

La construcción del dato con valores lingüísticos usando subconjuntos borrosos

Julio C. Acosta*
julioa@exa.unne.edu.ar

La Teoría de los Subconjuntos Borrosos es, de hecho, un paso hacia un acercamiento entre la precisión de las matemáticas clásicas y la sutil imprecisión del mundo real, un acercamiento nacido de la incesante búsqueda humana por lograr una mejor comprensión de los procesos mentales del conocimiento.

Lofti Zadeh (1982)

Resumen

Se presenta una ajustada síntesis de elementos de lógica difusa y subconjuntos borrosos, sus orígenes y desarrollo hasta la actualidad; elementos de las técnicas de formación y operaciones elementales, para rescatar una analogía existente entre las componentes del indicador de la matriz de datos (dimensión – procedimiento) y los elementos que conforman un subconjunto borroso (calificación – sustento – regla de formación). Se muestra la correspondencia existente entre el modelo cuatripartito de matriz de datos con la formación de los subconjuntos borrosos como modelo para ponderar unidades de análisis que requieran el uso de magnitudes lingüísticas y, eventualmente, cuando sus determinaciones de pertenencia se hallen en un margen de incertidumbre. Desde esa perspectiva se resalta que los modelos difusos no son meras construcciones formales, sino que tienen una interpretación concreta en la realidad de las percepciones humanas. Finalmente, se proveen algunos elementos básicos de las primeras herramientas necesarias para abrir el interrogante si será posible formalizar a través de lógica difusa el modo en que sucede el conocimiento y la valoración en la mente humana.

Palabras clave: matriz de datos – magnitudes lingüísticas – indicador

* Julio C. Acosta es Ingeniero en Construcciones UNNE, docente de los departamentos Matemática e Informática de la Facultad de Ciencias Exactas y Naturales y Agrimensura UNNE. Cursa la Maestría en Metodología de la Investigación Científica UNLa con subsidio del C.S. UNNE.

Abstract

It is presented an adjusted synthesis of elements of fuzzy logic and fuzzy subsets, their origins and development until the present time; elements of the formation techniques and elementary operations, to rescue an existent analogy among the components of the indicator of the matrix of data (dimension – procedure) and the elements that conform a blurred subset (qualification – I sustain – membership function). The existent correspondence is shown among the pattern cuatripartito of matrix data with the formation of the fuzzy subsets as model to ponder analysis units that require the use of linguistic magnitudes and possibly when its determinations of ownership are in a margin of uncertainty; and from that perspective it is stood out that the fuzzy models are not mere formal constructions, but rather they have a concrete interpretation in the reality of the human perceptions. Finally, some basic elements of the first necessary tools are provided to open the query if it will be possible to formalize through fuzzy logic the way in that it happens the knowledge and the valuation in the human mind.

Keywords: matrix data – linguistic magnitudes – indicator

1. Lógica Difusa

Una de las disciplinas matemáticas de reciente aparición es la llamada lógica difusa o borrosa, (Lofti Zadeh – 1965), que a diferencia de la lógica simbólica, en la que las expresiones son verdadero (1) o falso (0), introduce herramientas que son capaces de manipular expresiones que no son ni totalmente ciertas ni completamente falsas, es decir, esta lógica puede ser aplicada a conceptos que pueden tomar un valor cualquiera de veracidad, dentro de un conjunto de valores que oscilan entre dos extremos: la verdad absoluta o la falsedad total. Se deja a salvo que no es la lógica como tal lo que resulta impreciso, difuso o vago, sino los objetos que con ella se estudian y se expresan de tal manera; se caracteriza por *la falta de definición del concepto al que se aplica*. La lógica difusa es usada para tratar información imprecisa en términos de subconjuntos borrosos que se combinan en reglas para definir acciones del tipo: *Si la velocidad es alta, entonces reduzca un poco la aceleración; si los medicamentos están bien prescriptos entonces el enfermo sanará pero (y) los medicamentos no están “tan bien prescriptos”, “es posible” que el enfermo no se sane.*

Puede considerarse como antecedente inmediato de la lógica difusa a la lógica trivaluada de Lukasiewicz, en la que a los valores de verdad biestables (0 – 1) y a las operaciones de disyunción (\vee) y conjunción (\wedge) de la lógica simbólica expresadas en la Figura 1, se le suma la alternativa de valores de verdad trivaluados (0 – $\frac{1}{2}$ – 1) para las expresiones; esto es que: a p falso se le asigna cero, a p verdad se le asigna uno, pero p puede no ser totalmente verdad ni completamente falso, y entonces se le asigna medio; frente a las tablas de verdad para la conjunción y disyunción de la lógica simbólica bivalente, Lukasiewicz construye las tablas de verdad de disyunción y conjunción para su lógica trivaluada como se muestra en la Figura 1.

p	q	$p \vee q$	p	q	$p \wedge q$
1	1	1	1	1	1
1	0	1	1	0	0
0	1	1	0	1	0
0	0	0	0	0	0

Lógica simbólica bivaluada

p	q	$p \vee q$	p	q	$p \wedge q$
1	1	1	1	1	1
1	½	1	1	½	½
1	0	1	1	0	0
½	1	1	½	1	½
½	½	½	½	½	½
½	0	½	½	0	0
0	1	1	0	1	0
0	½	½	0	½	0
0	0	0	0	0	0

Lógica simbólica trivaluada

FIGURA 1

El ingeniero Lofti Zadeh en la Universidad de Berkeley (California EEUU), en el año 1965 publicó en la revista *Information and Control* un trabajo que resultó ser el fundador de esta disciplina, “Fuzzy Sets”; la teoría de los subconjuntos borrosos al principio encontró resistencia en la comunidad científica, no obstante el propio Zadeh y algunos seguidores de la nueva lógica continuaron desarrollando los fundamentos de la teoría de los subconjuntos borrosos, a principio de 1970 se establecieron en Japón varios grupos de trabajo en lógica difusa, también en ambientes hostiles en los primeros años. En 1974 en el Reino Unido, E. H. Mamdani y Seto Assilian desarrollan el primer controlador difuso para una máquina de vapor, pero la primera implementación real de esta tecnología debió esperar hasta 1980 en una planta cementera en Dinamarca. Japón logra las primeras aplicaciones en 1983 cuando la empresa Fuji aplica lógica difusa para el control de inyección química en plantas depuradoras de agua y en 1987 la firma Hitachi pone en marcha un controlador difuso para la supervisión del tren-metro de Sendai. Se introduce también la lógica difusa en las redes neuronales y Bart Kosko hace importantes aportes a los llamados *neuro-fuzzy systems*, (sistemas neuronales difusos). Hoy el control difuso está presente en la mayoría de los electrodomésticos de última generación (las cámaras filmadoras y fotográficas regulan sus lentes con lógica difusa, los lavarropas inyectan mayor o menor cantidad de jabón según el nivel de suciedad de la ropa con la misma técnica), está presente en la industria metalúrgica, en los puertos que usan controladores difusos en sus grúas y también en los aviones que vuelan con controladores difusos.

