa de f.

3.4. Relaciones difusas y grafos.

Al igual que en la subsección 3.1, voy a comentar las definiciones que más relevantes me han parecido:

Definición 16: Sea $X, Y \subseteq \mathbb{R}$ conjuntos universales, entonces:

$$\tilde{R} = \{((x,y), \mu_{\tilde{R}}(x,y))|(x,y) \subseteq X \times Y\} \tag{9}$$

es una relación difusa en $X \times Y$.

Definición 17: Una relación difusa que es reflexiva, simétrica y transitiva se denomina relación de similitud.

3.5. Análisis difuso.

Una función difusa es una generalización de una función clásica. Una función clásica es un mapeo del dominio de definición de la función en un espacio. Hay diferentes características de la definición clásica que pueden ser consideradas difusas en vez de *crisp*:

1. Puede haber un mapeo *crisp* de un conjunto difuso que lleve consigo la borrosidad del dominio y genere un conjunto difuso.
2. El mapeo en sí puede ser borroso, lo que “fuzzyfica” la imagen de un argumento nítido. Dubois y Prade la llaman “fuzzifying function”.
3. Las funciones clásicas pueden tener propiedades o restricciones difusas.

4. Aplicaciones.

En esta sección voy a comentar las tres aplicaciones que, desde mi punto de vista, me han pareco más interesantes y más en relación con lo estudiado en esta asignatura.

4.1. Lógica difusa, razonamiento aproximado y razonamiento pausible.

En la lógica dual, los valores de verdad pueden ser *verdadero* o *falso* y los operadores se definen mediante tablas de verdad. En la lógica difusa, los valores de verdad se expresan mediante las variables lingüísticas *verdadero* y *falso*. Además, en el razonamiento aproximado las frases pueden ser conjuntos difusos. En el razonamiento plausible, una sentencia A en la premisa no tiene que ser idéntica a la sentencia A en la implicación. En todas las formas de razonamiento difuso, las implicaciones pueden ser modeladas de diferentes formas. El elegir la más apropiado depende del contexto, puede ser elegido de forma empírica, axiomáticamente o en base a su eficiencia, según el problema que estemos abordando.

4.2. Sistemas basados en reglas difusas (sistemas expertos difusos y control difuso).

Los sistemas basados en el conocimiento son sistemas en los que los algoritmos son sustituidos por una base de conocimientos y un motor de inferencia. Existen diferentes maneras de adquirir y almacenar conocimientos especializados. Normalemente, este conocimiento se almacena mediante reglas del tipo *si X entonces Y*. Estas reglas se procesan en el motor de inferencia para obtener una decisión. A este tipo de sistemas se les llama *sistemas expertos*.

En los años 70, las reglas clásicas fueron sustituidas por afirmaciones difusas y el motor de inferencia tuvo que ser adaptado para ser capaz de inferir usando afirmaciones difusas. A estos nuevos sistemas se les denomina *unidades de cálculo*. Las primeras aplicaciones exitosas de estos sistemas fueron en la ingeniería de control. La entrada a estos sistemas eran los valores *crisp* producidos por el proceso y tenía que ser "fuzzificada" en valores difusas. Estos valores con las sentencias difusas se procesan en la unidad de cálculo generando conjuntos difusos como salida. Estos conjuntos difusos debías ser transformados en valores *crisp* para poder ser usados por los sistemas. Este proceso se denominó "desfuzzificación".

Hay sistemas que buscan reemplazar a los expertos humanos y, por lo tanto, el resultado debe ser lingüístico. En estos casos, se sustituye la "desfuzzificación" por una "aproximación lingüística", para proporcionar un resultado entendible por un ser humano.

4.3. Minería de datos difusa.

Es evidente que, con el desarrollo del procesaminto de datos, cada vez disponemos de más datos digitales, lo cual nos lleva a una situación en la cual era imposible extraer información y conocimiento de tal cantidad de datos. Los métodos clásicos para extraer datos seguían existiendo pero a no se adaptaban a las necesidades. Por ejemplos, las técnicas de *clústering* suponían que los datos podían dividirse *clústeres*, los cuales que no se adaptaban a la realidad. Con la teoría de conjuntos difusos aparecen maneras de mejorar los técnicas existentes. Bezdek fue uno de los primeros en desarrollar métodos de clústeres difusos con los objetivos de buscar estructuras en los datos que permitan reducir la complejidad de los algoritmos, aportando información para el control y la toma de decisiones. Bezdek desarrolló el algoritmo *FCM: fuzzy c-means*.

En este campo se ha investigado en diferentes enfoques: algoritmos jerárquicos, algoritmos heurísticos semiformales y *clústering* guiado por funciones objetivo. Aún así, dado que el área de la minería de datos está en auge, es un campo en el que se está investigando mucho en la actualidad.

5. Conclusión.

Como podemos ver, el estudio de los conjuntos difusos comenzó hace bastante tiempo, en la década de los 60. Además, considero bastante relevante comentar que es algo que surgió de manera natural, simplementa por la necesidad de representar, de una manera más real, la información en un sistema. De esta necesidad de representar información que no se adapta a un mundo dual (si o no, verdadero o falso, cero o uno), surgen los conjuntos difusos.

Uno de los aspectos más interesantes de este artículo es la reflexión sobre la formalidad matemática que hay detrás de los conjuntos difusos. La principal característica de éstos es la capacidad de representar información imprecisa. Aún así, tal y como comenta H.-J. Zimmermann, no hay nada de imprecisión en la base matemática subycente a la teoría de conjuntos difusos. Todo está perfectamente estudiado, demostrado y definido. Esto lo podemos ver reflejado en la gran sección de teoría matemática presente en el artículo y toda la investiagación realizada en esta temática.

Finalmente, una de la sección que más me ha gustado es la sección de aplicaciones ya que, desde mi punto de vista, siempre que surge una nueva técnica, un nuevo algoritmo o, en general, una nueva rama de investiagción, lo más importante es la aplicación de la misma. Desde mi punto de vista, en el caso de los conjuntos difusos, este árticulo comenta algunas aplicaciones más "tradicionales" y algunas más "innovadoras". De la primera categoría me gustaría destacar la lógica difusa y los sitemas basados en reglas. De la segunda, la minería de datos difusa y, en particular, me ha impresionado la posibilidad de crear algoritmos de *clústering* difusos y, por lo tanto, adaptar algoritmos clásicos para poder tratar con información difusa.