

**ESCOLA TÈCNICA SUPERIOR D'ENGINYERIA DE
TELECOMUNICACIÓ DE BARCELONA (UPC)**

Departament de Teoria del Senyal i Comunicacions

**SISTEMAS DIFUSOS DINÁMICOS
PARA EL
TRATAMIENTO DE INFORMACIÓN
TEMPORAL IMPRECISA**

Autor: Orestes Mas i Casals
Director: Joan Maria Miró Sans

Barcelona, 1997

2. Aspectos de teoría de conjuntos difusos

2.1 Introducción

En el contenido de este capítulo se revisan los fundamentos de la teoría de conjuntos difusos, recogiendo con especial énfasis aquellos aspectos de mayor interés en ingeniería, y muy en particular los que se utilizarán en el diseño e implementación de los sistemas temporales difusos, objetivo de esta tesis. Merecerán atención especial el concepto de relación difusa, la composición de relaciones y el principio de extensión. También se aprovechará este capítulo para fijar la notación a emplear a lo largo del texto.

Para destacar las conexiones existentes entre la teoría de conjuntos difusos y la matemática clásica, los conceptos se definen y desarrollan primeramente en el marco de la teoría clásica de conjuntos, para posteriormente extenderlos al dominio difuso. De esta forma, la teoría de conjuntos difusos se presenta como una generalización de la teoría clásica que, a su vez, es un caso particular de aquella.

Por último, cabe insistir en que no se ha pretendido un tratamiento exhaustivo de estas amplias disciplinas, que superaría con creces el ámbito de un capítulo. El objetivo perseguido en este caso es enfatizar aquellos aspectos de mayor interés para la tesis, así como presentar lo difuso como una generalización bien fundamentada de lo clásico.

2.2 Terminología

A continuación se relacionan las traducciones castellanas de términos en inglés que se han adoptado en este texto:

- ◆ **Fuzzy**→**Difuso o Borroso**: El uso de estos vocablos tiene su origen en la acepción fotográfica del término *fuzzy*, que designa una imagen poco definida o desenfocada. El uso de *difuso* o *borroso* será indistinto en el presente texto, si bien personalmente el autor prefiere el primero de los dos términos, puesto que está libre de las connotaciones “negativas” del segundo.
- ◆ **Crisp**→**Nítido**: En la literatura sobre conjuntos difusos se usa el término *crisp* como antónimo de *fuzzy*. En esta tesis hemos optado por traducirlo como *nítido*, puesto que es coherente con el símil

fotográfico utilizado en la traducción del vocablo *fuzzy*. En ocasiones usamos también la palabra *clásico*.

- ◆ **Fuzzification**→**Borrosificación**: Muchos autores españoles traducen de esta forma el término *fuzzification*, que designa a la operación que permite convertir un conjunto nítido en borroso. Sin embargo, la relativa complejidad del vocablo -junto con la de su antónimo, desborrosificación- hace que a menudo se sustituya por el anglicismo “fuzzificación”.
- ◆ **Defuzzification**→**Desborrosificación**: Con frecuencia también se utiliza el anglicismo “defuzzificación”, por las mismas razones que se exponen en el párrafo anterior.

Además de las traducciones que acaban de señalarse, se relacionan a continuación un conjunto de términos que forman parte del vocabulario habitual sobre conjuntos difusos:

- ◆ **t-norma y t-conorma**: El uso de estos vocablos, derivados respectivamente de *norma triangular* y *conorma triangular*, fue introducido en 1942 por Menger, en un estudio sobre espacios métricos estadísticos [Klir 1995]. Su uso se ha generalizado hoy en día en los textos sobre teoría de conjuntos difusos. Designan, respectivamente, cualquier operación válida de **intersección** o **unión**, respectivamente.
- ◆ **Universo del discurso**: Se trata de un útil artificio, debido a Augustus De Morgan [Prior 1976] -y adoptado generalmente desde entonces-, para designar una colección cualquiera de elementos con los que se trabaja en un contexto dado, en el que estos elementos poseen las mismas características. Se denota generalmente al universo con la letra X , y a sus elementos individuales como x . Sobre este concepto general se definen los conjuntos, tanto nítidos como difusos, que se tratan en las secciones siguientes.

2.3 Conjuntos

2.3.1 Conjuntos clásicos: Definiciones básicas.

Conjunto nítido.

G.Cantor, el introductor de la teoría de conjuntos, definió el concepto de conjunto como “el agrupamiento en un todo de objetos bien definidos y distintos de nuestra percepción o de nuestro pensamiento, los cuales se denominan elementos del conjunto”. Siguiendo esta definición, llamaremos *conjunto clásico* o *nítido* a una colección cualquiera de elementos de un universo dado. Los conjuntos así definidos dicotomizan los elementos del universo en dos categorías: miembros (aquellos que pertenecen al conjunto) y no miembros. Como regla general, escribiremos los nombres de los conjuntos con la letra inicial en mayúsculas, y los de sus elementos en minúsculas y cursiva. Se utiliza el símbolo “ \in ” para indicar la pertenencia de un elemento a un conjunto, de forma que la expresión $x \in A$ indica que el elemento x pertenece al conjunto A , mientras que $x \notin A$ significa que no pertenece a él.

Tiene interés destacar que, con el concepto de conjunto así definido, el universo del discurso es a su vez un conjunto nítido, denominado *conjunto universal*, que agrupa a todos los elementos concernientes a una aplicación o contexto dado. Por el contrario, el conjunto que no contiene elementos se designa como *conjunto vacío*, denotado por \emptyset .

Dado un cierto universo X , existen básicamente tres métodos para definir conjuntos en él:

1. Mediante la **enumeración** de todos los miembros del conjunto, encerrados entre llaves (método de la lista). Por razones obvias, este método sólo puede usarse para conjuntos *finitos*. Así el conjunto A , cuyos miembros son a_1, a_2, \dots, a_n , se escribe como

$$A = \{a_1, a_2, \dots, a_n\}$$

2. El segundo método consiste en definir un conjunto A mediante una **propiedad** que deben satisfacer todos sus miembros. Se trata del

denominado método de la regla, para el que se utiliza usualmente la notación

$$A = \{x \mid P(x)\}$$

en la que el símbolo \mid denota la frase “tal que”, y $P(x)$ designa una proposición de la forma “ x tiene la propiedad P ”. De esta forma el conjunto A queda descrito como el conjunto de aquellos elementos $x \in X$ para los cuales la proposición $P(x)$ es verdadera.

3. Por último, un tercer método se basa en definir un conjunto A mediante una **función**, llamada *función característica*, $\chi_A(x)$, que establece qué elementos de X son miembros de A y cuáles no. Dicha función es una aplicación⁵ de X al conjunto $\{0,1\}$. Devuelve el valor 1 si el elemento x pertenece a A , y 0 si no pertenece

$$\chi_A(x) : X \rightarrow \{0,1\}, \quad \chi_A(x) = \begin{cases} 1 & \text{para } x \in A \\ 0 & \text{para } x \notin A \end{cases}$$

Por el hecho de que la función característica permite distinguir si un elemento es o no miembro de un conjunto, recibe también el nombre de *función de discriminación*. Como se verá con detalle en la Sección 2.3.2, la generalización de esta función abre la puerta a la definición de conjuntos difusos. Por este motivo es el método que suele utilizarse con mayor frecuencia para describir conjuntos nítidos en los textos de lógica difusa.

Cardinalidad.

En sentido amplio, se define la cardinalidad de un conjunto como aquella cualidad que tiene en común con otros conjuntos con los que es posible establecer una correspondencia uno-a-uno entre sus elementos. En la práctica, para un conjunto finito A la cardinalidad coincide con el número de elementos que contiene, y suele denotarse con $\#A$ o $|A|$.

⁵ El concepto de aplicación se describe en la sección 2.6.

Convexidad.

Un conjunto universal importante es el formado por todos los puntos del espacio Euclídeo n -dimensional, \mathbb{R}^n . A los conjuntos definidos sobre \mathbb{R}^n interesa exigirles la propiedad de convexidad. Se dice que un conjunto A en \mathbb{R}^n es *convexo* si, para cada par de puntos $r, s \in A$ se cumple que los puntos t definidos como

$$t = \lambda r + (1-\lambda)s, \lambda \in [0,1]$$

pertenecen también a A . En otras palabras, el conjunto A en \mathbb{R}^n es convexo si para cada par de puntos r y s de A se cumple que todos los puntos intermedios de la recta que conecta r y s pertenecen también a A .

Subconjunto. Conjunto Potencia.

Para dos conjuntos cualesquiera A y B definidos sobre un mismo universo X , si cada miembro del conjunto B lo es también del conjunto A decimos que B es *subconjunto de A* o, alternativamente, que B está contenido en A escribiendo $B \subseteq A$. Un *subconjunto* será pues una colección de elementos de un conjunto. Conjuntos y subconjuntos son términos que se usan a menudo de forma sinónima, puesto que todo conjunto es subconjunto de sí mismo y a su vez subconjunto del universo X .

Un conjunto cuyos elementos son a su vez conjuntos suele denominarse *familia de conjuntos*. La familia de todos los posibles subconjuntos de un conjunto A recibe el nombre de *conjunto potencia* de A , y se denota usualmente por $\wp(A)$. La familia de todos los subconjuntos de $\wp(A)$ se llama *conjunto potencia de segundo orden* de A , escribiéndose $\wp(\wp(A)) = \wp^2(A)$. De forma similar pueden definirse conjuntos potencia de orden superior, aunque no serán de interés en la presente tesis. Resulta obvio que el conjunto potencia del universo finito X , $\wp(X)$, agrupa a todos los conjuntos posibles que podemos definir en ese universo y que por tanto su cardinalidad es $|\wp(A)| = 2^{|A|}$.

El concepto de conjunto potencia tiene su utilidad en la definición de funciones cuyos argumentos sean conjuntos, en lugar de elementos simples. En este caso la función se define sobre $\wp(X)$ en vez de hacerlo sobre X . En la sección 2.6.2 se expone este concepto con más detalle.

2.3.2 Conjuntos difusos: Definiciones básicas.

Conjunto difuso.

Al igual que para los conjuntos nítidos, un conjunto difuso se define como una colección de elementos escogidos entre los de un cierto universo X . La diferencia fundamental con los conjuntos nítidos es que en los difusos los distintos elementos de X pueden pertenecer **parcialmente** al conjunto. Por convenio, el concepto de pertenencia parcial se expresa mediante un número real del intervalo $[0,1]$, al que se denomina *grado de pertenencia* y que está asociado indisolublemente a cada elemento del conjunto difuso.

Al igual que sucedía en los conjuntos nítidos, a los difusos se les asigna un nombre con la letra inicial en mayúsculas, y a los elementos un nombre en minúsculas escrita en cursiva. Algunos autores simbolizan el carácter difuso de un conjunto mediante una tilde debajo de su nombre. Sin embargo, en este trabajo se ha prescindido de tal convenio, ya que complica innecesariamente la notación y al no existir, en general, ambigüedad sobre el tipo de conjunto del que se trata.