Es sorprendente que el éxito de lógica difusa haya sucedido principalmente en el campo del control automático de procesos (tecnología), cuando en rigor, lo que posiblemente buscaba Zadeh era crear un formalismo para manipular de forma mas eficiente la imprecisión y la vaguedad del razonamiento humano expresado lingüísticamente.

2. Subconjuntos borrosos

Un conjunto (colección de objetos) se puede denotar por extensión o por comprensión; por extensión lo hacemos listando todos los elementos que lo conforman y por comprensión

definiendo las propiedades que tienen en común los elementos (todos ellos y solo ellos). Se suele llamar al modo de describir los elementos por comprensión, definición a través de una *función característica* de los elementos del conjunto.

Un conjunto borroso extiende la idea de conjunto convencional (al que por contraposición se lo suele denominar conjuntos *nítido* –crisp–) permitiendo que sus elementos tomen “valores” diferentes de la pertenencia total o la no pertenencia. Por ejemplo, para denotar el conjunto de las personas menores de 20 años, simbólicamente:

$$A = \{x/x \text{ es persona menor de 20 años}\}$$

es suficiente tratar con conjuntos clásicos –nítidos–, en razón que los personas menores de 20 años pertenecerán al conjunto A y los mayores de 20 años (o los que tengan 20 años), no pertenecerán a dicho conjunto.

Si deseamos caracterizar el conjunto de las personas jóvenes, simbólicamente:

$$B = \{x/x \text{ es persona joven}\}$$

aparecen en principio dos “inconvenientes”: ¿qué se entiende exactamente por persona joven?, cuando se haya logrado consenso respecto de esto, supongamos por ejemplo, son jóvenes todos aquellos cuyas edades oscilan entre 16 y 30 años, aparece el otro aspecto “difuso” de la cuestión: ¿es acaso una persona de 18 años “tan joven” como otra de 29 años?; y aparece ahora la necesidad de referirnos a los subconjuntos borrosos, donde quede expresado (¿cuantificado?) por ejemplo, que no se pasa de joven a adulto “en un día”, en nuestro caso precisamente al cumplir los 31 años¹. A través de un subconjunto borroso puede quedar definido el *valor* (etiqueta) *joven*; entonces en el conjunto de las personas será posible reconocer elementos que “están mas incluidos que otros” en este subconjunto borroso *joven*.

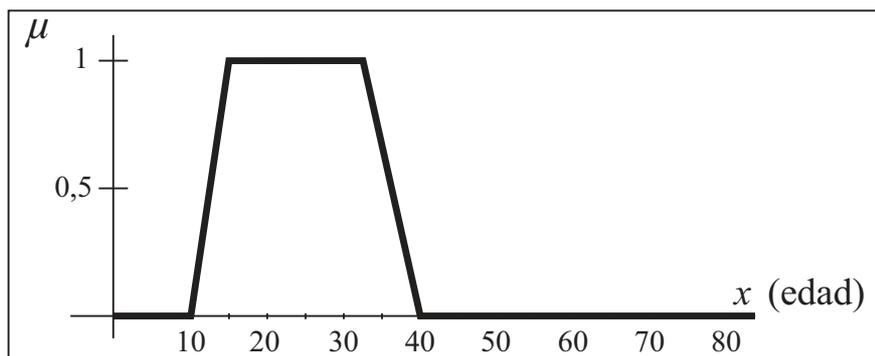


FIGURA 2

¹ Estos dos aspectos, principalmente el primero introducirá en la ponderación del dato con subconjuntos borrosos la presencia (=influencia) del contexto.

En el gráfico se observa que a las edades comprendidas entre 16 y 30 años les corresponde $\mu = 1$, esto es porque al intervalo que contiene esas edades le asignamos la plenitud de juventud; pero se aprecia también que la persona no se hace joven en un día, sino que, por ejemplo, desde los 12 años comienza a poseer “rasgos” de juventud, hasta que a los 16 es “plenamente” joven. En igual sentido, a los 30 años se inicia un período de extinción de juventud que culmina, en nuestro ejemplo, a los 40 años.²

La asignación de *sustento* a las diferentes edades queda definida por la función:

$$\mu_A(x) = \begin{cases} 0 & \text{si } 0 < x < 12 \\ 0.25x - 3 & \text{si } 12 \leq x < 16 \\ 1 & \text{si } 16 \leq x \leq 30 \\ -0.1x + 4 & \text{si } 30 < x \leq 40 \\ 0 & \text{si } x > 40 \end{cases}$$

Informalmente, diremos entonces que los subconjuntos borrosos se tratan de conjuntos cuyas fronteras no están definidas de manera totalmente nítida y la pertenencia de los elementos quedan expresadas en una función de asignación de sustento.

Así $A = \int_U \frac{\mu_A(x)}{x}$ cuando los x pertenecen a un universo continuo,³

si x pertenece a un universo discreto se escribe $A = \sum_U \frac{\mu_A(x)}{x}$

en las expresiones precedentes ni el símbolo \int debe entenderse como integral en el sentido clásico que le asigna el análisis matemático, ni el símbolo \sum debe entenderse como sumatoria en el sentido clásico que le asigna el álgebra, sino en ambos casos debe entenderse como la simple colección de valores continuos en un caso y discretos en el otro.

3. Dato

A la complejidad estructural del dato científico que propone Johan Galtung⁴ con las tres componentes: la unidad de análisis (UA), las variables (V) y los valores (R), se le adiciona una cuarta que Juan Samaja⁵ llama *indicadores*. En rigor, Galtung propone que el término variable solo es un sustituto del término *estímulo* cuando nos encontramos frente a conceptos teóricos no susceptibles de actos de medición. El hecho es que al estudiar las componentes

² La definición de las edades que abarcan los períodos es totalmente arbitraria y puestas por el autor, pero debe ser conformada por especialistas y en función del tipo de trabajo que se trate.

³ Todas las edades posibles en nuestro caso.

⁴ Cfr. Galtung, Johan. *Teoría y Método de la Investigación Social*, Buenos Aires, Eudeba, 1978.

