Introducción a la lógica difusa

Guillermo Morales-Luna
Centro de Investigación y Estudios Avanzados del IPN
(CINVESTAV-IPN)
gmorales@cs.cinvestav.mx

17 de febrero de 2002

Resumen

Hacemos una presentación elemental de la lógica difusa. Introducimos la noción de conjuntos difusos para luego presentar diversos cálculos proposicionales de tipo difuso.

Las lógicas difusas se han desarrollado rápidamente debido a sus potencialidades de aplicación, entre otras muchas áreas, en el diseño de controladores electrónicos. En este texto presentamos los conceptos básicos e invitaremos al lector a introducirse con mayor profundidad en estos temas a través de lecturas suplementarias.

1 Introducción

La *lógica difusa* ha cobrado una fama grande por la variedad de sus aplicaciones, las cuales van desde el control de complejos procesos industriales, hasta el diseño de dispositivos artificiales de deducción automática, pasando por la construcción de artefactos electrónicos de uso doméstico y de entretenimiento, así como también de sistemas de diagnóstico. De hecho, desde hace ya, al menos, década y media, la expedición de patentes industriales de mecanismos basados en la lógica difusa tiene un crecimiento sumamente rápido en todas las naciones industrializadas del orbe.

Se ha considerado de manera general que el concepto de lógica difusa apareció en 1965, en la Universidad de California en Berkeley, introducido por Lotfi A. Zadeh [7].

Las lógicas difusas, pues de hecho hay que hablar de ellas en plural, son esencialmente lógicas multivaluadas que extienden a las lógicas clásicas. Estas últimas imponen a sus enunciados únicamente valores *falso* o *verdadero*. Bien que éstas han modelado satisfactoriamente a una gran parte del razonamiento "natural", es cierto que el razonamiento humano utiliza valores de verdad que no necesariamente son "tan deterministas". Por ejemplo, al calificar que "el cielo es azul" uno está tentado a graduar qué tan "azul", en efecto, es el cielo, e igualmente, si "un vehículo se mueve rápido", también se está obligado a considerar qué tan rápido es el vehículo, aunque esto último no implique necesariamente cuantificar la velocidad del vehículo con toda precisión.

Las lógicas difusas procuran crear aproximaciones matemáticas en la resolución de ciertos tipos de problemas. Pretenden producir resultados exactos a partir de datos imprecisos, por lo cual son particularmente útiles en aplicaciones electrónicas o computacionales. El adjetivo "difuso" aplicado a ellas se debe a que los valores de verdad no-deterministas utilizados en ellas tienen, por lo general, una connotación de incertidumbre. Un vaso medio lleno, independientemente de que también esté medio vacío, no está lleno completamente ni está vacío completamente. Qué tan lleno puede estar es un elemento de incertidumbre, es decir, de difusidad, entendida esta última como una propiedad de indeterminismo. Ahora bien, los valores de verdad asumidos por enunciados aunque no son deterministas, no necesariamente son desconocidos. Por otra parte, desde un punto de vista optimista, lo difuso puede entenderse como la posibilidad de asignar más valores de verdad a los enunciados que los clásicos "falso" o "verdadero". Así pues, reiteramos, las lógicas difusas son tipos especiales de lógicas multivaluadas.

Las lógicas difusas han tenido aplicaciones de suma relevancia en el procesamiento electrónico de datos. En determinadas áreas de conocimiento, a sus enunciados se les asocia valores de verdad que son grados de veracidad o falsedad, mucho más amplios que los meros "verdadero" y "falso". En un sistema deductivo se distingue enunciados "de entrada" y enunciados "de salida". El objetivo de todo sistema manejador de una lógica difusa es describir los grados de los enunciados de salida en términos de los de entrada. Más aún, algunos sistemas son capaces de refinar los grados de veracidad de los enunciados de salida conforme se refinan los de los de entrada. Por estas propiedades es que ciertos sistemas de lógica difusa aparentan una labor de aprendizaje, y son excelentes mecanismos de control de procesos. Desde el punto de vista tecnológico, las lógicas difusas se encuadran en el área de la llamada Inteligencia Artificial y han dado origen a sistemas expertos de tipo difuso y a sistemas de control automático.

En esta presentación haremos énfasis en el carácter multivaluado de las lógicas difusas. Introduciremos primero la noción de *conjunto difuso*, y las operaciones usuales en ese tipo de conjuntos. Inmediatamente después, presentaremos ciertos tipos de cálculos proposicionales de tipo difuso y de cuantificación difusa.

2 Conjuntos difusos

De manera intuitiva se tiene el concepto de *conjunto* como una colección bien definida de elementos, en la que es posible determinar para un objeto cualquiera, en un universo dado, si acaso éste pertenece o no al conjunto. La decisión, naturalmente, es "*sí pertenece*" o bien "*no pertenece*".