De forma similar al caso nítido, dado un cierto universo X , existen distintos métodos para definir conjuntos difusos en él. Estos son:

1. Se define el conjunto A mediante una función, llamada *función de pertenencia*, $\mu_A(x)$, que devuelve el grado de pertenencia del elemento $x \in X$ al conjunto A . La función $\mu_A(x)$ es una aplicación de X en el intervalo $[0,1]$

$$\mu_A(x) : X \rightarrow [0,1].$$

2. Dado que $\mu_A(x)$ es una función, a menudo resulta útil describir el conjunto difuso a través de su representación gráfica. Esto puede dar lugar a gráficas continuas o discretas, según sea el universo del discurso. Un ejemplo de este tipo de representación se da en la Figura 2.1

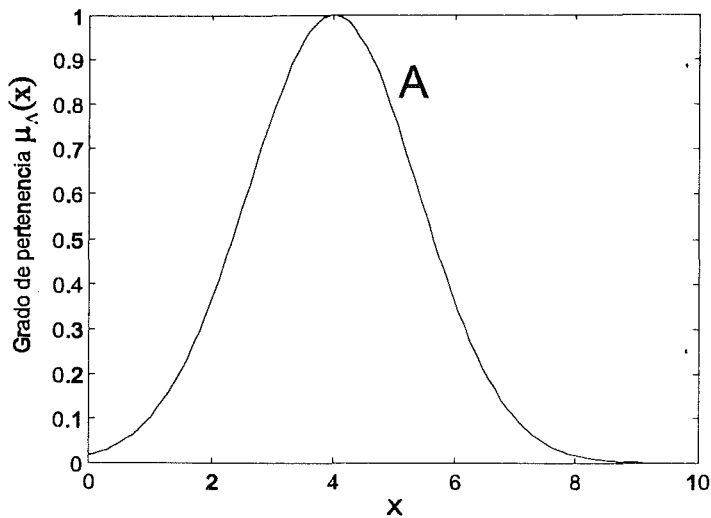


Figura 2.1: Representación gráfica de la función de pertenencia de un conjunto difuso⁶

3. Otra forma de definir conjuntos difusos se basa en la enumeración de todos sus miembros junto con sus grados de pertenencia obtenidos a partir de la función de pertenencia (método de la lista). Cada par elemento-grado se escribe en forma de fracción, con el nombre del elemento en el denominador y su grado de pertenencia en el numerador, uniéndose simbólicamente todas las fracciones con el símbolo +. Ni éste símbolo ni el de fracción implican la realización de suma o división alguna. Este método fue el utilizado por Zadeh en su primer artículo sobre conjuntos difusos [Zadeh 1965], y se usa para conjuntos *finitos*. El conjunto difuso A, cuyos miembros son a_1, a_2, \dots, a_n , con grados de pertenencia $\mu_A(a_1), \mu_A(a_2), \dots, \mu_A(a_n)$, se escribe como

$$A = \left\{ \frac{\mu_A(x_1)}{x_1} + \frac{\mu_A(x_2)}{x_2} + \dots + \frac{\mu_A(x_n)}{x_n} \right\} = \left\{ \sum_{x_i \in X} \frac{\mu_A(x_i)}{x_i} \right\} \quad (2.1)$$

En caso de que no exista ambigüedad en cuanto al número, posición y valor de los elementos del conjunto difuso, éstos pueden omitirse

⁶ El valor máximo no tiene necesariamente que ser 1.

en la ecuación (2.1) y escribir el conjunto en forma de un *vector* cuyas componentes sean los grados de pertenencia $\mu_A(x_i)$:

$$A = (\mu_A(x_1), \mu_A(x_2), \dots, \mu_A(x_n)) \quad (2.2)$$

Una variante del método para representar conjuntos difusos con infinitos elementos utiliza el símbolo de la integral en lugar del sumatorio sin que, al igual que en el caso anterior, ello implique la realización de integral alguna. En ese caso escribiremos

$$A = \int_X \frac{\mu_A(x)}{x} \quad (2.3)$$

Resulta evidente que los tres métodos descritos para definir conjuntos difusos descansan de una forma u otra en el concepto de función de pertenencia. Esto es así porque es una forma natural de expresar el grado de pertenencia de un elemento a un conjunto, especialmente para universos continuos. Además, así definida la función de pertenencia resulta ser una generalización de la función característica de los conjuntos nítidos, y su uso ofrece una vía conveniente para generalizar al caso difuso las distintas propiedades de los conjuntos nítidos. Por consiguiente, en lo sucesivo utilizaremos siempre esta función para manejar conjuntos difusos e incluso para los no difusos.

Cardinalidad

Dado que en un conjunto difuso los elementos no pertenecen a él plenamente, no tiene sentido definir el concepto de cardinalidad de la misma forma que en el caso nítido, es decir, como el número de sus elementos. Sin embargo, parece lógico extender el concepto de cardinalidad utilizando la noción de grado de pertenencia, y definir así la *cardinalidad escalar* de un conjunto difuso A como la acumulación de todos los grados de pertenencia de los elementos al conjunto

$$|A| = \sum_{x \in X} \mu_A(x) \quad (2.4)$$

definición que es compatible con la dada para el caso nítido, ya que al ser en ellos el grado de pertenencia 0 o 1, $|A|$ es igual al número de elementos de A. El concepto de

cardinalidad escalar de un conjunto difuso aparecerá posteriormente en las expresiones del grado de verdad de sentencias temporales difusas, tal y como se explica con todo detalle en el capítulo siguiente.

Convexidad.

La propiedad de convexidad para conjuntos difusos es también una generalización de la homónima para conjuntos nítidos. Al igual que aquella, la estableceremos para conjuntos difusos definidos en \mathbb{R}^n , que es donde tiene su principal aplicación. Diremos que un conjunto difuso A definido en \mathbb{R}^n es convexo si y sólo si para cualquier par de puntos n -dimensionales \mathbf{r} y \mathbf{s} que pertenecen al conjunto A con grados $\mu_A(\mathbf{r})$ y $\mu_A(\mathbf{s})$, respectivamente, el punto \mathbf{t} definido como

$$\mathbf{t} = \lambda\mathbf{r} + (1-\lambda)\mathbf{s}, \lambda \in [0,1] \quad (2.5)$$

pertenece al conjunto difuso A con un grado superior o igual al menor de entre $\mu_A(\mathbf{r})$ y $\mu_A(\mathbf{s})$. Matemáticamente lo expresaremos con

$$\mu_A(\mathbf{t}) = \mu_A(\lambda\mathbf{r} + (1-\lambda)\mathbf{s}) \geq \min(\mu_A(\mathbf{r}), \mu_A(\mathbf{s})) \quad (2.6)$$

para todo $\mathbf{r}, \mathbf{s} \in \mathbb{R}^n$ y siendo $\lambda \in [0,1]$.

En otras palabras, el conjunto A en \mathbb{R}^n es convexo si para todo par de puntos \mathbf{r} y \mathbf{s} de A se cumple que todos los puntos ubicados en la recta que conecta \mathbf{r} y \mathbf{s} pertenecen también en A con un grado superior o igual al menor de entre $\mu_A(\mathbf{r})$ y $\mu_A(\mathbf{s})$. Para espacios uni- o bidimensionales, la representación gráfica de las funciones de pertenencia correspondientes a conjuntos difusos convexos da lugar a curvas o superficies con un único máximo (unimodales) o, a lo sumo, con los máximos ubicados en puntos contiguos del espacio. La Figura 2.2 muestra un ejemplo de conjunto difuso convexo y otro no convexo. Cierta tipo de conjuntos difusos convexos servirá posteriormente de base para definir el concepto de *número difuso*, usado regularmente en la literatura y en la presente tesis.

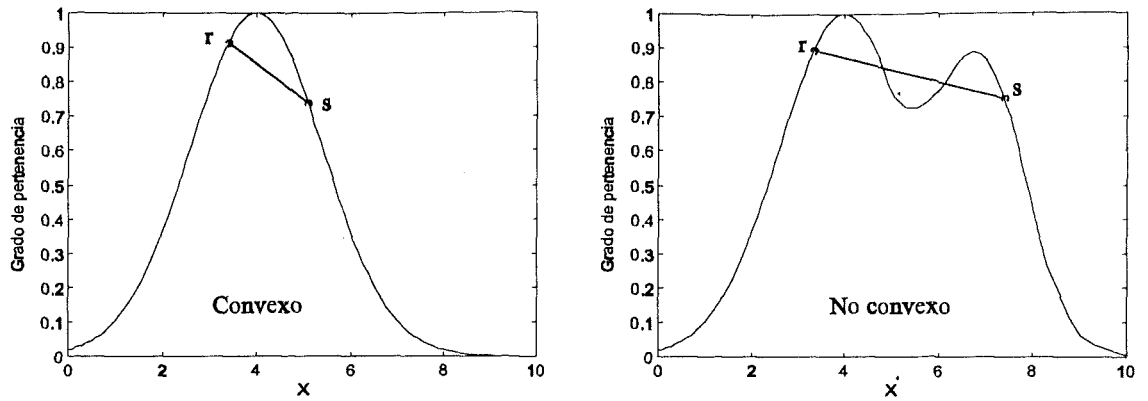


Figura 2.2: Conjuntos difusos convexos y no convexos

Subconjunto difuso. Conjunto potencia difuso. Grado de subconjunto.

El concepto de subconjunto definido en el ámbito de conjuntos nítidos se extiende para conjuntos difusos de la siguiente forma: Sean dos conjuntos difusos A y B definidos sobre el mismo universo X . Diremos que B es subconjunto de A o, alternativamente, que B está contenido en A (escrito $B \subseteq A$) si todos los elementos de B pertenecen también a A con un grado igual o superior, es decir

$$B \subseteq A \Leftrightarrow \mu_B(x) \leq \mu_A(x) \quad \forall x \in X \quad (2.7)$$

Esta definición se justifica por el hecho de que no es lógico que un elemento pertenezca al subconjunto con grado mayor al del propio conjunto. Una vez se hayan definido en la sección 2.4.2 las operaciones estándar de unión e intersección en el dominio de los conjuntos difusos, se podrá demostrar que, de la misma forma que para el caso nítido si B es subconjunto de A se cumplen las relaciones $A \cap B = B$ y $A \cup B = A$.

En otro orden de cosas, se define el *conjunto potencia difuso* de X como la familia de todos los posibles subconjuntos difusos del universo X , denotándose usualmente por $\mathcal{F}(X)$. De forma similar al caso nítido, la familia de todos los subconjuntos de $\mathcal{F}(X)$ se llama *conjunto potencia difuso de segundo orden* de X , escribiéndose $\mathcal{F}(\mathcal{F}(X)) = \mathcal{F}^2(X)$, y así sucesivamente. El concepto de conjunto potencia difuso se utilizará en la sección 2.6.2, cuando se discuta el principio de extensión. Allí, los conjuntos potencia difusa de orden n participan en la definición de funciones que toman como argumentos n conjuntos difusos.

Es de destacar que la definición de subconjunto dada por (2.7) es nítida, es decir, que la condición de subconjunto se da completamente o no se da en absoluto, en función de que la desigualdad de (2.7) la cumplan **todos** los elementos o no. Obviamente, si tal desigualdad la cumplen *casi* todos los elementos excepto unos pocos, puede considerarse que el conjunto B es *casi* un subconjunto del conjunto A, lo que conduce a definir un *grado de subconjunto* que refleje adecuadamente esta situación. Dicho grado se define de la siguiente forma: Dado que $B \subseteq A$ equivale a decir $\mu_B(x) \leq \mu_A(x)$, valores positivos de la expresión $\mu_B(x) - \mu_A(x)$ indican el *grado de violación de la inclusión* para el elemento x . Tomando sólo esos valores positivos y acumulando para todo $x \in X$ se tiene un índice de violación de la inclusión global, dado por

$$\sum_{x \in X} \max(0, \mu_B(x) - \mu_A(x)) \quad (2.8)$$

Si se sustrae este valor de la cardinalidad de B, se tiene un índice global de no-violación de la inclusión, pudiéndose normalizar el valor obtenido al intervalo $[0,1]$ dividiendo por la misma cardinalidad. De esta forma, el *grado de inclusión* del conjunto B en A, denotado por $S(B,A)$ se define como

$$S(B,A) = \frac{1}{|B|} \left(|B| - \sum_{x \in X} \max\{0, \mu_B(x) - \mu_A(x)\} \right) \quad (2.9)$$

De esta ultima expresión se deduce que si para todos los elementos del universo se cumple $\mu_B(x) \leq \mu_A(x)$, el grado de inclusión de B en A es igual a la unidad, resultado coherente con la primera definición que se dió de subconjunto difuso. Mediante ciertas manipulaciones algebraicas puede convertirse la expresión (2.9) en otra más compacta y de mayor interés. En efecto, dado que

$$\max\{0, \mu_B(x) - \mu_A(x)\} = \begin{cases} 0 & \text{si } \mu_B(x) \leq \mu_A(x) \\ \mu_B(x) - \mu_A(x) & \text{si } \mu_B(x) > \mu_A(x) \end{cases} \quad (2.10)$$

entonces

$$|B| - \sum_{x \in X} \max\{0, \mu_B(x) - \mu_A(x)\} =$$

$$\begin{aligned}
&= \begin{cases} \sum_{x_i \in X} \mu_B(x_i) & \forall x_i \mid \mu_B(x_i) \leq \mu_A(x_i) \\ \sum_{x_j \in X} \mu_B(x_j) - \sum_{x_j \in X} (\mu_B(x_j) - \mu_A(x_j)) & \forall x_j \mid \mu_B(x_j) > \mu_A(x_j) \end{cases} \\
&= \begin{cases} \sum_{x_i \in X} \mu_B(x_i) & \forall x_i \mid \mu_B(x_i) \leq \mu_A(x_i) \\ \sum_{x_j \in X} \mu_A(x_j) & \forall x_j \mid \mu_B(x_j) > \mu_A(x_j) \end{cases} \\
&= \sum_{x \in X} \min\{\mu_B(x), \mu_A(x)\} \tag{2.11}
\end{aligned}$$

con lo que la expresión (2.9) queda de la forma

$$\begin{aligned}
S(B, A) &= \frac{1}{|B|} \left(\sum_{x \in X} \min\{\mu_B(x), \mu_A(x)\} \right) \\
&= \frac{\sum_{x \in X} \min\{\mu_B(x), \mu_A(x)\}}{\sum_{x \in X} \mu_B(x)} \tag{2.12} \\
&= \frac{|B \cap A|}{|B|}
\end{aligned}$$

en donde la expresión $B \cap A$ representa la intersección difusa de los conjuntos A y B, efectuada mediante la función *mínimo*, según se expone en la sección 2.4.2. La expresión final obtenida (2.12), será de gran importancia en los desarrollos teóricos del capítulo 4.