⁵ Cfr. Samaja, Juan. *Epistemología y Metodología. Elementos para una teoría de la investigación científica*. Buenos Aires, Eudeba, 2005.

del dato desde la perspectiva de la conformación de los valores expresados en subconjuntos borrosos se encuentra que, *necesariamente* ante la presencia de variables lingüísticas, aparecen los *indicadores* en la conformación de los valores lingüísticos expresados en subconjuntos borrosos; y precisamente, *si mediante subconjuntos borrosos podemos llegar a cuantificar expresiones subjetivas (no susceptibles de medición en matemática clásica), entonces, ya no habrán mas conceptos teóricos no susceptibles de medición.*

Si definimos el *dato* como aquella información infradeterminada por la praxis⁶ y que es una respuesta a un código usado para interpretar o interpelar a la realidad; entonces trataremos de ahondar en la construcción del código como tema de discusión en la generación del dato. Autores como Galtung ponen todo el interés en la selección de la o las dimensiones relevantes para interpretar el sentido global de la variable, pero no le asignan rango epistemológico a los procedimientos que se deben realizar para “observar” el hecho. Para estos autores, la dimensión puede ser observable por si misma; Samaja sostiene la tesis kantiana, según la cual ninguna dimensión, en tanto es concepto, puede ser observable por si misma, en consecuencia es necesario *un término medio* que *enlace* las *sensaciones* (que es lo único dado a los sentidos) con *la dimensión*, esto es *el esquema*.

El investigador se encuentra frecuentemente en situaciones donde es posible determinar (y “visualizar”) perfectamente cada una de las componentes de una matriz de dato, si no todas, al menos alguna; si la Unidad de Análisis aparece nítida, se buscan cuales variables interesan al trabajo y a partir de ahí se definen los valores que puede tomar esa variable; o bien puede suceder que desde la definición de la(s) variable(s) que nos interesa conocer nos imbuimos de las unidades de análisis que son afectadas por ellas; el hecho es que *siempre* con la definición de la variable aparece la definición de los valores que ella puede tomar; y llegamos ahora al centro de la problemática a la que busca aportar este trabajo; si los valores que pueden asignarse a la variable resultan expresados en *cifras* (números), todo parece⁷ estar “bajo control”. Y en efecto esto sucede toda vez que si se trata de estudiar el desplazamiento de un objeto O identificado como unidad de análisis y de esa unidad de análisis, abordamos en particular la variable *velocidad* de desplazamiento (podría ser otra variable la trayectoria recorrida en el desplazamiento, por ejemplo); la velocidad a la que se espera se desplace el objeto varía de 0 (en reposo) a 2 m/s, y por alguna razón sabemos que el objeto no puede desplazarse a más de esa velocidad (y si lo hiciera no nos interesa), entonces, es claro que el valor que tome la UA_1 será por ejemplo 1,8 m/s; la misma variable para la UA_2 valdrá 0,7 m/s y así sucesivamente para cada UA_i . En este ejemplo, la cuarta componente *indicador* surge de una dimensión a la que se aplica algún procedimiento para construir el indicador, así en nuestro ejemplo dimensiones de la variable velocidad pueden ser el espacio recorrido y el tiempo empleado en recorrerlo, y el procedimiento algún método para medir las dimensiones, cuyos resultados aplicados en una formula adecuada nos dan el valor buscado.

⁶ Entre la teoría y la empiria hay algo mas básico: la praxis. Esta no es empiria ni teoría.

⁷ Veremos más adelante que no siempre es así.

4. Valores

Los valores que pueden ser asignados a una variable dependerán principalmente de la variable que se trate, como también de las herramientas que se usen en la medición, de las condiciones de medición entre otras, diremos que toman valores nítidos (numéricos) cuando los valores se expresan en números que resultan de una sola medición; pero también pueden expresarse en números que resultan de varias mediciones y de la aplicación de dichas mediciones en expresiones (formulas) previamente determinadas, estos se llaman índices y se expresan también como valores nítidos, pero puede suceder que frente a una determinada ponderación (por ejemplo capacidad del director de la obra) los resultados posibles sean: bueno, regular, malo; aparecen acá *valores lingüísticos*, que no siempre pueden ser expresados en números (índices)⁸ y son estos valores los que serán tratados con subconjuntos borrosos.

Puede ser el caso de un trabajo donde los resultados arrojen los siguientes valores de la variable que en la lectura de los resultados (o sus consecuencias) estén “estratificados”; por ejemplo, a valores de velocidad comprendidos entre 0 m/s y 1 m/s y por otro lado entre 1 m/s y 2 m/s. Se podrían proponer “rangos” de variación, ejemplo:

0.00 - 1.00		1.00 - 2.00		
0.00 - 0.50	0.50 - 1.50		1.50 - 2.00	
0.00 - 0.25	0.25 - 0.75	0.75 - 1.25	1.25 - 1.75	1.75 - 2.00

FIGURA 3

0.00 - 1.00		baja	0.00 - 0.25		muy baja
1.00 - 2.00		alta	0.25 - 0.75	baja	
			0.50 - 1.50	media	
			0.75 - 1.25	media	
			1.25 - 1.75	alta	
			1.75 - 2.00	muy alta	

FIGURA 4

⁸ En un planteo muy simplista, si el director de obra tiene buenos conocimientos teóricos (le asignamos 1) pero es pésimo en experiencia (le asignamos 0), el promedio nos daría que es regular (0.5). Lo cual no es parámetro válido para atribuir el *valor* “regular” a la variable “capacidad del director de obra”; puede suceder que en determinadas obras donde tenga necesidad de apelar frecuentemente a los conocimientos teóricos y existan buenos ayudantes, tendrá desempeño mejor que regular probablemente, como puede suceder que aún con buenos ayudantes, y buena experiencia práctica, pero incapaz de resolver cuestiones por falta de conocimientos teóricos, resulte en el balance malo.

Esto genera al menos dos situaciones: a) Que la variable *velocidad* de desplazamiento puede adoptar valores “*nítidos*” comprendidos entre 0 m/s y 2 m/s, que bien podrían “medirse” con “*pseudo valores lingüísticos*”⁹, dependiendo de la escala que se adopte. Esta situación puede presentarse en gráficos cartesianos donde se registran en el eje de las abscisas los valores que puede tomar la variable (0 – 2) y en el eje de las ordenadas le asignamos los valores {*baja, alta*}; entre ellos, puede aparecer el valor *media*, y será entonces {*baja, media, alta*}; y si buscáramos aún mas precisión aparecerán otros valores lingüísticos {*muy baja, media baja, media alta, muy alta*}. La figura 4 muestra los tres casos y b) hemos consignado las expresiones lingüísticas dadas como *pseudo valores literales* entendiendo que si bien se expresan en términos literales, resultan categóricamente deterministas, en razón de que la pertinencia de los *valores* encuentra límites nítidos en valores numéricos de velocidad, de manera que se registra como *baja* una velocidad de 0,99 m/s y como *alta* 1,01 m/s (Figura 5), se presenta así una situación que parece no ofrecernos “*toda*” la información disponible y/o que podría ser necesaria en la construcción del dato.