Por ejemplo, consideremos como universo a la *población económicamente activa*[1] *en México*, es decir, al conjunto formado por las personas residentes en ese país con una edad entre 18 años (cumplidos) y 66 años (por cumplir). Consideremos un mes cualquiera, digamos, diciembre de 2000 (y no porque entonces hubiera habido un cambio, sino porque era ése el último mes del siglo XX). El conjunto de personas *empleadas* en México en ese mes, podríamos pensar, está bien determinado: una persona en nuestro universo que entonces hubiera vendido su fuerza de trabajo, bajo un contrato de empleo, a una empresa legalmente constituída, sin duda alguna era una persona *empleada*, y alguien que no tuvo salario alguno

[1] No aspiro a usar definiciones técnicas precisas de tipo económico, así que los colegas especialistas en esas disciplinas habrán de disculparme las licencias que aquí me tome.

en ese mes y no estuvo vinculado a ningún patrón bajo una relación contractual, pues no era *empleado*. El lector observará la sobresimplificación del criterio de pertenencia enunciado. En efecto, ni falta el funcionario de la Secretaría del Trabajo que dirá: "Todo ciudadano que haya trabajado al menos una hora en ese mes y por eso haya recibido un pago, es un empleado", y tampoco faltará quien diga: "¿Qué empleo? No hallé trabajo en todo el 2000 y sólo en su último mes, mi primo me empleó a destajo para envolver regalos en su tienda: Yo no soy ningún empleado".

La noción intuitiva de conjunto puede, así, ser muy estrecha. En un *conjunto difuso* a cada elemento del universo se le asocia un *grado de pertenencia*, que es un número entre 0 y 1, a ese conjunto. Un conjunto difuso es pues una correspondencia (o función) que a cada elemento del universo le asocia su grado de pertenencia. Enunciada así esta definición parece ser cíclica, mas no lo es: un conjunto difuso es una función cuyo dominio es el universo y cuyo contradominio es el intervalo $[0, 1]$. En tanto el grado de pertenencia sea más cercano a 1 tanto más estará el elemento en el conjunto y en tanto el grado de pertenencia sea más cercano a 0 tanto menos estará el elemento en el conjunto.

Por ejemplo, los siguientes son conjuntos difusos, dados como funciones g, que pueden abarcar el concepto de *empleado*:

De estadística optimista "Uno es empleado si trabaja al menos una hora, bajo pago, en un mes." Para cada persona x sea $t(x)$ el número de horas trabajadas bajo pago el mes en cuestión. Hagamos $g_E(x) = 1$ si $t(x) \geq 1$ y $g_E(x) = 0$ si $t(x) = 0$.

De porcentaje en tiempo "Uno es empleado en proporción al tiempo trabajado." Supongamos que el total de horas posibles a ser laboradas en un mes sea $40 \times 4 = 160$. Hagamos $g_{PT}(x)$ igual al valor mínimo que resulte de comparar 1 con la razón $t(x)/160$.

De porcentaje en ingreso "Uno es empleado en proporción con que pueda adquirir los bienes de consumo necesarios para su familia." Denotemos por $p(x)$ a la paga que recibe el ciudadano x por hora de su trabajo. Supongamos que la "canasta básica" la evalúa la Secretaría de Comercio en M pesos al mes, por persona, y que cada trabajador tiene en promedio 2 dependientes económicos, además de él mismo. El salario del trabajador ha de mantener a 3 personas. Hagamos $g_{PI}(x)$ igual al valor mínimo que resulte de comparar 1 con la razón $t(x)p(x)/(3M)$.

Ponderación de tiempo e ingreso "Uno es empleado cuando trabaje mucho aunque no coma o no tenga apuros económicos aunque no trabaje." Sean a y b dos coeficientes entre 0 y 1 tales que $a + b = 1$. Hagamos $g_{Pon}(x) = a \cdot g_{PT}(x) + b \cdot g_{PI}(x)$.

El grado de pertenencia g_D a un conjunto difuso D puede ser interpretado de diversas maneras, en contextos diferentes. Las siguientes son sólo algunas posibles interpretaciones:

Proporción en la que se posee un atributo Si consideramos que D es un atributo, entonces para cada objeto x, $100 \cdot g_D(x)$ es el "porcentaje" con el que x posee D.

Probabilidad Si consideramos que D es un evento probabilista (una *variable aleatoria*, según se dice en la Teoría de la Probabilidad, con valores en el conjunto de partes del universo), entonces para cada objeto x, $g_D(x)$ es la probabilidad de que x ocurra en el evento D, es decir, $g_D(x) = \text{Prob}(x \in D)$.

Medida de creencia Si consideramos que D es un atributo, entonces para cada objeto x, $g_D(x)$ es un grado con el que se cree que x posee el atributo D.

Por ejemplo, g_{PT}, definida anteriormente, es ciertamente una proporción del tiempo laborado. $g_{Pon} = a \cdot g_{PT} + b \cdot g_{PI}$ es una medida de creencia (y la selección de pesos a y b sesga el énfasis que se le quiera dar al tiempo laborado o al ingreso obtenido). Para ilustrar la connotación probabilista, consideremos el conjunto difuso $D = \{'\text{empleados felices}'\}$. Entonces, para cada x, $g_D(x)$ sería una probabilidad de que x sea feliz.