α -Cortes. Normalidad.

Los α -cortes son conjuntos nítidos obtenidos a partir de conjuntos difusos siguiendo ciertas reglas. Matemáticamente, un α -corte definido sobre un cierto conjunto difuso A -denotado con ${}^\alpha A$, $\alpha \in [0, 1]$ - es el conjunto nítido formado por todos aquellos elementos del universo que pertenecen a A con un grado de pertenencia mayor o igual a α . Existe también el concepto de α -corte *fuerte*, ligeramente diferente del anterior, que se define como el conjunto nítido que contiene todos los elementos x del universo que

pertenecen a A con un grado *estrictamente mayor* a α . El α -corte fuerte se denota con ${}^{\alpha+}A$. En términos formales escribimos

$${}^{\alpha}A = \{x \mid \mu_A(x) \geq \alpha\} \quad (2.13)$$

$${}^{\alpha+}A = \{x \mid \mu_A(x) > \alpha\} \quad (2.14)$$

Resulta inmediato ver que dado un conjunto difuso cualquiera, pueden definirse a partir de él infinitos α -cortes, con sólo variar α entre 0 y 1. En la Figura 2.3 puede apreciarse un conjunto difuso cualquiera y su 0.6-corte.

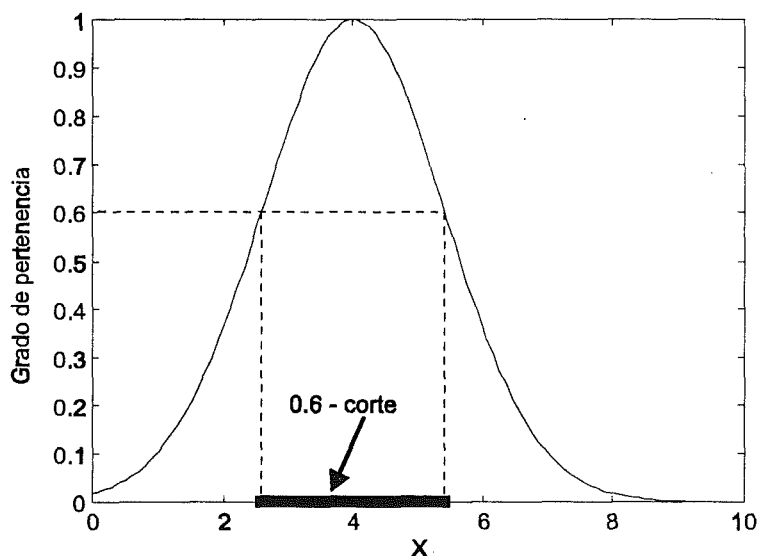


Figura 2.3: Conjunto difuso y su α -corte por 0.6.

La mayor utilidad de los α -cortes es la de permitir definir un conjunto difuso a partir de sus α -cortes de 0 a 1 mediante la llamada *propiedad de descomposición*⁷, que aplica a cualquier conjunto difuso, independientemente de que esté definido sobre un universo finito o infinito. Esta propiedad permite reconstruir el conjunto difuso a partir de la unión de todos sus α -cortes, asignando a cada elemento de un α -corte dado un grado de pertenencia igual a α :

⁷Para un tratamiento más profundo de la propiedad de descomposición, así como de otras propiedades de los α -cortes, véase, por ejemplo, [Klir 1995] pp. 35 y siguientes.

$$A = \bigcup_{\alpha \in [0,1]} \{\alpha \cdot A\} \quad (2.15)$$

expresión en la que $\alpha \cdot A$ es el conjunto difuso obtenido al asignar un grado de pertenencia α a todos los elementos del α -corte de A , y el símbolo de unión se refiere a la unión difusa estándar, definida en 2.4.2.

Por otra parte, podemos basarnos en el concepto de α -corte para definir otros conceptos útiles en teoría de conjuntos difusos, algunos de los cuales se enuncian a continuación:

- ◆ Se denomina *soporte* del conjunto difuso A definido en X al conjunto nítido que contiene todos los elementos que pertenecen a A en algún grado distinto de cero, o sea, al conjunto ^{0+}A .
- ◆ El *núcleo* del conjunto difuso A definido en X es el conjunto nítido que contiene todos los elementos que pertenecen totalmente a A , o sea, al conjunto 1A . Tanto el núcleo como el soporte pueden ser conjuntos vacíos.
- ◆ La *altura* $h(A)$ de un conjunto difuso A se define como el mayor grado de pertenencia al conjunto A de cualquier elemento del soporte. Puede ser menor que 1, lo que indica que ningún elemento pertenece totalmente a A .

El concepto de *normalidad* juega un papel importante en la definición de números difusos. Un conjunto difuso A recibe el nombre de *normal* si su núcleo no es el conjunto vacío, es decir, si existe al menos un elemento $x \in X$ tal que $\mu_A(x)=1$ o, lo que es lo mismo, $h(A)=1$. Llamaremos *sub-normal* a un conjunto difuso que no cumpla la anterior condición.

Singletons.

El concepto de *singleton* es fundamental en la teoría de conjuntos difusos puesto que permite expresar un elemento cualquiera como un conjunto difuso, extendiendo por tanto de forma natural muchos conceptos propios de los conjuntos clásicos al campo de

los conjuntos difusos. Matemáticamente se define el *singleton* como un conjunto difuso cuyo soporte contiene sólo un elemento. Su función de pertenencia es, pues,

$$\mu_A(x) = \begin{cases} \mu_0 & x = x_0 \\ 0 & x \neq x_0 \end{cases} \quad (2.16)$$

y se representa generalmente con una línea vertical según muestra la Figura 2.4.

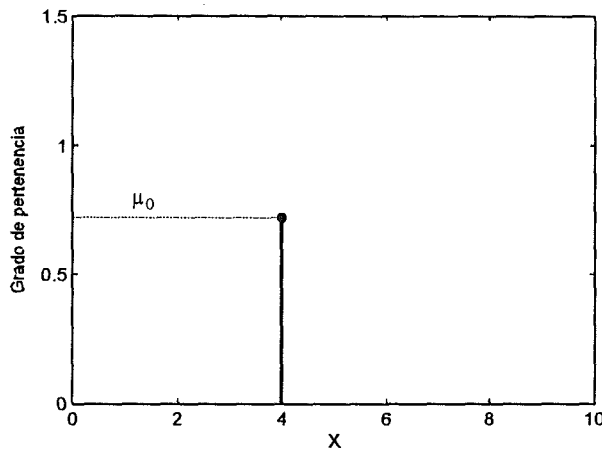


Figura 2.4: Representación gráfica de un singleton en X

Nótese que si el grado de pertenencia del elemento al singleton difuso es 1, entonces el singleton puede interpretarse también como un conjunto clásico que contiene un sólo elemento. Esta propiedad permite, por ejemplo, expresar directamente un número cualquiera como un conjunto difuso, lo cual es sumamente importante en operaciones como la *fuzzificación*, que se trata más adelante.

Números (valores) difusos y variables lingüísticas.

Los conjuntos difusos permiten generalizar el concepto de *número*, de la siguiente forma: Definimos un *número difuso* en un universo⁸ X como un conjunto difuso *normal* y *convexo* en X . Así definido, un número difuso modela a un número

⁸ El universo X es generalmente el conjunto de los números reales, aunque pueden definirse números difusos sobre espacios de distintas características y dimensiones, al igual que se hace para los números nitidos. Un número difuso n -dimensional recibe el nombre de *punto difuso*.

conocido con cierto grado de indeterminación, por lo que suele usarse para representar conceptos numéricos imprecisos, como “unos 30 grados”, “alrededor del 20 de Octubre”, o lingüísticos y dependientes del contexto, como “alto”, “pequeño”, etc.

Así como se definen valores (números) difusos, pueden definirse también *variables difusas*. Una *variable difusa* es una etiqueta lingüística que puede tomar distintos valores difusos, siendo estos últimos, como acabamos de ver, conjuntos difusos. Así, por ejemplo, en un contexto en el que nos refiramos a la temperatura del agua de un baño (universo $T=[0,100]$) podemos considerar la variable difusa “Temperatura”, cuyos valores difusos podrían ser los conceptos “baja”, “media” y “alta” definidos por las funciones de pertenencia $\mu_{baja}(T)$, $\mu_{media}(T)$, $\mu_{alta}(T)$, tal y como se muestra en la Figura 2.5:

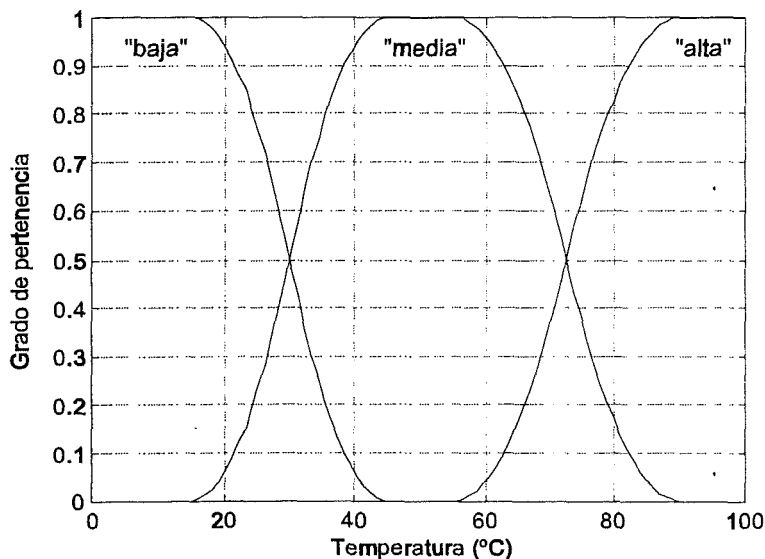


Figura 2.5: Variable difusa “Temperatura” y sus respectivos valores difusos

Si el número que se quiere representar no presenta incertidumbre, podemos considerarlo como un conjunto clásico que contiene un sólo elemento, o sea un singleton *normal*. Dado que un singleton es un caso particular de conjunto difuso, resulta obvio constatar que las variables difusas resultan un concepto idóneo para manejar valores con distintos grados de incertidumbre. Como consecuencia, los conceptos de valor y variable difusa se utilizan constantemente en la manipulación de expresiones difusas.

En esta figura también se aprecia una de las características más singulares de los conjuntos difusos, a saber, que las funciones de pertenencia asociadas a ellos pueden

solaparse. Como se pondrá de manifiesto más adelante, esta característica permite distribuir sobre más de un valor a la vez las decisiones que se toman en base a los valores de la variable, lo que contribuye en gran manera a la robustez de este tipo de razonamientos.

Fuzzificación y Desfuzzificación

El proceso de *fuzzificación* o *borrosificación* es el que permite convertir un valor nítido en difuso. Existen varias posibilidades de efectuar esta conversión, siendo la más sencilla de ellas considerar el valor nítido como un *singleton* difuso centrado en ese valor. Así, el número nítido 4 se fuzzifica identificándolo con el singleton definido por

$$\mu_4(x) = \begin{cases} 1 & x = 4 \\ 0 & x \neq 4 \end{cases} \quad (2.17)$$

Este método de fuzzificación, que presenta la ventaja de su sencillez, tiene la importante limitación de no poder incorporar la posible imprecisión o ambigüedad que suele acompañar a todo valor obtenido del mundo real, por ejemplo, como resultado de una medida. En efecto, dado que usamos la operación de fuzzificación para traducir valores nítidos al lenguaje difuso, parece lógico pensar que el proceso de fuzzificación debería tener en cuenta la ambigüedad presente y reflejarla de algún modo en el resultado final. Con esta idea se definen otras técnicas de fuzzificación que devuelven un número difuso centrado en el valor a fuzzificar. Por ejemplo, la ecuación (2.18) podría ser la salida de un fuzzificador a cuya entrada se presente el valor nítido 4:

$$\mu_4(x) = \begin{cases} 1 - \frac{|4-x|}{\xi} & x \in [(4-\xi), (4+\xi)] \\ 0 & \text{resto} \end{cases} \quad (2.18)$$

Los conjuntos difusos resultantes de fuzzificar un cierto valor x se representan en la Figura 2.6, para los dos tipos de fuzzificador antes citados.