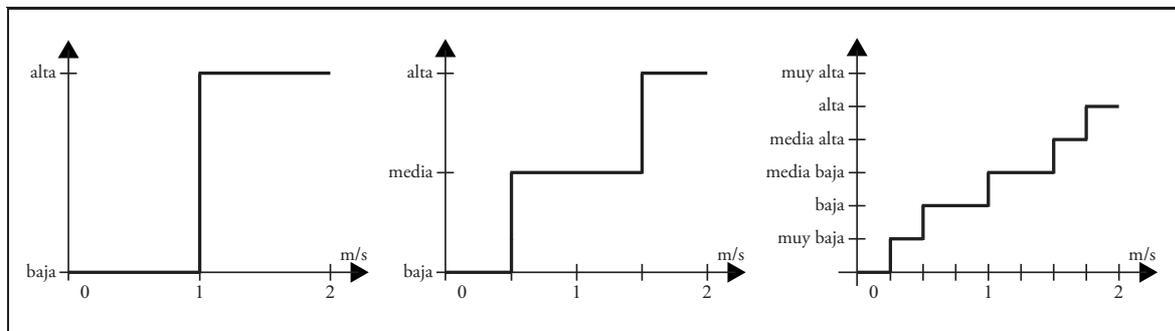


FIGURA 5

Se intuye que puede existir una “zona gris” donde atribuir *valor bajo* a la variable pareciera “no estar suficientemente justificada” porque si bien el valor 0,99 m/s pertenece al rango definido como *bajo*, *está “tan” próximo a valor alto*, que podemos preguntarnos ¿bajo, igualmente bajo que 0,10 m/s? Es acá donde podemos empezar a pensar en términos de subconjuntos borrosos.

Veamos la misma información: velocidad de movimiento del objeto expresada en m/s v.s. expresión en valores lingüísticos {*bajo, medio, alto*} expresados en la figura 6, donde se introduce en el eje de las ordenadas una “*variable*” μ , que toma valores en el intervalo [0, 1] y que significa “el grado de pertenencia” del *valor “numérico”* de la variable al *valor lingüístico*; muestran que:

- a) a la variable lingüística *baja* le corresponden las velocidades que van de 0 a 1, pero a las velocidades del intervalo [0, 0.5] les “corresponde” el valor lingüístico *bajo* “en mayor grado” que a las velocidades comprendidas en el intervalo [0.5, 1], a las que también les

⁹ Introduzco la expresión “*pseudo valores lingüísticos*” para referir que no cualquier valor expresado en letras resulta un valor lingüístico.

corresponde el valor lingüístico *baja*. Valor lingüístico *media* le corresponde al intervalo [0.5, 1.5], en diferentes “grados”, así, la figura 6 muestra que la velocidad 1.0 m/s es “mas media” ($\mu = 1$) que 1.4 m/s ($\mu < 1$).

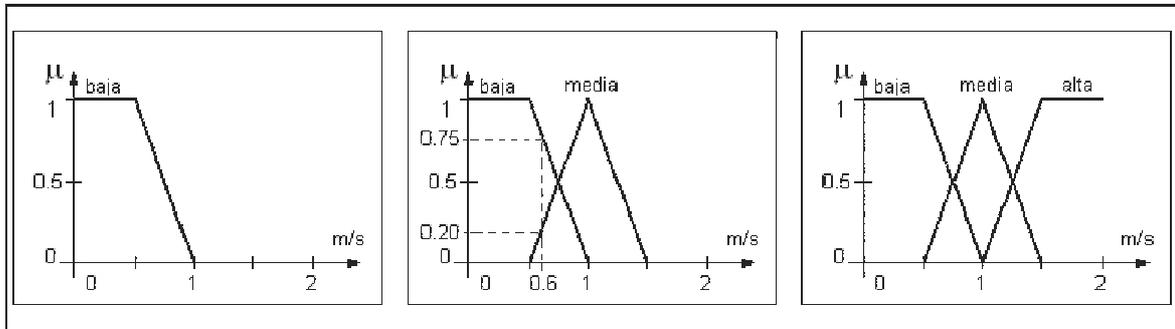


FIGURA 6

- b) a los valores de velocidad comprendidos en [0.5, 1] le corresponden los valores lingüísticos: *baja* y *media* pero en “diferentes grados”. Vea en la figura 3.b que a valor “nítido” de velocidad 0.6 m/s le corresponden $\mu = 0.75$ para valor lingüístico *baja* y $\mu = 0.20$ para valor lingüístico *media*.
- c) los valores nítidos de velocidad comprendidos en [1, 2] corresponden a valor lingüístico *alta*, la cual “comparte” con el valor lingüístico *media* el intervalo [1, 1.5].

¿Qué es μ ?. Informalmente asumimos que es un “auxiliar” que nos permite asignar un “sustento” de pertenencia de cada valor nítido al valor lingüístico que definimos, esto es así toda vez que con la Figura 6 decimos por ejemplo, que al valor lingüístico *alta* le corresponde con sustento 1 ($\mu = 1$) los valores nítidos de velocidad del intervalo [1.5, 2.0]; una velocidad de 1.25 m/s da un sustento $\mu = 0.5$ al valor lingüístico *alta* y al mismo tiempo $\mu = 0.5$ a valor lingüístico *media*, significa esto que el valor nítido de velocidad 1.25 m/s tiene “rasgos” de velocidad expresada en valores lingüísticas como alta y media.