Un conjunto, en el sentido intuitivo, posee una función *característica*: En cada elemento, la característica vale 1 ("*sí*") si el elemento está en el conjunto y vale 0 ("*no*") en caso contrario. En consecuencia, todo conjunto intuitivo es en sí un conjunto difuso.

Recíprocamente, dado un conjunto difuso D con función de pertenencia g_D, se puede fijar un *umbral* z entre 0 y 1, inclusive, para formar el conjunto, en el sentido intuitivo, de elementos con grado de pertenencia al menos z: x está en D_z si y sólo si $g_D(x) \geq z$. Diremos que D_z es el *corte a altura* z de D. El corte a altura 0 es entonces todo el universo, en tanto que el corte a altura 1 consta de los elementos con valor de pertenencia 1 al conjunto.

Es bien sabido que los conjuntos intuitivos pueden combinarse mediante las operaciones, llamadas *booleanas*, de complemento, unión e intersección: El complemento de un conjunto está formado por los elementos del universo que no están en él, la unión de dos conjuntos la forman los elementos que están en uno o en otro conjunto y la intersección la conforman los elementos en ambos conjuntos. Si nos referimos a funciones características, se tiene que la característica del complemento posee el valor opuesto al de la característica del conjunto, la característica de la unión de dos conjuntos vale uno si al menos una de las características de los conjuntos vale uno, y la característica de la intersección vale uno si las características de ambos conjuntos valen uno.

Esto puede servir de motivación para definir operadores de composición de conjuntos difusos. De hecho, para cada una de las interpretaciones descritas arriba se puede introducir una colección particular de operadores. Veamos en cada caso operaciones de complemento, unión e intersección:

Proporción en la que se posee un atributo En este caso, los grados de pertenencia se interpretan como proporciones, por lo cual se definen las operaciones como sigue:

Complemento El complemento de un conjunto difuso D asigna a cada objeto x el grado "complementario": $g_{\overline{D}}(x) = 1 - g_D(x)$.

Intersección La intersección de dos conjuntos difusos D, E asocia el mínimo de los grados de pertenencia, es decir, para cada objeto x: $g_{D\cap E}(x) = \mathrm{Min}(g_D(x), g_E(x))$.

Unión De manera similar, la unión de dos conjuntos difusos D, E asocia el máximo de los grados de pertenencia, es decir, para cada objeto x: $g_{D\cup E}(x) = \mathrm{Max}(g_D(x), g_E(x))$.

Probabilidad Vistos los grados de pertenencia como probabilidades, se tiene:

Complemento La probabilidad del complemento de un conjunto difuso D es la probabilidad "complementaria": $g_{\overline{D}}(x) = 1 - g_D(x)$.

Intersección Ésta es la probabilidad de la ocurrencia simultánea de dos eventos. La intersección está muy ligada al concepto de *probabilidad condicional*. Si denotamos por $\text{Prob}(A|B)$ a la probabilidad de que "ocurra A dado que ha ocurrido B", entonces por un célebre resultado de la Teoría de la Probabilidad, llamado el *Teorema de Bayes*, ha de valer la identidad $\text{Prob}(A|B)\text{Prob}(B) = \text{Prob}(B|A)\text{Prob}(A)$. El valor común en esta igualdad es, precisamente, la probabilidad de la intersección $\text{Prob}(A \cap B)$.

Así pues, teniendo una función d que a dos eventos cualesquiera A, B les asocia una "densidad de probabilidad condicional" $d(A|B)$ tal que a cada objeto x le asocia un valor $d(A|B)(x)$ de manera que

$$d(A|B)(x) \cdot g_B(x) = d(B|A)(x) \cdot g_A(x) \tag{1}$$

entonces para dos conjuntos difusos cualesquiera D, E se puede definir el grado de pertenencia a la intersección como $g_{D \cap E}(x) = d(D|E)(x) \cdot g_E(x)$.

Una densidad de probabilidad condicional que satisfaga la ec. (1) podría definirse haciendo, por ejemplo, que para cualesquiera dos eventos distintos e "independientes" A y B: $d(A|B)(x) = g_A(x)$. Evidentemente, la noción de independencia dependerá del universo en cuestión.

Por ejemplo, si consideramos a la población económicamente activa, el conjunto A "de profesores que enseñan en instituciones de educación superior" y el conjunto B de "empleados con salarios altos" pueden ser considerados independientes pues ciertamente se dan los casos de profesores universitarios con bajos salarios, de profesores universitarios con altos salarios, de empleados con altos salarios que no son profesores y de empleados con bajos salarios que no son profesores. Así pues, para cualquier ciudadano x, $d(A|B)(x) = g_A(x)$ y $d(B|A)(x) = g_B(x)$, es decir, se cumple la relación (1).

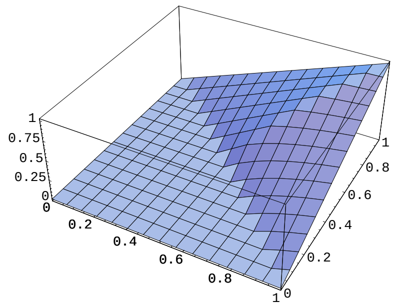
Vemos pues que para definir la operación de intersección, basta tener un operador de "probabilidad condicional". De manera recíproca, si se tiene definida de alguna manera al operador de intersección, entonces siguiendo el teorema de Bayes se puede definir un operador de "probabilidad condicional". Por tanto, las nociones de intersección (probabilista) de conjuntos difusos y la de probabilidad condicional son reducibles una a la otra.