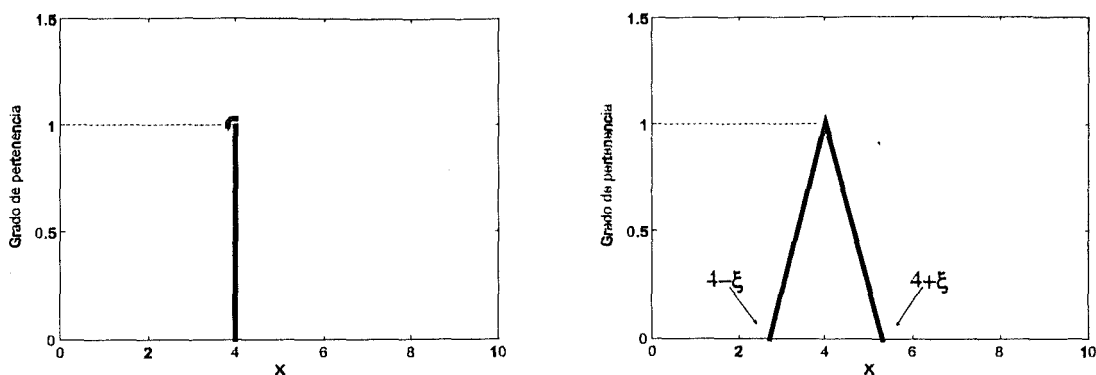


Figura 2.6: Fuzzificador por singleton y por conjunto difuso

El fuzzificador por conjunto difuso es, pues, capaz de reflejar la incertidumbre asociada al valor de entrada, pero es más complejo de construir y requiere la toma de ciertas decisiones de diseño, tales como la elección de la *forma* de la función de pertenencia asociada al número difuso y de valores de los parámetros que, como ξ en el ejemplo propuesto, la determinan. Todo ello con el objetivo de que la salida del fuzzificador exprese adecuadamente nuestra percepción de la incertidumbre.

En cuanto a la representatividad de resultados que producen, cabe decir que ambos métodos son prácticamente indistinguibles, pudiendo en todo caso trabajar mejor el fuzzificador por conjunto difuso en entornos contaminados por ruido [Wang 1994], aunque este último punto requiere un estudio más profundo para poder afirmarse con rotundidad. Puesto que en la práctica se busca la simplicidad de construcción y no suele considerarse explícitamente la presencia de ruido en los sistemas de inferencia difusa, el segundo tipo de fuzzificador se utiliza muy raramente, en beneficio del fuzzificador por singleton.

La operación inversa, o *desfuzzificación*, trata de encontrar el valor nítido que mejor represente a un conjunto difuso dado. Constituye una operación más compleja y delicada, puesto que es imposible reducir a un sólo número las sutilezas de un conjunto difuso -especialmente si éste tiene una función de pertenencia compleja-, con lo que el resultado obtenido puede no ser de mucha utilidad. En la literatura se han propuesto diferentes métodos de desfuzzificación, siendo el más utilizado el del *centroide*, que representa a un conjunto difuso por la abscisa del centroide de área bajo la curva (o volumen bajo la superficie si el conjunto es multidimensional) delimitada por la función

de pertenencia del conjunto. Así pues, si B es un conjunto difuso en Y representado por su función de pertenencia $\mu_B(y)$, el valor y_0 que lo desfuzzifica es:

$$y_0 = \frac{\int_Y \mu_B(y) \cdot y \cdot dy}{\int_Y \mu_B(y) \cdot dy} \quad (2.19)$$

Como se ha dicho anteriormente, en la gran mayoría de sistemas difusos se prefiere el método del centroide a cualquier otro para realizar la desfuzzificación. Esto es así porque este método es el que mejor sintetiza en un único valor toda la información contenida en el conjunto difuso. Sin embargo no está exento de problemas, ya que resulta complejo de implementar de forma circuital al contener su expresión un cociente. Por ello se han desarrollado otros métodos de desfuzzificación como el del *mayor máximo*, *menor máximo*, *media de los máximos*, *bisectriz del área*, etc., los cuales no se tratan aquí al no ser tan utilizados.

2.4 Operaciones sobre conjuntos

2.4.1 Operaciones sobre conjuntos nítidos.

Complemento, unión e intersección.

El *complemento relativo* del conjunto A respecto al conjunto B, denotado por B-A, es el conjunto de todos los elementos de B que no pertenezcan a A. Si el conjunto B es el conjunto universal X, hablamos de *complemento absoluto* o, simplemente, *complemento*, denotándose por \bar{A} . El complemento absoluto es *involutivo*, es decir, el complemento del complemento es el conjunto original ($\overline{\bar{A}} = A$). El complemento del conjunto vacío es el conjunto universal, y viceversa.

La *unión* de dos conjuntos A y B, denotada por $A \cup B$, es el conjunto que contiene los elementos que pertenecen a cualquiera de los conjuntos A y B, o a ambos a la vez. La *intersección* de dos conjuntos A y B, denotada $A \cap B$, es el conjunto que contiene los elementos que pertenecen tanto a A como a B. Tanto la unión como la intersección pueden generalizarse para un número cualquiera de conjuntos. La definición formal de todas estas operaciones se muestra a continuación

$$B-A = \{x \mid x \in B \text{ y } x \notin A\} \quad (2.20)$$

$$\bar{A} = \{x \mid x \notin A\} \quad (2.21)$$

$$A \cup B = \{x \mid x \in A \text{ o } x \in B\} \quad (2.22)$$

$$A \cap B = \{x \mid x \in A \text{ y } x \in B\} \quad (2.23)$$

Si se describen los conjuntos en base a sus funciones características y se aprovecha el hecho de que éstas solamente toman los valores 0 y 1, las operaciones de complemento, unión e intersección pueden escribirse como

$$\chi_{\bar{A}}(x) = 1 - \chi_A(x) \quad (2.24)$$

$$\chi_{A \cup B}(x) = \max\{\chi_A(x), \chi_B(x)\} \quad (2.25)$$

$$\chi_{A \cap B}(x) = \min\{\chi_A(x), \chi_B(x)\} \quad (2.26)$$

Es de destacar que las operaciones de complemento, unión e intersección dadas en (2.24)-(2.26) no son las únicas que pueden definirse. Por ejemplo, en (2.26) podría usarse el producto en lugar del mínimo, con idéntico resultado. Por convenio suele llamarse *t-norma* a cualquier operador de intersección válido, y *t-conorma* a cualquier operador de unión válido. Sin embargo, dado que todas las posibles t-normas, t-conormas y complementos devuelven los mismos resultados, el estudio de otros operadores distintos a (2.24)-(2.26) carece de sentido para conjuntos nítidos. Por el contrario, será un tema de gran interés en el estudio de conjuntos difusos, en donde cada operador devuelve resultados distintos.

Propiedades.

Las operaciones que se acaban de exponer presentan propiedades de gran interés en el tratamiento de expresiones matemáticas con conjuntos. En la Tabla 2-1 se muestran de forma compacta algunas de las más importantes.

Tabla 2-1:
Propiedades fundamentales de las operaciones en conjuntos nítidos

<i>Conmutatividad</i>	$A \cup B = B \cup A$ $A \cap B = B \cap A$
<i>Asociatividad</i>	$(A \cup B) \cup C = A \cup (B \cup C)$ $(A \cap B) \cap C = A \cap (B \cap C)$
<i>Distributividad</i>	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
<i>Idempotencia</i>	$A \cup A = A$ $A \cap A = A$
<i>Identidad</i> <i>(condiciones de frontera)</i>	$A \cup \emptyset = A$ $A \cap X = A$
<i>Absorción</i>	$A \cup X = X$ $A \cap \emptyset = \emptyset$ $A \cup (A \cap B) = A$ $A \cap (A \cup B) = A$
<i>Ley de contradicción</i>	$A \cap \bar{A} = \emptyset$
<i>Ley del tercero excluido</i>	$A \cup \bar{A} = X$
<i>Leyes de De Morgan</i>	$\overline{A \cup B} = \bar{A} \cap \bar{B}$ $\overline{A \cap B} = \bar{A} \cup \bar{B}$

De todas estas propiedades, muchas son extensibles a un número cualquiera de conjuntos, como sucede con las Leyes de De Morgan. En estas últimas puede apreciarse el carácter dual de las operaciones unión e intersección respecto a la de complemento, lo que es de gran utilidad en los casos en que resulta más cómodo definir un conjunto a partir de los elementos que *no* pertenecen a él, que hacerlo a partir de los que sí pertenecen.

Aunque muchas de estas propiedades seguirán siendo válidas en el dominio de los conjuntos difusos, unas pocas dejarán significativamente de serlo, marcando diferencias esenciales entre el punto de vista clásico y el difuso. De estas diferencias se hablará con mayor profundidad en la sección siguiente.

2.4.2 Operaciones sobre conjuntos difusos.

La discusión de las operaciones que pueden realizarse con conjuntos difusos es paralela a la desarrollada para sus homólogos nítidos, definiéndose operaciones de unión,

intersección, complemento, etc. Dichas operaciones son generalizaciones de las definidas para los conjuntos nítidos, de tal forma que si se restringe el rango de grados de pertenencia al conjunto $\{0,1\}$, se deben obtener los mismos resultados que en el caso nítido.

Al igual que para los conjuntos nítidos, también para los difusos existen varios operadores que satisfacen las condiciones exigibles para ser considerados complementos, uniones o intersecciones. La gran diferencia estriba en que en el caso nítido todos los operadores posibles para una misma operación devolvían idénticos resultados, mientras que esto no ocurre en el caso difuso, en donde distintos operadores devuelven valores distintos cuando sus argumentos son números *entre* 0 y 1.

Este comportamiento ha generado numerosos estudios sobre qué operadores emplear en el tratamiento de conjuntos difusos, sin que se haya llegado a conclusiones definitivas. En la mayoría de situaciones se utilizan los llamados *operadores estándar*, pero se acepta el uso de otros en aquellas aplicaciones en las que se demuestre que conducen a mejores resultados. Es más, en las numerosas aplicaciones prácticas de la lógica difusa, la elección de un operador u otro es considerada por muchos autores como un parámetro más de diseño.

Debido a ésto, se presentan a continuación los operadores difusos separados en dos apartados: primeramente los operadores estándar y en segundo lugar se discuten otros, distinguiendo entre complementos, t-normas y t-conormas.

Operadores estándar.

Los operadores que se presentan en este apartado son, con mucho, los de utilización más frecuente, por lo que reciben el nombre de estándar. Para dos conjuntos difusos A y B definidos en el universo X, los conjuntos complemento (absoluto), unión e intersección se definen como los conjuntos difusos cuyas funciones de pertenencia están dadas por

Complemento: $\mu_{\bar{A}}(x) = 1 - \mu_A(x)$ (2.27)

Unión: $\mu_{A \cup B}(x) = \max\{\mu_A(x), \mu_B(x)\}$ (2.28)

Intersección: $\mu_{A \cap B}(x) = \min\{\mu_A(x), \mu_B(x)\}$ (2.29)

La elección de estas funciones para implementar el complemento, unión e intersección difusa está plenamente justificada. Por una parte parece lógico definir el conjunto *complemento de A* como aquel conjunto difuso cuyos elementos pertenecen a él con un grado igual al de *no* pertenencia a A. Para la *unión* y la *intersección* la justificación es parecida. El uso de la función *máximo* como operador de unión salva el contrasentido de que un elemento pertenezca a $A \cup B$ con un grado inferior al de su pertenencia a uno de los dos conjuntos. De forma similar al elegir la función *mínimo* para la intersección evita el absurdo que un elemento pueda pertenecer a $A \cap B$ con un grado superior que a cualquiera de ellos.

Es interesante destacar que las expresiones de los operadores estándar para los conjuntos difusos coinciden con las definidas para los nítidos. Esto no es en modo alguno casual: puesto que cualquier operador de complemento, unión o intersección devuelve idénticos resultados en el caso nítido, se escogen precisamente los definidos en (2.21)-(2.23) para que no varíen al extender su campo de operación al caso difuso.

Los operadores estándar fueron propuestos originalmente por Zadeh en su primer artículo sobre conjuntos difusos [Zadeh 1965], y posteriormente otros autores han estudiado en profundidad sus propiedades. La Tabla 2-2 resume las más relevantes de ellas, que son esencialmente las mismas que se cumplen para el caso nítido.