¿Cómo es posible que se considere que un valor nítido aparezca en dos (o mas) valores lingüísticos de una variable a la vez?. En esto consiste la “novedad” de esta forma de “valorar” el dato, porque esa “medida” está indicándonos que un valor nítido de alguna dimensión, no tiene por qué pertenecer siempre “totalmente” una clase, ni tiene por qué ser “exclusiva” de esa clase. Por ejemplo, si fuera un médico el que encuentra que el valor nítido que arrojan los resultados de un análisis clínico se encuentra en el límite de lo normal y lo patológico, con razón concluye que tanto el estado de saludable como el de enfermo que diagnostique en el paciente puede tener algún “margen de error”; interpreta que si bien el valor está en el rango de lo aceptable, no es “tan aceptable” como si el resultado hubiera estado por encima de lo hallado; puede decirse que el médico suele adoptar una decisión *lógica difusa*, cuando no considera al paciente ni saludable ni enfermo, sino que “construye” su diagnóstico, por ejemplo, apelando a los valores que toman otras variables de la unidad de análisis para diagnosticar.

5. Definición de los valores lingüísticos

Los valores lingüísticos de la variable velocidad en nuestro ejemplo anterior pueden ser {*baja*, *media*, *alta*} y son expresados mediante una función de $\mu_{(x)}$; llamamos formalmente a cada μ en el valor lingüístico que representa de la variable lingüística V , *sustento* de la *calificación* x . Así un subconjunto borroso que expresa el valor lingüístico *baja* estará conformado por ejemplo por:

$$B = \int_U \frac{\mu_B(x)}{x}$$

asumiendo que los x pertenecen a un universo continuo, (el símbolo \int no debe entenderse como integral en el sentido clásico que le asigna el análisis matemático) y la expresión $\frac{\mu_B(x)}{x}$ debe leerse: “en el subconjunto borroso B a la calificación x le corresponde el sustento $\mu_B(x)$ ”

$$\text{donde } \mu_B(x) = \begin{cases} 1 & 0 \leq x \leq 0,5 \\ -2x + 2 & 0,5 < x \leq 1 \\ 0 & 1 < x \leq 2 \end{cases}$$

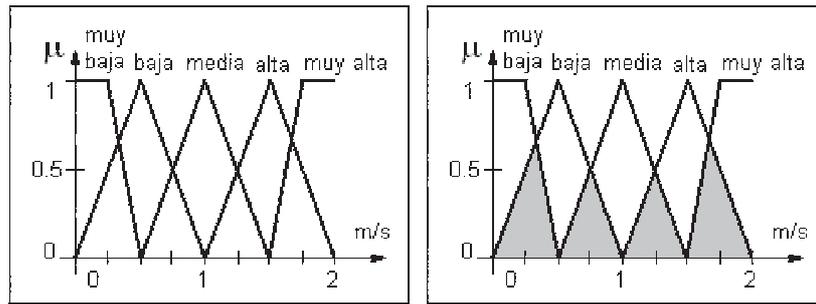
el subconjunto borroso que expresa el valor lingüístico *media*, se escribe: $M = \int_U \frac{\mu_M(x)}{x}$,

$$\text{donde } \mu_M(x) = \begin{cases} 0 & 0 \leq x \leq 0,5 \\ 2x - 1 & 0,5 < x \leq 1 \\ -2x + 3 & 1 < x \leq 1,5 \\ 0 & 1,5 < x \leq 2 \end{cases}$$

el subconjunto borroso que expresa el valor lingüístico *alto*, se escribe $A = \int_U \frac{\mu_A(x)}{x}$,

$$\text{donde } \mu_A(x) = \begin{cases} 0 & 0 \leq x \leq 1 \\ 2x - 2 & 1 < x \leq 1,5 \\ 1 & 1,5 < x \leq 2 \end{cases}$$

La ley de variación del sustento μ tomará la forma que el caso requiera, cada caso en particular; la forma adecuada será la que mejor se adapte al modelo (situación) de la realidad que queremos representar, y es precisamente en este punto, otro momento donde decimos que en la construcción del *sustento* μ es donde se manifiesta ese *indicador* que introduce el modelo cuatripartito de la estructura del dato. Puede ser el caso que necesitemos mayor precisión en los valores lingüísticos, entonces apelaremos a introducir mas subconjuntos borrosos, que definen los valores lingüísticos posibles, por ejemplo: {*muy baja*, *baja*, *media*, *alta*, *muy alta*}, situación que se representa en la figura 7. En la figura 8 las regiones sombreadas que indican la zona “borrosa” donde se superponen sustentos de calificaciones de valores lingüísticos diferentes.



\bar{B} = lo que no es baja.

Si μ_i es el sustento de las calificaciones x_i en B; el sustento de las calificaciones x_i en $\neg B$ es $\mu'_i = 1 - \mu$, entonces si:

$$B = \int_U \frac{\mu_B(x)}{x} \quad \text{con} \quad \mu_B(x) = \begin{cases} 1 & 0 \leq x \leq 0,5 \\ -2x + 2 & 0,5 < x \leq 1 \end{cases}$$

$$\bar{B} = \int_U \frac{\mu'_B(x)}{x} = \int_U \frac{(1 - \mu_B(x))}{x} \quad \text{con} \quad \mu'_B(x) = \mu_{\bar{B}}(x) = \begin{cases} 0 & 0 \leq x \leq 0,5 \\ 2x - 1 & 0,5 < x \leq 1 \\ 1 & 1 < x \leq 2 \end{cases}$$

Se hace así muy visible que hay una calificación x para la que tanto el subconjunto borroso *baja* como en el subconjunto borroso *no baja* tienen el mismo sustento μ (figura 9) y asimismo hay un intervalo de calificaciones x_i que tienen sustentos mayores que 0 tanto en el subconjunto borroso *baja* como en el subconjunto borroso *no baja*; esto suele ser interpretado como que se están rompiendo los principios de la lógica clásica de “no contradicción y del tercero excluido”. En álgebra de conjuntos el complemento del conjunto B (que se escribe \bar{B}) es el conjunto conformado por todos los elementos del universo que no pertenecen al conjunto B, formalmente:

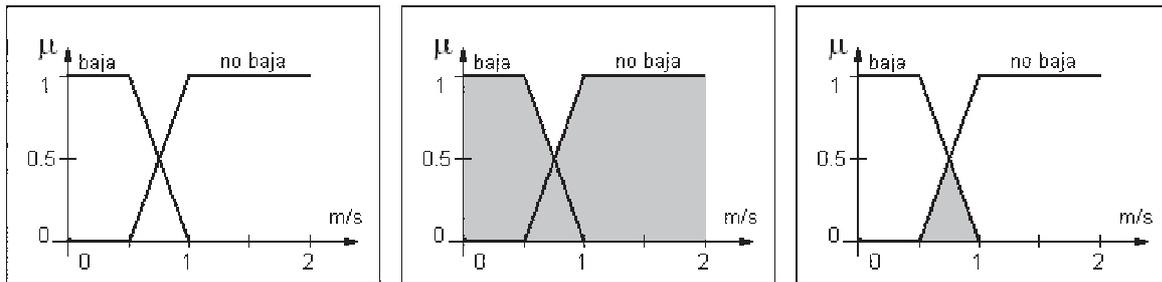
$$\bar{B} = \{ \forall x / x \in U \wedge x \notin B \}$$

de manera que en teoría de conjuntos clásica, nítidos – crisp, y lógica clásica, la unión de las partes (todas), es igual al todo (universo), lo que parece no verificarse en lógica difusa con los subconjuntos borrosos.