Unión La probabilidad de que ocurra uno u otro evento es la probabilidad de uno, más la probabilidad del otro, menos la probabilidad de que ocurran ambos eventos: $g_{D \cup E}(x) = g_D(x) + g_E(x) - g_{D \cap E}(x)$.

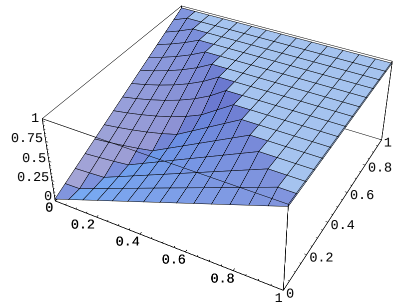
Medida de creencia Las siguientes funciones pueden parecer definidas de manera arbitraria, pero ciertamente tienen una motivación intuitiva:

Complemento Como en los casos anteriores, se hace: $g_{\overline{D}}(x) = 1 - g_D(x)$.

Intersección Dados dos conjuntos difusos A, B con sendos grados de pertenencia g_A y g_B, si para un punto dado x, la suma $g_A(x) + g_B(x)$ es menor que 1 entonces descartamos que ese punto sea común a ambos conjuntos, es decir, no debe estar "en la intersección". En otro caso, se toma como grado de pertenencia, a la



(a)



(b)

Figura 1: Grados de pertenencia, según el enfoque de "medida de creencia", de (a) intersección, $g_{A\cap B}(x)$, y (b) unión, $g_{A\cup B}(x)$, en términos de los grados de creencia $g_A(x)$ y $g_B(x)$.

intersección, a la razón de la diferencia $[g_A(x) + g_B(x)] - 1$ entre el máximo de $g_A(x)$ y $g_B(x)$. En símbolos

$$g_{A\cap B}(x) = \begin{cases} 0 & \text{si } g_A(x) + g_B(x) < 1 \\ \frac{g_A(x)+g_B(x)-1}{\max(g_A(x),g_B(x))} & \text{si } g_A(x) + g_B(x) \geq 1 \end{cases}$$

Unión Dados dos conjuntos difusos A, B con sendos grados de pertenencia g_A y g_B, si para un punto dado x, la suma $g_A(x) + g_B(x)$ es mayor que 1 entonces convenimos en que ese punto está "en la unión". En otro caso, se toma como grado de pertenencia, a la unión, al máximo de las razones $g_A(x)/(1 - g_B(x))$ y $g_B(x)/(1 - g_A(x))$. En símbolos

$$g_{A\cup B}(x) = \begin{cases} 1 & \text{si } g_A(x) + g_B(x) \geq 1 \\ \max\left(\frac{g_A(x)}{(1-g_B(x))}, \frac{g_B(x)}{(1-g_A(x))}\right) & \text{si } g_A(x) + g_B(x) < 1 \end{cases}$$

En la figura 1 mostramos las gráficas correspondientes a estos operadores de interseccón y de unión.

A partir de operaciones de complemento, unión e intersección, se obtienen conjuntos difusos "más complejos" como resultado de aplicar sucesivamente estos operadores partiendo de una colección de conjuntos difusos, digamos, "primitivos". Para hablar con un poco de más precisión: Si $A_1, \ldots, A_n$ son conjuntos difusos primitivos, para operadores de complemento, de intersección y de unión fijos, la clase de *conjuntos definibles*, partiendo de los conjuntos primitivos, son los que se obtienen mediante las reglas siguientes:

1. Todo conjunto primitivo es definible.

2. El complemento de todo definible es, a su vez, definible.

3. La intersección y la unión de dos conjuntos definibles, son, a su vez, definibles, también.

Así, por ejemplo, si A_1, A_2, A_3 son tres conjuntos difusos primitivos, los siguientes son meros ejemplos de conjuntos difusos definibles a partir de ellos:

$$\overline{A_2}$$
$$\overline{A_1 \cap \overline{A_2}} \cup A_3$$
$$(\overline{A_1} \cup A_2 \cup A_3) \cap (A_1 \cup \overline{A_2} \cup A_3) \cap (A_1 \cup A_2 \cup \overline{A_3})$$

Denotemos a un conjunto definible como $F(A_1, \ldots, A_n)$, sólo para enfatizar el hecho de que se obtiene de los conjuntos A_i. Cada tal conjunto tiene asociada una función $g_{F(A_1,\ldots,A_n)}$ que a cada objeto x del universo le asocia un grado de pertenencia $g_{F(A_1,\ldots,A_n)}(x)$ al conjunto definible, la cual, naturalmente, se escribe como una composición de los grados de pertenencia de los conjuntos primitivos.