Tabla 2-2:
Propiedades fundamentales de las operaciones estándar en conjuntos difusos

<i>Conmutatividad</i>	$A \cup B = B \cup A$ $A \cap B = B \cap A$
<i>Asociatividad</i>	$(A \cup B) \cup C = A \cup (B \cup C)$ $(A \cap B) \cap C = A \cap (B \cap C)$
<i>Distributividad</i>	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Tabla 2-2:
Propiedades fundamentales de las operaciones estándar en conjuntos difusos
(continuación)

<i>Idempotencia</i>	$A \cup A = A$ $A \cap A = A$
<i>Identidad</i>	$A \cup \emptyset = A$ $A \cap X = A$
<i>Absorción</i>	$A \cup X = X$ $A \cap \emptyset = \emptyset$ $A \cup (A \cap B) = A$ $A \cap (A \cup B) = A$
<i>Leyes de De Morgan</i>	$\overline{A \cup B} = \overline{A} \cap \overline{B}$ $\overline{A \cap B} = \overline{A} \cup \overline{B}$

De la comparación entre esta tabla y la Tabla 2-1 se observa que existen dos leyes fundamentales que no se conservan: la ley de contradicción y la ley del tercero excluido (*tertium non datur*), que son duales entre sí como ponen de manifiesto las leyes de De Morgan. Es decir, que para un conjunto difuso A definido en X se tiene que, en general, $A \cap \overline{A} \neq \emptyset$ y $A \cup \overline{A} \neq X$. Esto es una consecuencia natural del hecho de abandonar el concepto de pertenencia o no pertenencia estrictas de un elemento a un conjunto, pero las implicaciones filosóficas de este hecho son numerosas y profundas, pues contradicen los cimientos en los que se basa la matemática tradicional y, por extensión, todo el edificio científico construido sobre ellos.

Por ejemplo, la no verificación de estas leyes va a significar que un elemento puede pertenecer a la vez a un conjunto y a su complementario, lo que extendido al campo de la lógica lleva a admitir que es posible aseverar simultáneamente un hecho y su contrario sin que ello constituya un absurdo. Una matemática basada en estos axiomas no puede, por ejemplo, basar la demostración de teoremas en una reducción al absurdo. De hecho, el mismo hecho de probar un teorema debe verse ahora desde un punto de vista difuso, y asignar grados a sus pertenencias al grupo de los teoremas probados y al de los refutados. Otra consecuencia importante es que en teoría de conjuntos difusos, el todo no es necesariamente la suma de sus partes, cosa que entra en conflicto, por ejemplo, con la teoría de probabilidades, siendo uno de los argumentos para refutar a los que consideran la teoría de conjuntos difusos como una forma encubierta de la primera.

Cabe señalar sin embargo que el punto de vista difuso no invalida el tradicional, pero relega su aplicabilidad a aquellos casos que se fundamentan en una visión bivalente de la teoría de conjuntos o del mundo físico en general.

Aunque los operadores difusos estándar de complemento, unión e intersección son generalizaciones de los definidos en (2.24), (2.25) y (2.26) para el caso nítido, numerosos estudios sobre el tema han mostrado que no son las únicas generalizaciones posibles, sino que para cada una de las operaciones mencionadas existen varias familias de funciones susceptibles de ser utilizadas, y que se comportan como sus equivalentes nítidas cuando el rango de grados de pertenencia manejados se reduce al conjunto $\{0,1\}$. El problema que ello plantea es que, mientras las distintas familias posibles de operadores de complemento, unión e intersección arrojan los mismos resultados cuando manejan sólo ceros o unos, devuelven resultados *distintos* cuando sus argumentos están entre 0 y 1. Sucede, por tanto, que la elección apropiada de un u otro operador es una decisión que depende del contexto del problema, y que puede llegar a tener tanta influencia sobre el resultado final como la definición que se haga de la función de pertenencia para expresar un determinado concepto difuso.

A continuación se muestran brevemente otras familias de operadores difusos, indicando las condiciones que deben de cumplir en cada caso.

Complementos difusos.

Los operadores de complemento difuso son funciones que toman como argumento el grado de pertenencia de un elemento x a un cierto conjunto difuso A y devuelven el grado en que x pertenece al conjunto complementario de A , que puede interpretarse como el grado en que x *no pertenece* a A . Denotando con la letra c a todos los posibles operadores de complemento, éstos pueden definirse como una aplicación del intervalo $[0,1]$ sobre si mismo, es decir

$$c: [0,1] \rightarrow [0,1] \quad (2.30)$$

Para que un operador de complemento tenga sentido como tal, tendrá que verificar además los dos axiomas siguientes:

Axioma 1: $c(0) = 1$ y $c(1) = 0$ (condiciones de frontera)

Axioma 2: Para todo $a, b \in [0,1]$, si $a \leq b$ entonces $c(a) \geq c(b)$ (monotonicidad)

El primer axioma establece las condiciones -llamadas *de frontera*- para que el complemento nítido sea un caso particular del difuso, mientras que el segundo indica que la pertenencia a un conjunto en un grado mayor conlleva una pertenencia menor al complemento.

A continuación se presentan algunas de las familias o *clases* de complementos que cumplen estos axiomas. En general cada familia se expresa en función de un parámetro cuya variación permite obtener distintos complementos de la familia.

$$\text{Clase Sugeno:} \quad c_{\lambda}(a) = \frac{1-a}{1+\lambda \cdot a} \quad \lambda \in (-1, \infty) \quad (2.31)$$

$$\text{Clase Yager:} \quad c_{\omega}(a) = (1-a^{\omega})^{1/\omega} \quad \omega \in (0, \infty) \quad (2.32)$$

$$\text{Clase umbral:} \quad c_t(a) = \begin{cases} 1 & \text{para } a \leq t \\ 0 & \text{para } a > t \end{cases} \quad t \in [0,1] \quad (2.33)$$

En algunos casos es conveniente que los operadores de complemento cumplan restricciones axiomáticas adicionales. Dos de las más comunes se recogen en los axiomas 3 y 4. Es fácil comprobar que el complemento estándar y las clases de Sugeno y Yager cumplen los cuatro axiomas expuestos, mientras que la clase umbral sólo cumple los dos básicos.

Axioma 3: c es una función continua

Axioma 4: $c(c(a)) = a$ para cada $a \in [0,1]$ (involución)

La Figura 2.7 muestra algunas de las funciones complemento correspondientes a la clase Yager, para distintos valores del parámetro ω . En ella puede observarse que para valores de $\omega \neq 1$ la función complemento no es una recta, -como sucede en el complemento estándar-, lo que da lugar a fenómenos curiosos, contrarios a la intuición. Por ejemplo, la Figura 2.8 muestra claramente cómo para valores $\omega \neq 1$, la suma de los grados de pertenencia de un elemento a un conjunto y a su Yager-complementario no es 1, en

contra de lo esperado. Un análisis similar muestra que lo mismo ocurre con el complemento Sugeno para $\lambda \neq 0$.

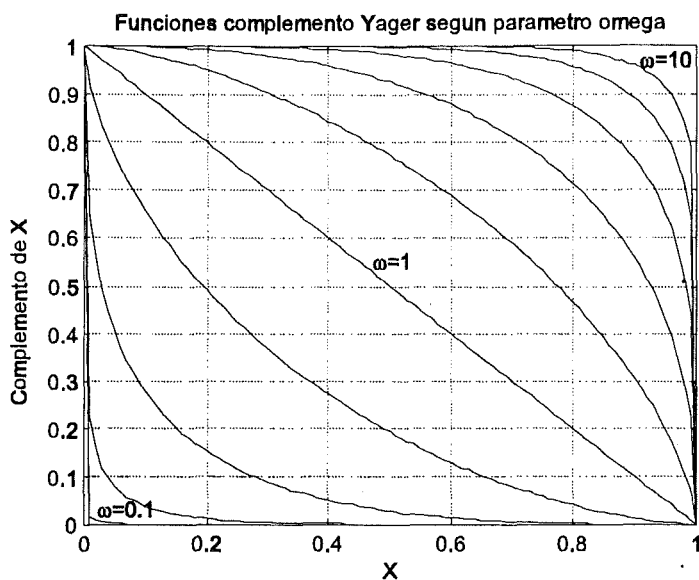


Figura 2.7: Gráfica de las funciones complemento para la clase Yager

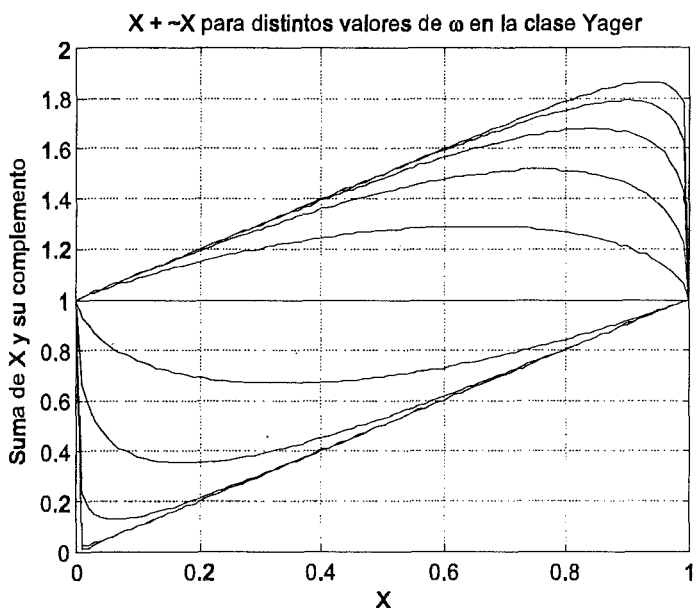


Figura 2.8: Suma del grado de pertenencia y su Yager-complementario para distintos valores del parámetro ω

El hecho de que $\mu_A(x) + \mu_{\bar{A}}(x) \neq 1$ puede tener interés en aquellas situaciones caracterizadas por un alto grado de incertidumbre, en las que un pequeño grado de pertenencia a un conjunto no implique necesariamente que el correspondiente al complementario deba ser alto. De todas formas, dado que es difícil encontrar una

situación práctica en la que esto se produzca, actualmente se ha generalizado el uso del operador complemento estándar (2.27) en la mayor parte de aplicaciones, a pesar de las múltiples alternativas disponibles.

Uniones (t-conormas) difusas.

Los operadores de unión difusa son funciones que toman como argumentos los grados de pertenencia de un elemento x a dos conjuntos difusos A, B y devuelven el grado en que x pertenece al conjunto unión $A \cup B$. Denotando con la letra u a todos los posibles operadores de unión difusa se tiene que éstos responden a la estructura de la ecuación (2.34), en donde el símbolo \times denota al producto cartesiano de dos conjuntos, concepto que se desarrollará con todo detalle en la sección 2.5.

$$u: [0,1] \times [0,1] \rightarrow [0,1] \tag{2.34}$$

Para dotar de sentido a los operadores de unión difusa, éstos deben satisfacer los siguientes axiomas para todo $a, b, d \in [0,1]$:

Axioma 1: $u(a,0) = a$ (condición de frontera)

Axioma 2: $b \leq d \Rightarrow u(a, b) \leq u(a, d)$ (monotonicidad)

Axioma 3: $u(a, b) = u(b, a)$ (conmutatividad)

Axioma 4: $u(a, u(b,d)) = u(u(a, b), d)$ (asociatividad)

A continuación se listan las t-conormas que, aparte de la estándar, suelen usarse con más frecuencia para implementar la unión difusa.

Suma algebraica: $u(a, b) = a+b-ab$ (2.35)

Suma acotada: $u(a, b) = \min\{1, a+b\}$ (2.36)

Unión drástica: $u(a, b) = \begin{cases} a & \text{si } b = 0 \\ b & \text{si } a = 0 \\ 1 & \text{en caso contrario} \end{cases}$ (2.37)

Clase Yager: $u_w(a, b) = \min\{1, [a^w + b^w]^{1/w}\}$ $w > 0$ (2.38)

Entre todas estas funciones no hay ninguna de uso mucho más extendido que el resto, por lo que en general se recurre a la unión estándar. Puede demostrarse también que, de entre todas las t-conormas existentes, la unión estándar es la que devuelve el valor menor de todas ellas, es decir, si $u(a,b)$ denota a cualquier t-conorma, se cumple que $\max(a,b) \leq u(a,b) \leq 1$.

Dada la multiplicidad de operaciones distintas existentes para la unión difusa, resulta conveniente establecer un símbolo común que denote a todas ellas. En este caso la elección más habitual es utilizar el símbolo \oplus , de tal forma que podemos escribir genéricamente:

$$\mu_{A \cup B}(x) = \mu_A(x) \oplus \mu_B(x) \quad (2.39)$$

Intersecciones (t-normas) difusas.