En efecto, se define la *unión* de valor *baja con no baja* como:

$$B \cup \bar{B} = \max [\mu_B(x), \mu'_B(x)]$$

se ve en el gráfico que la unión de B y su complemento no “cubre” el todo (figura 10).



$$B \cap \bar{B} = \min [\mu_B(x), \mu_{\bar{B}}(x)]$$

se ve en la figura 11 que la intersección de B y su complemento “parece” tener alguna región en común, lo que nos pone en duda el principio de no contradicción, porque la intersección de lo que es (*baja*) con lo que no es (*no baja*), no es vacía. Se advierte que “se rompen” así algunos principios de la lógica clásica, pero el tratamiento de esa cuestión no es relevante para el tema motivo de este artículo.

6. Subconjuntos borrosos para valores lingüísticos. El indicador.

Un *valor lingüístico* es aquel que se expresa en palabras del lenguaje natural; estas palabras en lógica difusa pueden ser caracterizadas como subconjuntos borrosos definidos en un *universo del discurso*.

Un valor lingüístico puede expresarse en el cuádruple $(N, U, T(N), M)$, donde:

N es el nombre (expresión coloquial) de la variable literal; por ejemplo, N es la velocidad de movimiento del objeto en el ejemplo que hemos tratado en la mayor parte de este trabajo.

U es el dominio actual en el que la variable lingüística N toma valores¹⁰ “cuantitativos”; en el ejemplo anterior $U = [0, 2]$ en m/s ¹¹.

$T(N)$ es el conjunto de valores lingüísticos que la variable N puede tomar; en nuestro ejemplo resulta $T(N) = \{ baja, media, alta \}$

M es la regla semántica que relaciona cada valor lingüístico en $T(N)$ con un subconjunto difuso en U ; en nuestro ejemplo M relaciona los *valores lingüísticos*: *baja*, *media* y *alta* con cada una de las funciones de pertenencia $\mu(x)$ definidas en U .

¹⁰ Pueden ser nítidos o no, en tal caso se establecerá una “cuantificación” que nos aporte la idea de pertenencia de la calificación al subconjunto borroso.

¹¹ Que al formularse en términos de calificaciones de un subconjunto borroso puede resultar útil que sea normalizado e inclusive transformado en “adimensional”.

De estas definiciones vemos que las variables lingüísticas pueden ser “extensiones” de las variables numéricas en el sentido de que permiten tomar subconjuntos borrosos como sus valores; así hoy podríamos estar caracterizando expresiones que pueden resultar altamente subjetivas como “alto, bello, largo” mediante subconjuntos borrosos, y con ellos podemos operar estableciendo reglas de inferencias difusas y relaciones, obteniendo resultados “aproximados”, tan aproximados como los resultados de los “razonamientos” que a diario hacemos los seres humanos, como por ejemplo:

Si al analizar *la calidad* de un examen surge que: a) las consignas son adecuadas respecto de los contenidos; b) las consignas solicitan cuestiones conceptuales y c) el tiempo disponible es adecuado; podemos decir que la calidad del examen es buena, pero: i) si las consignas son adecuadas respecto de los contenidos, y ii) las consignas solicitan cuestiones conceptuales, pero iii) el tiempo disponible no es el adecuado a ese temario, estaríamos frente a una situación donde calificar al examen de malo podría resultar temerario, pero en la que indudablemente la calidad del examen no puede ser calificada buena.

En el caso, la variable lingüística N es *calidad del examen*; pueden ser valores lingüísticos de esa variable, por ejemplo:

$$T(N) = \{ \text{no adecuado, adecuad, muy adecuado} \}$$

El dominio U será el Universo de *calificaciones* x , que se utiliza para definir cada valor lingüístico de la variable, por ejemplo puede ser el caso que hayamos establecido calificación de 0 a 5, donde los sustentos mayores para las menores calificaciones se condicen con los valores literales mas desfavorables y viceversa, así por ejemplo pueden quedar definidos los subconjuntos borrosos.

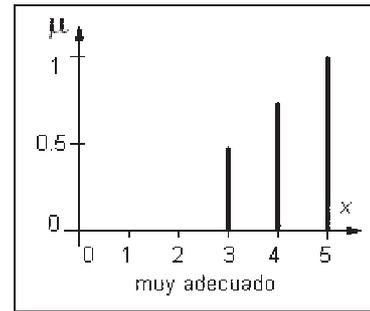
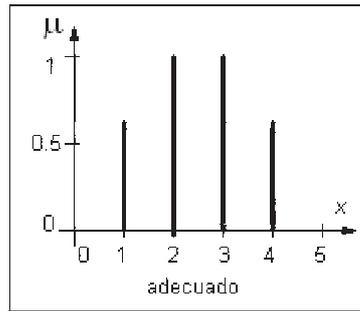
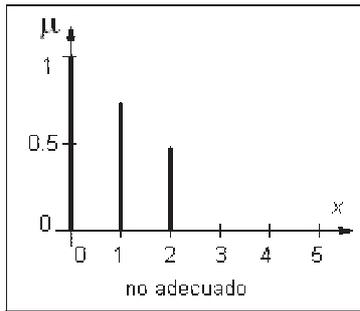
En estos casos $U = \{1, 2, 3, 4, 5\}$ hemos adoptado un universo discreto (que bien podría ser continuo). Por ejemplo, puede ser el caso que los valores literales estén representados en los subconjuntos borrosos $A = \text{no adecuado}$; $B = \text{adecuado}$ y $C = \text{muy adecuado}$, representados en las Figuras 12, 13 y 14.

Estos subconjuntos borrosos se expresan simbólicamente:

$$A = \frac{1}{0} + \frac{0,75}{1} + \frac{0,5}{2} \qquad B = \frac{0,6}{1} + \frac{1}{2} + \frac{1}{3} + \frac{0,6}{4} \qquad C = \frac{0,5}{3} + \frac{0,75}{4} + \frac{1}{5}$$

La regla M es la que asigna el *sustento* μ a cada calificación x del Universo definido para los valores de las variables, y es precisamente en la construcción de esta regla donde aparece *el indicador*, ese indicador del modelo cuatripartito de la matriz de datos que se propone a partir de la aplicación de algún procedimiento a una(s) dimensión(es).