Existen dos problemas fundamentales en cualquier teoría de conjuntos difusos:

Problema 2.1 (de deducción) *Para un conjunto definible $F(A_1, \ldots, A_n)$ y un objeto dado x, si se sabe que cada grado de pertenencia $g_{A_i}(x)$ cae en un intervalo $[a_i, b_i]$, entonces se ha de estimar en qué intervalo I ha de caer el grado de pertenencia $g_{F(A_1,\ldots,A_n)}(x)$.*

Utilizando una jerga técnica actual, podemos decir que éste es un problema de tipo "*hacia adelante*": conociendo los valores iniciales $g_{A_i}(x)$, mediante las funciones de los conectivos de complemento, unión e intersección, se calcula consecutivamente los grados de pertenencia de los conjuntos involucrados hasta obtener el valor $g_{F(A_1,\ldots,A_n)}(x)$.

En etapas de aplicación, un resolvedor de este problema se ve como un agente que realiza "pronósticos": "Si los valores iniciales son de tales características, los finales han de ser de tales cuales".

Problema 2.2 (de inferencia) *Para un conjunto definible $F(A_1, \ldots, A_n)$ y un objeto dado x, si se sabe que el grado de pertenencia $g_{F(A_1,\ldots,A_n)}(x)$ cae en un intervalo I, y que para algunos conjuntos primitivos $A_k, A_{k+1}, \ldots, A_n$ sus correspondientes grados de pertenencia $g_{A_i}(x)$ caen en intervalos $[a_i, b_i]$, $i = k, \ldots, n$, entonces se ha de estimar para los otros índices $i = 1, \ldots, k-1$ en qué intervalos $[a_i, b_i]$ debieron caer los correspondientes grados de pertenencia $g_{A_i}(x)$.*

También en jerga técnica, podemos decir que éste es un problema de tipo "*hacia atrás*": conociendo los valores finales $g_{F(A_1,\ldots,A_n)}(x)$ y algunos iniciales $g_{A_i}(x)$, teniendo en cuenta las funciones de los conectivos de complemento, unión e intersección, se busca determinar los valores que debieron asumir los demás grados de pertenencia iniciales para obtener el valor final.

En etapas de aplicación, un resolvedor de este problema se ve como un agente que realiza "diagnósticos": "Si los valores observados (finales) son de tales características en unas ciertas condiciones (iniciales), entonces las demás variables iniciales han de haber cumplido con tales hipótesis".

Vemos pues que los conjuntos difusos involucran de manera esencial procedimientos de cálculo numérico o simbólico. Al contrario de una primera idea sugerida por su nombre, veremos que la lógica difusa es un área de cálculo preciso. Ambos problemas, de deducción y de inferencia, pueden ser resueltos, efectiva y eficientemente, analizando los tipos de las funciones matemáticas involucradas en los grados de pertenencia y en los conectivos lógicos.

3 Cálculo proposicional difuso

Recordamos que Gottfried Wilhelm Leibniz[2] planteó en el s. XVII la necesidad de un *calculus ratiocinator* que sería un sistema con un propio formalismo que permitiera la manipulación simbólica de enunciados, en concordancia con las leyes de la lógica, para descubrir nuevas verdades o bien para verificar como verdaderas a tesis postuladas como tales. Por lo menos desde entonces se reconoció la relevancia del manejo (procesamiento) simbólico de enunciados lógicos. Naturalmente, George Boole[3] y Gottlob Frege[4] contribuyeron de manera notabilísima en el origen mismo del razonamiento automático. Mas, desde la perspectiva de las lógicas difusas, en los antecedentes de ellas es necesario mencionar a Jan Łukasiewicz[5]. En su sistema trivaluado, si a los valores de verdad *Falso*, *Desconocido* y *Verdadero* se les representa, respectivamente, por los valores numéricos 0, 1/2 y 1, entonces las nociones de *complemento* $n(x) = 1 - x$, *conjunción* $c(x, y) = \min(x, y)$ y *disyunción* $d(x, y) = \max(x, y)$, corresponden a la interpretación intuitiva de esos conectivos lógicos (invitamos al lector a que escriba las tablas de verdad de los conectivos y a que se cerciore de esta aseveración). Este sistema trivaluado entraña, ciertamente, una noción de difusidad.

En un cálculo proposicional difuso se tiene inicialmente una colección de proposiciones primitivas, o "atómicas", una serie de conectivos lógicos y reglas definidas de "buena formación" de proposiciones "compuestas" a partir de las atómicas. Cada proposición puede asumir un valor de verdad que puede ser *Falso*, o *Verdadero* o alguno otro "entre" esos dos valores extremos de verdad. Cada conectivo lógico tiene asociada una función que determina el valor de verdad de la proposición resultante de él en términos de los valores de verdad de las proposiciones que componen a esa proposición resultante. Resultan entonces sendos problemas "de deducción" (dado que los valores de las proposiciones atómicas quedan caracterizados, se ha de caracterizar los de proposiciones compuestas partiendo de ellas) y "de inferencia" (dado que los valores de las proposiciones compuestas han sido observados, se ha de formular hipótesis, y además probarlas, respecto a los valores de las proposiciones atómicas involucradas que debieron dar origen a los valores observados). En lo que sigue, detallaremos esta construcción de un cálculo proposicional difuso.