Los operadores de intersección difusa son idénticos en su estructura a las t-conormas difusas que acaban de tratarse. Como aquéllas, son funciones que toman como argumentos los grados de pertenencia de un elemento x a dos conjuntos difusos A , B y devuelven el grado en que x pertenece al conjunto intersección $A \cap B$. Denotando con la letra i a todos los posibles operadores de intersección se tiene que

$$i: [0,1] \times [0,1] \rightarrow [0,1] \quad (2.40)$$

También los operadores de intersección difusa deben satisfacer los siguientes axiomas para todo $a, b, d \in [0,1]$:

Axioma 1: $i(a,1) = a$ (condición de frontera)

Axioma 2: $b \leq d \Rightarrow i(a, b) \leq i(a, d)$ (monotonicidad)

Axioma 3: $i(a, b) = i(b, a)$ (conmutatividad)

Axioma 4: $i(a, i(b,d)) = i(i(a, b), d)$ (asociatividad)

A continuación se listan las t-normas que se suelen usar con más frecuencia (después de la intersección estándar) para implementar la intersección difusa.

Producto algebraico: $i(a, b) = a \cdot b$ (2.41)

Diferencia acotada: $i(a, b) = \max(0, a+b-1)$ (2.42)

Intersección drástica: $i(a, b) = \begin{cases} a & \text{si } b = 1 \\ b & \text{si } a = 1 \\ 0 & \text{en caso contrario} \end{cases}$ (2.43)

Clase Yager: $i_w(a, b) = 1 - \min\{1, [(1-a)^w + (1-b)^w]^{1/w}\}$ $w > 0$ (2.44)

Como ocurre para las t-conormas, en general se prefiere el uso de la intersección estándar a cualquier otra t-norma. Sin embargo, de entre las no estándares, el **producto algebraico** es con mucho la más extendida, quizás porque conduce a expresiones con un tratamiento analítico más sencillo.

De forma dual a lo que ocurre para las t-conormas, puede demostrarse también que, de entre todas las t-normas existentes, la intersección estándar es la que devuelve el valor mayor de todas ellas, es decir, si $i(a, b)$ denota a cualquier t-norma, se cumple que $0 \leq i(a, b) \leq \min(a, b)$.

Es de interés resaltar que no todos los complementos, t-normas y t-conormas definidos hasta ahora verifican las leyes de De Morgan, hecho que podría constituir un problema en ciertas aplicaciones. Dado que las leyes de De Morgan involucran siempre a las tres operaciones mencionadas, en general procuraremos escoger aquellas ternas que verifiquen dichas leyes, hablando entonces de la t-norma i y la t-conorma u como *duales* respecto del complemento c . La Tabla 2-3 lista algunas de estas ternas.

Tabla 2-3: Ternas de operadores difusos duales.

<i>Nombre</i>	<i>complemento</i>	<i>t-conorma</i>	<i>t-norma</i>
Zadeh (estándar)	1-x	max(x,y)	min(x,y)
Probabilístico	1-x	x·y	x+y-xy
Lukasiewicz	1-x	min{1, a+b}	max(0, a+b-1)
Weber (drástico)	1-x	Intersección drástica	Unión drástica

Cabe recordar que, tal y como se expone en la Tabla 2-2, los operadores estándar de unión e intersección son también duales respecto al complemento estándar.

Al igual que sucede para las t-conormas, la multiplicidad de operaciones distintas existentes para la intersección difusa, denotaremos mediante el símbolo genérico $*$ a todas ellas, de tal forma que podemos escribir genéricamente:

$$\mu_{A \cap B}(x) = \mu_A(x) * \mu_B(x) \quad (2.45)$$

2.5 Relaciones

El concepto de relación nítida juega un papel fundamental en la teoría de conjuntos, puesto que a partir de él se derivan conceptos tan importantes como el de aplicación o función. Además, mediante su generalización se llega al concepto de *relación difusa*, que es una de las bases del razonamiento aproximado, parte esencial de todo sistema de inferencia difusa.

En esencia, una *relación* entre dos o más conjuntos es una correspondencia que se establece entre ciertos elementos de dichos conjuntos. Esta correspondencia fija los criterios en base a los cuales se relacionan los elementos. Para definir matemáticamente las relaciones, es preciso introducir previamente el concepto de *producto cartesiano*.

2.5.1 Producto Cartesiano de conjuntos nítidos.

El producto cartesiano de n universos es un concepto simple, consistente en un conjunto que agrupa de forma ordenada a todas las combinaciones posibles de elementos de los n universos, formando las denominadas *n-plas*. Antes de dar una definición más formal, definiremos una *n-pla* como una secuencia ordenada de n elementos, escrita en la forma (x_1, x_2, \dots, x_n) . Puesto que un universo no es más que un tipo de conjunto, en adelante se trata el producto cartesiano de conjuntos en general.

Para los n conjuntos nítidos A_1, A_2, \dots, A_n , definimos el *Producto Cartesiano* de dichos conjuntos (en el orden indicado) como el conjunto de todos las *n-pla* posibles tales que $x_1 \in A_1, \dots, x_n \in A_n$, escribiéndose $A_1 \times A_2 \times \dots \times A_n$ o, simplemente

$$\prod_{1 \leq i \leq n} A_i = \{(x_1, x_2, \dots, x_n) \mid x_i \in A_i \quad \forall i = 1, 2, \dots, n\} \quad (2.46)$$

De la definición se deduce fácilmente que la cardinalidad del conjunto producto cartesiano es igual a el producto de cardinalidades de los conjuntos que lo forman.

2.5.2 Relaciones nítidas.

Relación entre n conjuntos

Con el producto Cartesiano definido en (2.46), puede ahora expresarse una relación entre los conjuntos A_1, A_2, \dots, A_n , como un **subconjunto** del producto cartesiano $A_1 \times A_2 \times \dots \times A_n$. Esta relación recibe el nombre de *relación n-aria* y se denota generalmente por $R(A_1, A_2, \dots, A_n)$ o, si no hay ambigüedades, simplemente por R . Llamamos *relación binaria* al caso particular de una relación entre dos conjuntos, el primero de los cuales recibe el nombre de *conjunto salida*, y el segundo *conjunto llegada*.

A continuación se exponen los métodos más usuales para representar las relaciones nítidas n-arias.

Representación de relaciones nítidas

Dado que definimos las relaciones como conjuntos nítidos, podemos utilizar cualquiera de los métodos descritos en 2.3.1 (enumeración, regla, función característica) para representarlas, teniendo en cuenta que cada elemento de la relación es ahora una n-pla. En particular, usando la función característica se tiene que para una relación $R(A_1, A_2, \dots, A_n)$

$$\chi_R(x_1, x_2, \dots, x_n) = \begin{cases} 1 & \text{si } (x_1, x_2, \dots, x_n) \in R \\ 0 & \text{si } (x_1, x_2, \dots, x_n) \notin R \end{cases} \quad (2.47)$$

en donde, evidentemente, $x_i \in A_i, \dots, x_n \in A_n$.

En el caso particular de que los conjuntos A_i sean finitos, existen otras formas convenientes para representar relaciones, como son los *diagramas sagitales* o las *matrices de relación*. En la primera de ellas se dibujan agrupados los elementos de los distintos conjuntos sobre los que está definida la relación, y se unen mediante líneas aquellos elementos pertenecientes a un mismo par ordenado. Así, por ejemplo, para los conjuntos P (países) y M (monedas) definidos como $P = \{\text{Alemania, Canadá, EE.UU., Francia}\}$ y $M = \{\text{Dólar, Peseta, Marco}\}$, el diagrama sagital que relaciona los países con monedas es el mostrado en la Figura 2.9:

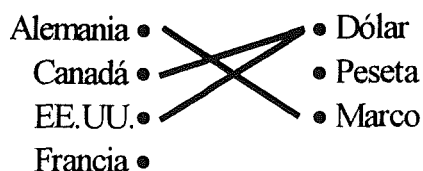


Figura 2.9: Representación de relaciones mediante diagramas sagitales

El método matricial consiste en crear una matriz de dimensiones $|A_1| \times |A_2| \times \dots \times |A_n|$ -donde $| \cdot |$ representa la cardinalidad-, y etiquetar cada una de las filas/columnas con los nombres de los elementos de los correspondientes conjuntos. De esta forma cada elemento de la matriz corresponde a una n-pla del producto cartesiano $A_1 \times A_2 \times \dots \times A_n$. Asignando a cada uno de los elementos el valor 1 si la n-pla pertenece a la relación y el valor 0 en caso contrario, se obtiene la matriz de relación $[R]$ asociada a la relación R .

Por razones obvias de representación, el método matricial se aplica principalmente a relaciones **binarias**, donde la matriz resultante es bidimensional. Para el ejemplo anterior, la matriz de relación se muestra en la Figura 2.10:

	<i>Dólar</i>	<i>Peseta</i>	<i>Marco</i>
Alemania	0	0	1
Canadá	1	0	0
EE.UU.	1	0	0
Francia	0	0	0

Figura 2.10: Representación de relaciones mediante matrices de relación

Dado que expresamos las relaciones como conjuntos, podemos definir sobre ellas uniones, intersecciones, complementos, etc. Las particularidades de estas operaciones ya se han tratado en la sección 2.4.1, por lo que es innecesario repetirlas para este caso.

2.5.3 Composición de relaciones nítidas.

Sea R una relación **binaria** entre elementos de dos conjuntos A y B , y S otra relación binaria entre los conjuntos B y C . Dado que el universo de llegada de R es el mismo que el de salida de S , queda establecida implícitamente una relación entre A y C , a la que denominaremos T . Se define entonces la *composición* como aquella operación matemática que permite describir la relación T a partir de las descripciones de las

relaciones R y S. La operación de composición se denota con el símbolo \circ , con lo cual se escribe $T = R \circ S$.

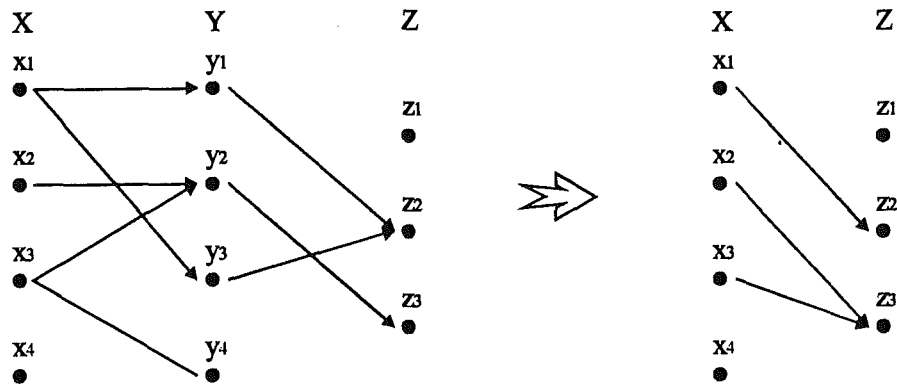


Figura 2.11: Composición de relaciones nítidas

En el diagrama sagital de la Figura 2.11 puede apreciarse que el elemento $x_i \in A$ estará relacionado con el elemento $z_k \in C$ si y sólo si existe alguna conexión entre ellos a través de *cualquier* elemento intermedio $y_j \in B$. Si expresamos las relaciones a través de sus funciones características, entonces la composición puede escribirse como

$$\chi_T(x, z) = \chi_{R \circ S}(x, z) = \max_{y \in B} (\min(\chi_R(x, y), \chi_S(y, z))) \quad (2.48)$$

o bien, usando el producto en lugar de la función mínimo

$$\chi_T(x, z) = \max_{y \in B} (\chi_R(x, y) \cdot \chi_S(y, z)) \quad (2.49)$$

Las expresiones anteriores reciben el nombre de *composición max-min* y *composición max-prod*. Las dos son equivalentes, ya que las funciones mínimo y producto devuelven idénticos resultados cuando sus argumentos sólo pueden tomar los valores 0 y 1. El algoritmo que siguen consiste en explorar los elementos y_j del conjunto intermedio para ver si existe o no un camino de relación entre x y z que pase por y_j (función mínimo o función producto). Esta comprobación se repite para todos los y_j , asignando finalmente un 1 a $\chi_T(x, z)$ si existe por lo menos un camino entre x y z (función máximo), y un 0 en caso contrario.

El algoritmo que acabamos de describir puede aplicarse en la composición de relaciones tanto entre conjuntos finitos como infinitos, si bien para estos últimos debe sustituirse la operación *máximo* por el *supremo* (la mínima cota superior). En el caso particular de relaciones entre universos finitos, existe una expresión cerrada para expresar la matriz de relación de la composición $T = R \circ S$ en función de las matrices de relación de R y S , consistente en hallar la matriz de relación resultante, $[T]$, como

$$[T] = [R] \otimes [S] \quad (2.50)$$

donde el operador \otimes es similar al producto matricial ordinario, pero sustituyendo la suma algebraica por el **máximo** y el producto algebraico por el **mínimo** o el **producto**, según queramos obtener la composición max-min o la max-prod. Para la relación de la Figura 2.11, la composición matricial se muestra en la Figura 2.12.