Perspectivas Metodológicas



conjunto T que contiene otros valores posibles y se revela así una variable que caracterizan a una unidad de análisis, proporcionando toda la información de un dato.

7. Regla semántica. Indicador

Todo *indicador* está conformado por algún *procedimiento* simple o complejo que se ejecuta sobre alguna(s) *dimensión*(es) de la variable para averiguar cual valor de la variable corresponde a cada unidad de análisis.

Suele suceder que la dimensión se reconoce “fácilmente” en las sub-variables, éstas deben ser “tratadas” y normalmente se las convierte en calificaciones x para asignárseles sustentos μ ; en el caso anterior, los “valores nítidos” que se les puede asignar a cada una de las *dimensiones* son una escala numérica, mediante algún *procedimiento* hemos determinado que si el *valor lingüístico* es *no adecuado*, las calificaciones x bajas tienen sustento μ alto y las calificaciones altas tienen bajo sustento (eventualmente nulo); en correspondencia, si el *valor lingüístico* es *adecuado*, le corresponden sustentos altos a las calificaciones altas. En este análisis sucede la aplicación de *procedimientos aplicados a dimensiones*, los procedimientos no son únicos, mas bien es posible afirmar que cada problema requiere “su” procedimiento, y por otra parte se hace evidente la presencia de la(s) dimensión(es) de la variable. Bignoli Arturo muestra un cuadro indicativo que transcribimos en el caso de una unidad de análisis donde los valores nítidos de la variable analizada van de 10 a 7000, se establecen “rangos” de asignación de calificaciones y de valores lingüísticos a la escala de valores nítidos del dominio subyacente.

Calificación (x)	Valores nítidos	Valores lingüísticos	Valores lingüísticos	Valores lingüísticos	Valores lingüísticos
0	10	ínfimo	muy pequeño	pequeño	pequeño
1	30	muy muy pequeño			
2	60	muy pequeño	pequeño	pequeño	pequeño
3	120	pequeño			
4	240	bastante pequeño	medio	medio	pequeño
5	480	medio			
6	960	bastante grande	grande	grande	grande
7	1900	grande			
8	3000	muy grande	muy grande	grande	grande
9	5500	muy muy grande			
10	7000	enorme			

Podríamos haber segmentado los valores hallados en dos franjas: grande y pequeño; pero también podrían ser grande, mediano y pequeño; en la construcción de las posibles variantes para los valores que puede tomar la variable aparece la presencia de G que es una gramática con la cual se formarán los demás valores, a partir de unos pocos formados desde la asignación de sustento a las calificaciones. Aparece de esta manera “la quinta” componente “G” de la variable lingüística definida por Zadeh (1975) ($N, U, T(N), M, G$), donde:

N es el nombre de la variable literal.

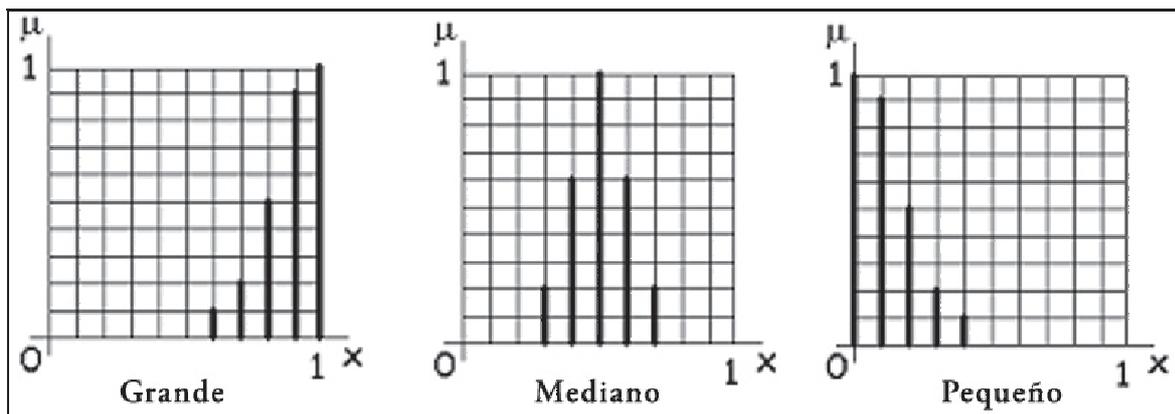
U es el dominio subyacente.

$T(N)$ es el conjunto de etiquetas que puede tomar N .

M es la regla semántica que relaciona cada elemento de $T(N)$ con un subconjunto difuso en U ; entre todos los posibles $M: T(N) \rightarrow F(U)$

G es una gramática para generar las etiquetas de $T(N)$

Ejemplo: Definidos tres valores mediante sus respectivos subconjuntos borrosos (etiquetas) Grande, Medio y Pequeño:



$$D_x(\text{Grande}) = 0.1/0.6 + 0.2/0.7 + 0.5/0.8 + 0.9/0.9 + 1/1$$

$$D_x(\text{Medio}) = 0.2/0.3 + 0.6/0.4 + 1/0.5 + 0.6/0.6 + 0.2/0.7$$

$$D_x(\text{Pequeño}) = 1/0 + 0.9/0.1 + 0.5/0.2 + 0.2/0.3 + 0.1/0.4$$

Es posible encontrar una gramática que nos permita generar otras etiquetas que expresan otros valores a partir de ellas. El procedimiento para componer etiquetas D'_x dado el subconjunto borroso

$$D_x = \mu_1 / x_1 + \mu_2 / x_2 + \dots + \mu_n / x_n$$

será aplicar la producción $D_x(\text{original}) = \sum \frac{\mu_i}{x_i} \rightarrow D'_x(\text{transf}) = \sum \frac{\mu_i^n}{x_i}$

donde será n el valor que “regula” la modificación de los valores de la etiqueta original.