Comencemos con conjuntos de valores de verdad a los cuales llamaremos de *valuaciones*. Un conjunto de valuación $\mathcal{V}$ puede ser finito o bien puede ser un contínuo.

Como ejemplos de valuaciones finitas consideremos primero un enfoque de creencias, que

[2](1646-1716) Filósofo, matemático y asesor político alemán, inventor del cálculo diferencial e integral (simultáneamente con Isaac Newton, aunque de manera independiente), con grandes aportaciones a la lógica y a la metafísica.

[3](1815-1864) Matemático inglés, inventor de la, así llamada en honor suyo, *álgebra booleana.* En 1847 publicó su *Mathematical Analysis of Logic* y en 1854 la célebre *An Investigation into the Laws of Thought, on Which Are Founded the Mathematical Theories of Logic and Probabilities*, bases de su álgebra.

[4](1848-1925) Lógico y matemático alemán, fundador de la lógica matemática. Su trabajo se desarrolló entre la matemática y la filosofía, de hecho se dice que él decía que "todo buen matemático es medio filósofo y todo buen filósofo es medio matemático". Sus convicciones políticas eran muy particulares, vistas éstas desde la perspectiva actual.

[5](1878-1956) Lógico y filósofo polaco, enseñó en la Universidad de Leopolis (Lwów) (1906-1915), en la Universidad de Varsovia (1915-1939) y fue profesor en la Academia de Ciencias de Irlanda, en Dublín, (1945-1956). Se ocupó de problemas de determinismo en lógica y en filosofía, y en los fundamentos de la teoría de la probabilidad. Fue iniciador de la lógica multivaluada: en 1920 presentó un sistema trivaluado, en 1922 uno con una infinidad de valores de verdad y en 1953 su sistema tetravaluado para una lógica modal.

a cada proposición atómica le asocia una *etiqueta lingüística* tal como:

Falso *Casi_falso* *Tal_vez_falso* *Desconocido* *Tal_vez_cierto* *Casi_cierto* *Verdadero*

o bien, con un enfoque de "posesión de atributos" del tipo "p es A", las etiquetas lingüísticas pueden ser

Definitivamente_no *Más_bien_no* *Al_parecer_no* *No_se_sabe*
Al_parecer_sí *Más_bien_sí* *Definitivamente_sí*

De este segundo tipo son las mediciones de variables de control cuando éstas varían de manera discreta, es decir, cuando los aparatos de medición detectan tan solo cambios de una determinada magnitud. El vendedor de telas en la mercería de la esquina mide longitudes con un metro marcado hasta centímetros. Sus "etiquetas lingüísticas" serían

$$.00 \quad .01 \quad .02 \quad \cdots \quad .5 \quad \cdots \quad .98 \quad .99 \quad 1.$$

Si $\mathcal{V}$ es un conjunto de valuación finito, entonces al enumerar a sus elementos, digamos $\mathcal{V} = \{v_0, \ldots, v_{n-1}\}$, se puede suponer que esa enumeración concuerda con el orden interno de $\mathcal{V}$, es decir, si $i < j$ entonces la etiqueta lingüística v_i es menor, en el orden de $\mathcal{V}$, que v_j. En tal caso, $\mathcal{V}$ puede identificarse con un subconjunto de números entre 0 y 1, incluyendo éstos, de manera que el primer elemento v_0 corresponda a 0 y el último, v_{n-1}, corresponda a 1 (esto se logra si a cada v_i se le asocia el número $i/(n-1)$ lo cual daría una distribución uniforme de las etiquetas en el intervalo $[0, 1]$ pero ciertamente ésta no es la única manera de insertar a $\mathcal{V}$ en $[0, 1]$, la manera en la que se inserte dependerá de la aplicación, evidentemente).

Como ejemplos de valuaciones contínuas tenemos los que resultan cuando se estima "probabilidades de ocurrencia". Si a una proposición p se le asocia la probabilidad de que ocurra (en cierto espacio de eventos) entonces su valor de verdad puede ser un número real entre 0 y 1 inclusive. O bien, cuando se observa una parámetro p en algún proceso, el cual varía de manera contínua y puede asumir valores entre uno mínimo a y otro máximo b, entonces el propio intervalo $[a, b]$ constituye un espacio de valuación contínuo. Utilizando una sencilla *regla de tres*, es decir, asociándole a cada número x entre a y b, el número $y = (x-a)/(b-a)$ se puede identificar al conjunto de valuación $\mathcal{V} = [a, b]$ con el intervalo $[0, 1]$.

Así pues, sin ninguna pérdida de generalidad, podremos suponer siempre que el conjunto de valuación $\mathcal{V}$ es un subconjunto del intervalo $[0, 1]$ y contiene a los extremos 0 y 1. Supondremos también que es simétrico respecto al punto medio $1/2$, es decir, supondremos que si para un número x se tiene que x está en $\mathcal{V}$ entonces $1-x$ está también en $\mathcal{V}$.