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Figura 2.12: Composición matricial de relaciones nítidas

Es de resaltar que hasta ahora se ha presentado la composición sólo para relaciones binarias, porque constituye un caso sencillo y fácil de ilustrar. Como es lógico, la discusión puede extenderse a una composición general entre relaciones n-arias de la siguiente forma: Sean R y S dos relaciones en $A_1 \times \dots \times A_r$ y $A_s \times \dots \times A_n$ respectivamente, siendo $s \leq r$ (teniendo, por tanto, $r-s+1$ dimensiones compartidas). Su composición $T=R \circ S$ es una relación en $A_1 \times \dots \times A_{r-1} \times A_{s+1} \times \dots \times A_n$ definida por la función característica:

$$\begin{aligned} \chi_T(x_1, \dots, x_{r-1}, x_{s+1}, \dots, x_n) = \\ = \max_{(x_r, \dots, x_s) \in A_r \times \dots \times A_s} \left\{ \min(\chi_R(x_1, \dots, x_r), \chi_S(x_s, \dots, x_n)) \right\} \end{aligned} \quad (2.51)$$

es decir, se eliminan las dimensiones comunes tomando el máximo respecto de las citadas dimensiones.

2.5.4 Relaciones difusas.

Relación difusa entre n conjuntos

En 2.5.2 se definió una relación nítida como un subconjunto de un cierto producto cartesiano. Ello permite una rápida generalización para el caso difuso de la siguiente forma: Definimos una *relación difusa* R entre los conjuntos nítidos A_1, A_2, \dots, A_n , como un subconjunto *difuso* del producto cartesiano $A_1 \times A_2 \times \dots \times A_n$, en donde las n-pla (x_1, x_2, \dots, x_n) que constituyen los elementos de la relación llevan asociado un grado de pertenencia $\mu_R(x_1, x_2, \dots, x_n)$ entre 0 y 1. Este grado de pertenencia suele interpretarse como la “intensidad” de la relación que liga entre sí los elementos de la n-pla.

Es de destacar el hecho de que en la definición que acaba de darse lo difuso es la relación en sí, pero no los conjuntos involucrados en el producto cartesiano. Recuérdese que la expresión que se utiliza para el producto cartesiano (2.46) se ha definido en base a conjuntos nítidos. Es posible también definir un *producto cartesiano difuso*⁹ tomando como base conjuntos difusos, pero ello no conlleva que las relaciones definidas sobre un tal producto sean automáticamente difusas. El producto cartesiano difuso se utiliza principalmente como base teórica para otros desarrollos como puede ser el principio de extensión, el cual se detalla en la sección 2.6.2.

De la definición de relación difusa también se desprende que toda relación difusa n-aria puede considerarse como un conjunto difuso n-dimensional en $A_1 \times A_2 \times \dots \times A_n$, con función de pertenencia $\mu_A(x_1, x_2, \dots, x_n)$. Este punto de vista permite intercambiar conjunto con relación en conceptos clave como la composición.

Un caso interesante lo constituyen las relaciones binarias designadas con $R(A, A)$ o $R(A^2)$ y caracterizadas por que los conjuntos salida y llegada coinciden. En estas relaciones se definen tres propiedades características: *reflexividad*, *simetría* y *transitividad*. La posesión o no de estas propiedades determina distintos tipos

⁹ Véase ecuación (2.71).

importantes de relación binaria (equivalencia, compatibilidad, orden, etc.) los cuales forman la base para construir interesantes estructuras matemáticas, como los retículos (*lattices*). Debido a su extensión y escaso impacto en la presente tesis, no se profundizará en el tratamiento de estas cuestiones.

Representación de relaciones difusas

Dada la similitud estructural entre las relaciones nítidas y las difusas, podemos usar para representar las segundas los mismos métodos que se han descrito para las primeras, adaptados -cuando se requiera- para mostrar con claridad el grado de pertenencia de cada elemento de la relación. Así, la forma más general de representar relaciones difusas de cualquier tipo es mediante la expresión matemática de la función de pertenencia $\mu_R(x_1, x_2, \dots, x_n)$, dada para toda n-pla (x_1, x_2, \dots, x_n) . Como en el caso nítido, para conjuntos finitos pueden usarse también los diagramas sagitales y las matrices de relación, no sin antes asociar el grado de pertenencia adecuado a cada línea del diagrama, en el primer caso, o a cada elemento de la matriz, en el segundo.

Cabe señalar que para el caso particular de relaciones *binarias* sobre universos con soporte infinito, una forma muy conveniente de representar la relación es mediante el gráfico tridimensional que resulta de dibujar la función $\mu_R(x, y)$. A modo de ejemplo, considérese la relación difusa entre los conjuntos $A=[-3,3]$ y $B=[-5,10]$ expresada verbalmente mediante la frase “*y es aproximadamente el cuadrado de x*”. Intuitivamente deberíamos asignar grados de pertenencia de cada tupla (x, y) a la relación que sean tanto más bajos cuanto más difiera y de x^2 , Una forma de expresar matemáticamente esta idea es mediante una función de pertenencia de tipo gaussiano, con el máximo en aquellos puntos en donde y sea exactamente x^2 , pero decreciendo suavemente a medida que y se aleje de ese valor, como puede ser la función

$$\mu_R(x, y) = e^{-\left(\frac{y - x^2}{\sigma}\right)^2} \quad (2.52)$$

en la que el parámetro σ debe de elegirse para que la relación exprese lo más fielmente posible nuestro concepto de “aproximadamente”. La Figura 2.13 muestra la función de pertenencia de la relación para $\sigma^2=5$.

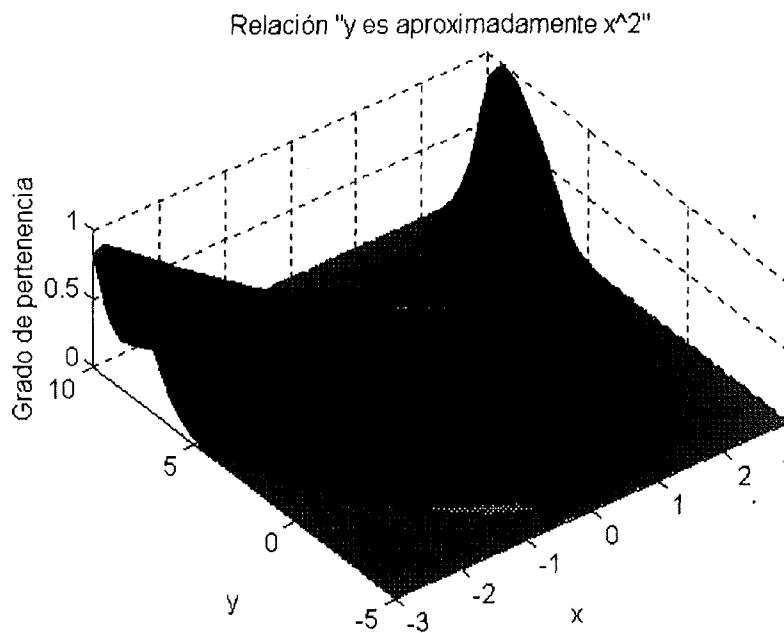


Figura 2.13: Representación 3-D de una relación difusa binaria

Evidentemente, podrían plantearse objeciones a esta descripción, como por ejemplo el hecho que se otorguen grados de pertenencia no nulos para valores de y negativos. La decisión de utilizar esta función u otra se tomará atendiendo a los resultados que se obtengan en cada aplicación concreta.

2.5.5 Proyección de una relación.

Sea una relación n -aria R definida sobre A_1, \dots, A_n . Sea también la secuencia $j \dots k$ una subsecuencia de $1 \dots n$. Se define la *proyección* de la relación R sobre el subespacio A_j, \dots, A_k (denotada $proj[R; A_j, \dots, A_k]$ o también $[R \downarrow A_j, \dots, A_k]$) como aquella relación obtenida mediante la expresión

$$proj[R; A_j, \dots, A_k] = \max_{(x_l, \dots, x_m) \in A_l \times \dots \times A_m} \{ \mu_R(x_1, \dots, x_n) \} \quad (2.53)$$

en la que $l \dots m$ representa la secuencia complementaria a $j \dots k$ en $1 \dots n$. En otras palabras, la proyección se obtiene eliminando la dependencia de R respecto de las variables (x_l, \dots, x_m) mediante el uso de la función *máximo* sólo sobre esas variables. Como en otras ocasiones se ha comentado, el máximo se sustituye por el *supremo* en caso de relaciones continuas.

La Figura 2.14 muestra el proceso para obtener la proyección de una relación binaria en $A \times B$ sobre B . La relación original es la dada en la ecuación (2.52) y representada en la Figura 2.13. En este caso la proyección resulta ser una relación unaria, por lo que puede considerarse un conjunto difuso unidimensional en B , el cual se representa en rojo en el gráfico.

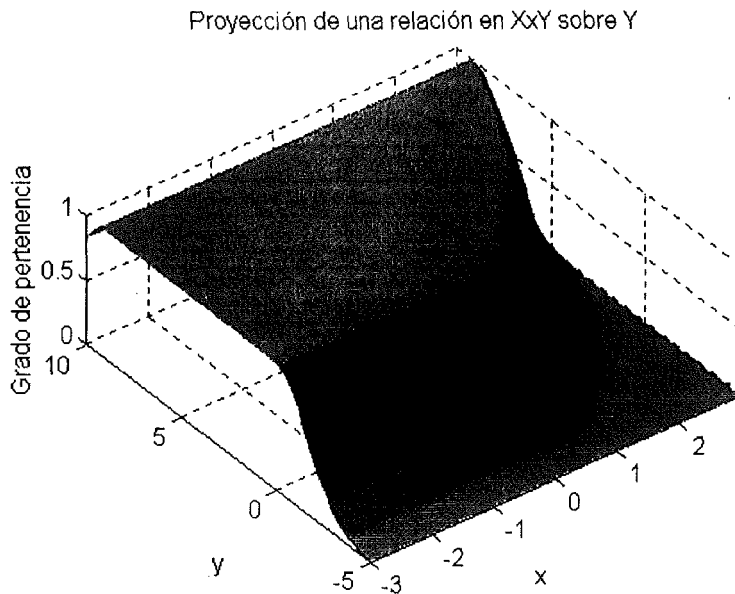


Figura 2.14: Proyección de una relación difusa en $A \times B$ sobre B

Bajo circunstancias especiales, el concepto de proyección puede generalizarse reemplazando el operador máximo por otra t-conorma [Klir 1995]. De hecho, el uso de cualquier función que efectúe una agregación de valores en las variables deseadas aunque

2.5.6 Extensión cilíndrica de una relación.

La extensión cilíndrica es el caso inverso a la proyección. Sea R una relación difusa definida en A_j, \dots, A_k . Su extensión cilíndrica al espacio A_1, \dots, A_n es una relación difusa dada por la función de pertenencia

$$\mu_R(x_1, \dots, x_n) = \mu_R(x_j, \dots, x_k) \quad (2.54)$$

o, en notación de Zadeh

$$\mu_R(x_1, \dots, x_n) = \int_{A_1 \times \dots \times A_n} \mu_R(x_j, \dots, x_k) / (x_1, \dots, x_n) \quad (2.55)$$

En resumen, la extensión cilíndrica se obtiene incrementando el número de variables de la función de pertenencia y haciendo dicha función independiente de las nuevas variables. La Figura 2.15 muestra en rojo a un conjunto difuso (que puede verse como una relación unaria) definido en A y su extensión al espacio definido por A×B:

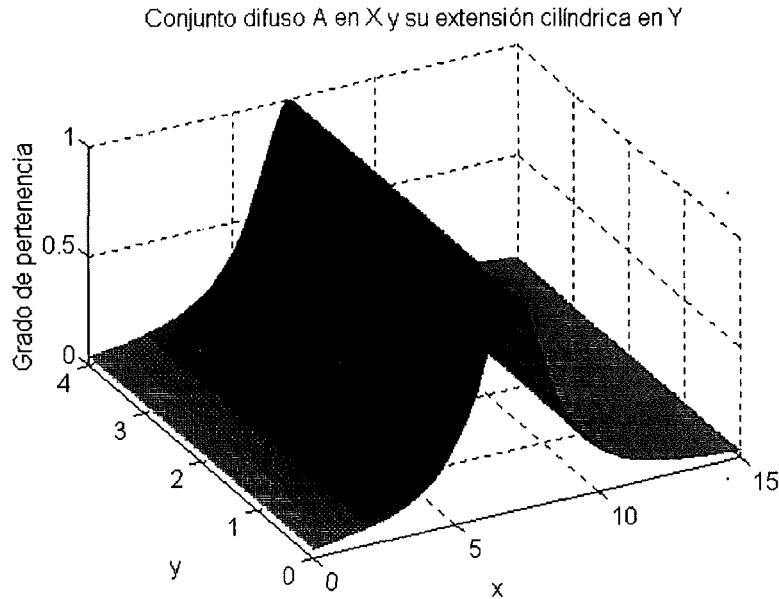


Figura 2.15: Extensión cilíndrica de un conjunto difuso

2.5.7 Composición de relaciones difusas.

Al igual que para muchos otros tópicos de la teoría de conjuntos, se llega a la composición de relaciones difusas por un proceso de generalización del mismo concepto definido en el caso nítido. La mayor parte de la exposición se hará utilizando relaciones binarias, mencionando al final el caso general de composición de relaciones n-arias.