Ejemplo: para $D_x(Grande) = 0.1/0.6 + 0.2/0.7 + 0.5/0.8 + 0.9/0.9 + 1/1$

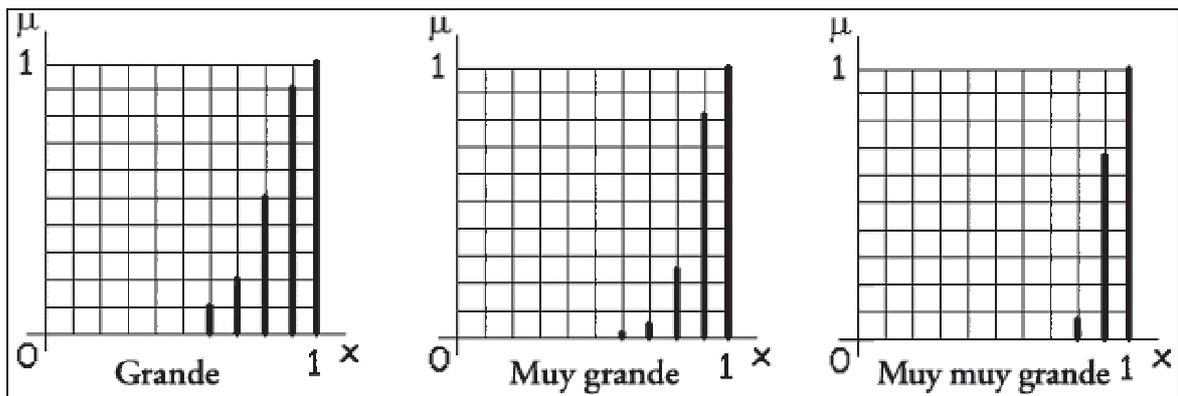
$$\text{será: } D_x(Grande) = \sum \frac{\mu_i}{x_i} \rightarrow D'_x(\text{muy grande}) = \sum \frac{\mu_i^2}{x_i}$$

entonces: $D_x(\text{muy grande}) = 0.1^2/0.6 + 0.2^2/0.7 + 0.5^2/0.8 + 0.9^2/0.9 + 1^2/1$
 que resulta: $D_x(\text{muy grande}) = 0.01/0.6 + 0.04/0.7 + 0.25/0.8 + 0.81/0.9 + 1/1$

Si buscamos una etiqueta que represente una situación aún mayor, tendremos:

$$D_x(Grande) = \sum \frac{\mu_i}{x_i} \rightarrow D'_x(\text{muy muy grande}) = \sum \frac{\mu_i^4}{x_i}$$

entonces: $D_x(\text{muy muy grande}) = 0.1^4/0.6 + 0.2^4/0.7 + 0.5^4/0.8 + 0.9^4/0.9 + 1^4/1$
 que resulta: $D_x(\text{muy muy grande}) = 0.06/0.8 + 0.66/0.9 + 1/1$



8. Conclusiones

Los subconjuntos borrosos, hasta aquí, no son mas que otra forma de “expresar situaciones” del mundo real, no sabemos que sucederá con ellos dentro de veinte siglos; los números irracionales fueron “descubiertos” por los pitagóricos en el siglo V a.C. y la comunidad científica “terminó” de incorporarlos como tales formalmente mucho tiempo después. La frecuencia y forma en que hemos “asimilado” la existencia y el uso de los números es probable que hayan provocado una “reconfiguración de significados”, que resulta útil, sin dudas, pero que también es posible que nos posicionen al margen de los cuestionamientos y de los nuevos descubrimientos que de ellos pueden derivarse.

Los subconjuntos borrosos nos permiten “expresar” (¿medir?) incertidumbres en marcos formales, donde la imprecisión aparece como consecuencia de la ausencia de criterios de pertenencia definidos con nitidez (conjuntos crisp), hasta ahora esto parecía reservado a la Teoría de la Probabilidad, pero sabido es que con ella es imposible escapar de la aleatoriedad de los sucesos, consecuentemente sus resultados son la probabilidad de ocurrencia, y solo eso (pero no menos); la Teoría de Subconjuntos borrosos trabaja en el marco de la *posibilidad*. ¿Será exagerado esperar que los subconjuntos borrosos “lleguen a ser números” en el próximo milenio?

Quedó expuesta una analogía entre la matriz de datos cuatripartita propuesta por Samaja (2005) y la lógica difusa tratada con subconjuntos borrosos; se “identifican” la variable de la matriz de datos, cuando acepta valores lingüísticos con las “variables lingüísticas” de la lógica difusa cuyos valores se expresan en las etiquetas del conjunto $T(N)$; los valores de ambos modelos pueden ser lingüísticos¹³ y el indicador de la matriz de datos se asocia a “la construcción” del subconjunto borroso, en el diseño de calificaciones se involucra (eventualmente) la dimensión (subvariable) y en la conformación del sustento se involucra algún procedimiento para asignar sustento μ_i a la calificación x_p , junto a la gramática (G) que se diseñe para proveer las diferentes etiquetas que expresarán los valores; son estas operaciones las que resultan cotidianamente más difíciles de visualizar, y en consecuencia el indicador de la matriz de datos (que siempre está), no siempre emerge con la nitidez suficiente para ser aceptado universalmente y sin discusiones.

Asumimos este aporte como una puerta que se abre para la exploración de la manera en que sucede el conocimiento y la valoración en la mente humana; nuestras *escalas de valores* suelen ser altamente subjetivas, aún en situaciones en las que la objetividad es deseable, como un examen o juicio. Parece ser que la formación del indicador puede tener muchas respuestas a estos interrogantes; habrá que indagar y ahondar en ellos y otra tarea será “cuantificarlos” con subconjuntos borrosos.

Bibliografía

- Bignoli, Arturo, *Nociones sobre el uso de magnitudes inciertas, variables literarias y conjuntos difusos*, Bs. As., Fondo de Ayuda a la educación, 1985.
- Meza, Araceli, *Tesis de Maestro en Ciencias: “Observadores Difusos y Control Adaptable Difuso Basado en Observadores”*, México, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, 2003, Director de tesis Dr. Wen Yu Liu.
- Pedrycz Witold & Gomide Fernando. *An Introduction to Fuzzy Sets: Analysis and Design*, Boston, Massachusetts Institute of Technology, 1998.
- Samaja, Juan “*Epistemología y Metodología. Elementos para una teoría de la investigación científica*”. Bs. As., Eudeba, 2005.

¹³ El autor de este trabajo inició el estudio de la identificación de ambos modelos desde esta analogía.

Zadeh, Lofti . “The Concept of a Linguistic Variable and its Application to Aproximate Reasoning”. *Information Sciences*, vol. 8, pag. 199-249 (part I), vol. 8 (part II), vol. 9 (part III), 1975;

“Fuzzy Sets”. *Information and Control*, vol. 8, www.lcc.uma.es/ppgg/FSS, LCC-ITI-2001-11, Dpto. de Lenguajes y Ciencias de la Computación de la Universidad de Málaga.