Para construir un cálculo proposicional difuso, elijamos un conjunto $\mathcal{P}_0$ de proposiciones *atómicas*. Una *asignación* es una correspondencia que a cada átomo p en $\mathcal{P}_0$ le asocia bien un valor $v(p)$ en $\mathcal{V}$ o bien lo deja sin ningún valor asociado.

Supongamos ahora que se tiene tres conectivos lógicos: *complemento* $\neg$, *conjunción* $\wedge$ y *disyunción* $\vee$, cada una con una respectiva operación de evaluación, digamos $f_\neg, f_\cap, f_\cup$. Entonces la colección de *proposiciones booleanas compuestas* $\mathcal{P}_B$ se define como sigue: Todo átomo p es un elemento de $\mathcal{P}_B$; el complemento de una proposición en $\mathcal{P}_B$ está también en $\mathcal{P}_B$; y la conjunción y la disyunción de dos proposiciones en $\mathcal{P}_B$ están también en $\mathcal{P}_B$.

$$\boxed{\begin{array}{rcl} f_\neg(x) & = & 1-x \\ f_\cap(x,y) & = & \min(x,y) \\ f_\cup(x,y) & = & \max(x,y) \end{array}}$$

(a) Proporcionalidad

$$\boxed{\begin{array}{rcl} f_\neg(x) & = & 1-x \\ f_\cap(x,y) & = & \begin{cases} 0 & \text{si } x+y<1 \\ \frac{x+y-1}{\max(x,y)} & \text{si } x+y\geq 1 \end{cases} \\ f_\cup(x,y) & = & \begin{cases} 1 & \text{si } x+y\geq 1 \\ \max\left(\frac{x}{(1-y)},\frac{y}{(1-x)}\right) & \text{si } x+y<1 \end{cases} \end{array}}$$

(b) Creencia

$$\boxed{\begin{array}{rcl} f_\neg(x) & = & 1-x \\ f_\cap(x,y) & = & x\cdot y \\ f_\cup(x,y) & = & x+y-xy \end{array}}$$

(c) Probabilístico

$$\boxed{\begin{array}{rcl} f_\neg(x) & = & 1-x \\ f_\cap(x,y) & = & \max(x+y-1,0) \\ f_\cup(x,y) & = & \min(x+y,1) \end{array}}$$

(d) Łukasiewicz

Tabla 1: Funciones de evaluación para conectivos de complemento, conjunción y disyunción.

Toda asignación v, definida sobre los átomos se extiende naturalmente a una asignación v^* definida en toda la colección $\mathcal{P}_B$. A saber: Si p es un átomo, entonces $v^*(p) := v(p)$. Si para una proposición p en $\mathcal{P}_B$ se tiene $x = v^*(p)$ entonces $v^*(\neg p) := f_\neg(v^*(p))$. Similarmente, si para dos proposiciones p, q en $\mathcal{P}_B$ se tiene $x = v^*(p)$, $y = v^*(q)$ entonces

$$v^*(p\cap q) := f_\cap(v^*(p), v^*(q)) \;\text{ y }\; v^*(p\cup q) := f_\cup(v^*(p), v^*(q)).$$

En jerga técnica, se dice que la asignación v^* está *propagando incertidumbres* partiendo de la asignación v.

Ya hemos visto ejemplos de funciones de evaluación en la sección anterior. En la tabla 1 presentamos un resumen de ellas en el contexto actual y además presentamos una nueva colección de funciones de evaluación debidas también a Łukasiewicz.

Hasta ahora hemos sólo considerado tres tipos de conectivos lógicos: complemento, conjunción y disyunción. Sin embargo, los demás quedan prácticamente determinados pues siempre se puede definir a la *implicación* haciendo $(p \to q) \equiv \neg p \vee q$ y a la *equivalencia lógica* haciendo, por ejemplo, $(p \leftrightarrow q) \equiv (p\wedge q)\vee(\neg p \wedge \neg q)$. Recíprocamente, si se tuviese definidas únicamente funciones de evaluación para el complemento y para una implicación "$\to$" entonces se podría definir a los demás conectivos haciendo: $(p\vee q) \equiv (\neg p) \to q$, $(p\wedge q) \equiv \neg(p \to (\neg q))$, y $(p \leftrightarrow q) \equiv (p\to q)\wedge(q\to p)$. En varias lógicas difusas se procede según esto último.

Así pues, cada posible selección de funciones de evaluación de conectivos da origen a un distinto cálculo proposicional difuso. Lo que tienen en común todos ellos es que cuando se les restringe a considerar sólo valores de verdad deterministas $0 =$ *Falso* y $1 =$ *Verdadero* entonces coinciden con el cálculo proposicional clásico. No se puede afirmar que ninguno de los presentados aquí es el correcto. Dependiendo de algunas aplicaciones acaso uno de ellos es el más adecuado, mas dejará de serlo, cuando se cambie de aplicaciones.

4 Cálculo de predicados difuso

De igual manera a como se procede en el cálculo de predicados clásico, se supone dada una *signatura*, consistente de una colección de *constantes*, una colección de *relaciones* y una colección de *funciones*. Cada relación R y cada función F tiene asociada una *aridad*, es decir, un número de argumentos. Los *términos* se forman considerando a las constantes y a las variables como términos y a la composición de símbolos de funciones con términos, como términos también. Las *fórmulas atómicas* se obtienen como composición de símbolos de relaciones con términos. Las *fórmulas* se obtienen considerando como tales a las fórmulas atómicas, a la composición de fórmulas con conectivos lógicos y a las cuantificaciones *universales* y *existenciales* de fórmulas (véase los detalles de construcción en el escrito de José Alfredo Amor en este mismo libro).

También, al igual que en el caso clásico, una *interpretación* consiste de un universo M, de una correspondencia de cada constante c en la signatura con un elemento m_c en M y de una correspondencia de símbolos de funciones con funciones en M: Si f es un símbolo de función de aridad n, entonces m_f es una función con dominio M^n y contradominio M, es decir, $m_f : M^n \to M$. De esta manera, a cada término t que no involucre variables, le corresponderá un elemento m_t en M.

Para completar la noción de interpretación, a cada símbolo de relación R, digamos que de aridad n, se le asocia un conjunto difuso m_R en el universo M^n; y a los conectivos lógicos se les asocia funciones específicas de evaluación.

Pues bien, una asignación v asocia a cada variable x un elemento en M (escribiremos $v \equiv_x u$ para denotar el hecho de que v y u coinciden en todas las variables, excepto, quizá, en x). A una fórmula atómica *cerrada*, es decir sin variables, $R(t_1, \ldots, t_n)$ la asignación le asocia como valor de verdad el grado de pertenencia de la n-ada $(m_{t_1}, \ldots, m_{t_n})$ al conjunto difuso m_R. En símbolos: $v(R(t_1, \ldots, t_n)) = g_{m_R}(m_{t_1}, \ldots, m_{t_n})$. Si $\Box$ es un conectivo lógico, con función de evaluación $f_\Box$ entonces para dos fórmulas ϕ, ψ se define $v(\phi \Box \psi) = f_\Box(v(\phi), v(\psi))$. Finalmente, para fórmulas cuantificadas se define:

$$
\begin{aligned}
v(\forall x\, \phi(x)) &= \min\{u(\phi) | u \equiv_x v\} \\
v(\exists x\, \phi(x)) &= \max\{u(\phi) | u \equiv_x v\}
\end{aligned}
$$

Así pues, ya sea en un cálculo difuso de proposiciones o en uno de predicados, se puede plantear los siguientes dos problemas:

Problema 4.1 (de deducción o de "pronóstico") *Para una fórmula ϕ, si se sabe que las fórmulas atómicas que involucra toman valores de verdad en ciertos intervalos, entonces se ha de estimar en qué intervalo I ha de caer el valor de verdad de ϕ.*

Problema 4.2 (de inferencia o de "diagnóstico") *Para una fórmula ϕ, si se sabe que el valor de verdad $v(\phi)$ cae en un intervalo I, y que para algunos átomos involucrados en ϕ sus correspondientes valores de verdad caen en ciertos intervalos, entonces se ha de estimar para los átomos restantes en qué intervalos debieron caer sus correspondientes valores de verdad.*

5 Problema fundamental de la lógica difusa

Se puede identificar como problema fundamental el de inferencia enunciado anteriormente. En efecto, un *programa de lógica difusa* corresponde de una lista de parejas $[(\phi_i, v(\phi_i))]_i$ consistente, cada una, de una fórmula y de un valor de verdad asociado. En estas condiciones se ha de resolver el problema de inferencia: Determinar los valores de verdad que debieron haber asumido los átomos involucrados para obtener las condiciones establecidas en el programa lógico. En jerga técnica, decimos que se ha de *satisfacer la consulta (query satisfaction)* planteada por el programa lógico.

Existen diversos procedimientos para localizar una solución a este problema, y tanto en el artículo de Mauricio Osorio como en el conjunto de José Alfredo Amor y Raymundo Morado se presentan técnicas, considerando lógicas deterministas. Propiamente las técnicas pueden repetirse aquí, mas en este caso, se ha de tener especial cuidado en el manejo de valores de incertidumbre. No abundamos más en este tema y remitimos al lector a la literatura especializada.

Lecturas recomendadas

[1] Dubois, D., Prade, H. Fuzzy sets in approximate reasoning II (Logical approaches), *Fuzzy sets and systems.*, **40**, pp. 203-244, 1991.

[2] Dubois, D., Prade, H. *Fuzzy sets and systems: Theory and applications*, Academic Press, 1980.

[3] Hájek, P., Godó Ll. *Deductive systems of fuzzy logic*, unpublished manuscript, 1997.

[4] Kantrowitz, M. *et al*, *FAQ: Fuzzy Logic and Fuzzy Expert Systems*, disponible en `ftp.cs.cmu.edu:/user/ai/pubs/faqs/fuzzy/fuzzy.faq`, (desde 1995).

[5] Kaufmann, A., *Introducción a la teoría de los subconjuntos borrosos*, Cía. Editorial Continental, 1982.

[6] Zadeh, L. Fuzzy sets, *Information & Control.*, **8**, 1965.

[7] Zadeh, L. Fuzzy logic, *IEEE Computer*, **1:83**, 1988.