Sean tres conjuntos A, B y C entre los cuales se han definido dos relaciones binarias difusas R y S, la primera de ellas entre A y B y la segunda entre B y C. Estas dos relaciones establecen *de facto* una relación binaria directa entre el conjunto A y el conjunto C, que denotaremos con la letra T. Los elementos de T están caracterizados por

su grado de pertenencia $\mu_T(x,z)$ a la relación, el cual, a su vez, depende de los grados de pertenencia $\mu_R(x,y)$ y $\mu_S(y,z)$ correspondientes a las relaciones difusas intermedias R y S.

De forma similar a lo que acontece para otros operadores difusos, el algoritmo que permite obtener $\mu_T(x,z)$ a partir de $\mu_R(x,y)$ y $\mu_S(y,z)$ no es único. De todos los que se han descrito y aplicado, el que se utiliza con mayor frecuencia responde a las siguientes tres reglas:

1. Si no hay conexión entre un elemento $x \in A$ y un $z \in C$, $\mu_T(x,z) = 0$.

Por ejemplo, en la relación representada en la Figura 2.16, $\mu_T(x_1, z_1) = 0$ porque no existe ningún camino que una x_1 con z_1 .

2. Si sólo existe un camino de conexión entre un elemento $x \in A$ y un $z \in C$, $\mu_T(x,z)$ se toma igual al menor de los grados de pertenencia asociados a cada uno de los tramos, ya que carecería de sentido que la intensidad de la relación entre x y z fuese mayor que la intensidad del tramo más débil. Por ejemplo, en la Figura 2.16 $\mu_T(x_2, z_3)$ vale $\mu_T(x_2, z_3) = \min(0.2, 0.6) = 0.2$.

3. Si existe más de un camino de conexión entre un elemento $x \in A$ y un $z \in C$, $\mu_T(x,z)$ se toma como el mayor de los grados de pertenencia asociados a cada camino, calculados según el punto anterior. Como ejemplo de este caso, en la Figura 2.16 el grado de pertenencia el par (x_1, z_2) a la relación T es

$$\begin{aligned} \mu_T(x_1, z_2) &= \max\left\{\min(\mu_R(x_1, y_1), \mu_S(y_1, z_2)), \min(\mu_R(x_1, y_3), \mu_S(y_3, z_2))\right\} \\ &= \max\left\{\min(0.7, 0.3), \min(0.4, 0.8)\right\} = 0.4 \end{aligned}$$

En general, pues, el grado de pertenencia de cualquier elemento del producto cartesiano $A \times C$ a la relación compuesta $T = R \circ S$ se puede determinar mediante la expresión

$$\mu_T(x, z) = \max_{y \in B} \left\{ \min(\mu_R(x, y), \mu_S(y, z)) \right\} \quad (2.56)$$

algoritmo que, por su forma, recibe el nombre de *composición max-min*. La situación descrita se muestra en la Figura 2.16, que es la versión difusa de la Figura 2.11.

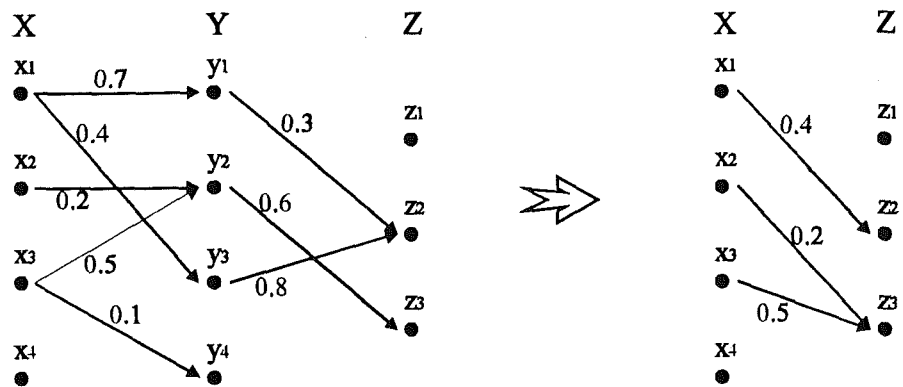


Figura 2.16: Composición max-min de relaciones difusas

También se utiliza con cierta frecuencia la *composición max-producto*, cuya expresión puede verse en la ecuación (2.57), la cual se muestra en la Figura 2.17 para la relación antes descrita.

$$\mu_T(x, z) = \max_{y \in B} \{ \mu_R(x, y) \cdot \mu_S(y, z) \} \quad (2.57)$$

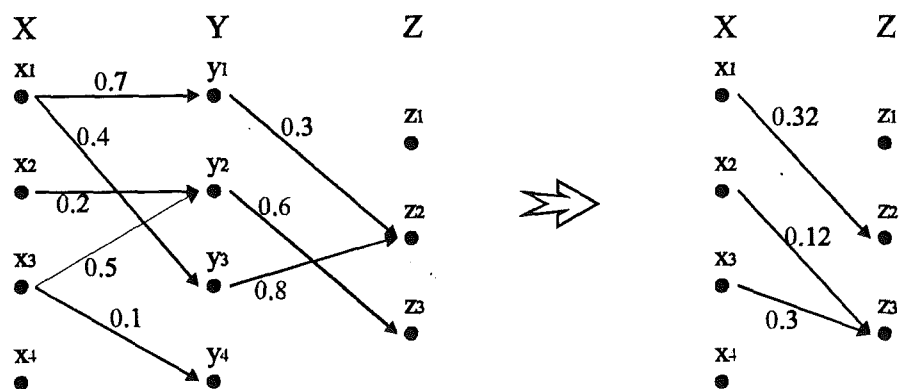


Figura 2.17: Composición max-producto de relaciones difusas

Al contrario de lo que sucedía en la composición de relaciones nítidas, los resultados obtenidos al utilizar la composición max-min o la max-producto son distintos, al ser el producto de dos números del intervalo $[0, 1]$ igual o inferior a su mínimo. Por otra parte, cabe señalar que las ecuaciones (2.56) y (2.57) se expresan para relaciones con un número finito de elementos. En el caso de relaciones con soporte infinito el operador *máximo* debe sustituirse por el *supremo*, que es la mínima cota superior.

Ross [Ross 1995] sugiere una interesante interpretación de la ecuación (2.56): $\mu_{R \circ S}(x, z)$ puede considerarse como la resistencia a la rotura de un conjunto de cuerdas que ligan el elemento x con el elemento z . Existen tantas cuerdas como caminos entre x y z . Cada una de las cuerdas tiene su origen en x y su final en z , pasando por un elemento intermedio y , con lo que se generan dos enlaces por cuerda. La resistencia de cada cuerda es la del enlace más débil, puesto que están en serie. La fuerza de la relación entre x y z es la de la cuerda más resistente de todas las que unen a los citados elementos, puesto que están en paralelo.

Teniendo en cuenta que los operadores *mínimo* y *producto* son casos particulares de t-normas difusas, y que el operador *máximo* lo es de las t-conormas, resulta de interés generalizar la composición de relaciones difusas mediante la expresión

$$\mu_{R \circ S}(x, z) = \mu_T(x, z) = \bigoplus_{y \in Y} \{ \mu_R(x, y) * \mu_S(y, z) \} \quad (2.58)$$

donde \oplus y $*$ son los símbolos genéricos que denotan a las t-conormas y t-normas respectivamente.

De igual modo que para el caso nítido, dado que muchas relaciones con número finito de elementos se expresan mediante sus respectivas matrices de relación, la matriz de la relación [T] puede obtenerse partiendo de las matrices [R] y [S] aplicando la siguiente regla nemotécnica

$$[T] = [R] \otimes [S] \quad (2.59)$$

donde el operador \otimes es similar al producto matricial ordinario, aunque sustituyendo la suma algebraica por una t-conorma \oplus , y el producto algebraico por una t-norma $*$. A modo de ejemplo en la Figura 2.18 se muestra la obtención de la matriz de relación [T] a partir de las matrices [R] y [S] descritas en la Figura 2.16, para el caso de tomar la pareja de operadores máximo y mínimo:

$$\begin{pmatrix} 0 & 0.4 & 0 \\ 0 & 0 & 0.2 \\ 0 & 0 & 0.5 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0.7 & 0 & 0.4 & 0 \\ 0 & 0.2 & 0 & 0 \\ 0 & 0.5 & 0 & 0.1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0.3 & 0 \\ 0 & 0 & 0.6 \\ 0 & 0.8 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Figura 2.18: Composición matricial de relaciones binarias difusas

Para el caso de relaciones n-arias, el proceso que debe seguirse es totalmente análogo al del caso nítido: Sean R y S dos relaciones difusas en $A_1 \times \dots \times A_r$ y $A_s \times \dots \times A_n$ respectivamente, siendo $s \leq r$ (teniendo, por tanto, $r-s+1$ dimensiones compartidas). Su composición $T=R \circ S$ es otra relación difusa en $A_1 \times \dots \times A_{r-1} \times A_{s+1} \times \dots \times A_n$ definida por la función de pertenencia:

$$\begin{aligned} \mu_T(x_1, \dots, x_{r-1}, x_{s+1}, \dots, x_n) &= \\ &= \bigoplus_{(x_r, \dots, x_s) \in A_r \times \dots \times A_s} \{ \mu_R(x_1, \dots, x_r) * \mu_S(x_s, \dots, x_n) \} \end{aligned} \quad (2.60)$$

es decir, se eliminan las dimensiones comunes agregando respecto de las citadas dimensiones mediante el uso del operador \bigoplus , usualmente el máximo.

Otra forma de expresar la composición de relaciones difusas utiliza algunos de los conceptos vistos en anteriores apartados. Según este método, la composición de las relaciones R y S puede obtenerse aplicando el siguiente algoritmo:

1. Encontrar la *extensión cilíndrica* de las dos relaciones R y S al espacio formado por el producto cartesiano total, $(A_1 \times \dots \times A_r) \times (A_s \times \dots \times A_n)$.
2. Calcular la relación *intersección* entre las relaciones extendidas, mediante el uso de una t-norma adecuada.
3. *Proyectar* dicha intersección al subespacio originalmente no común, eliminando de esta forma el común.

Como puede apreciarse, este algoritmo interpreta el operador \oplus como una proyección. Ya se comentó en la sección 2.5.5 que para la operación de proyección debe utilizarse una t-conorma. Sin embargo, cualquier función que realice algún tipo de agregación de los valores adecuados sería un candidato válido para encarnar una proyección. Por ello, algunos autores han propuesto posibles funciones cuya uso puede dar resultados significativamente mejores en ciertas aplicaciones. La utilidad de este punto de vista quedará plenamente justificada en el posterior desarrollo de esta tesis.

Por último, tiene mucho interés detenerse en un punto de especial relevancia en lógica difusa: Dada la estructura de conjunto que posee toda relación difusa, es posible ver un conjunto difuso A en X como una relación difusa en $\xi \times X$, en donde el símbolo ξ representa un conjunto con un único elemento, y éste está relacionado con cada elemento de A con una intensidad dada por $\mu_A(x)$. Visto así, un conjunto difuso puede “activar” una relación difusa R cualquiera en $X \times Y$ mediante la composición entre ambos, obteniéndose una relación en $\xi \times Y$. Matemáticamente esto se expresa de la forma

$$\mu_{A \circ R}(\xi, y) = \bigoplus_{x \in X} \{ \mu_A(\xi, x) * \mu_R(x, y) \} \quad (2.61)$$

cuyo resultado puede interpretarse a su vez como conjunto difuso B en Y , obteniéndose en definitiva

$$\mu_{A \circ R}(y) = \mu_B(y) = \bigoplus_{x \in X} \{ \mu_A(x) * \mu_R(x, y) \} \quad (2.62)$$

que puede reescribirse en forma sintética como

$$B = A \circ R \quad (2.63)$$

La ecuación (2.63) es de una importancia capital, puesto que es la base de dos técnicas clave en teoría de conjuntos difusos: Por un lado generaliza el concepto de “encontrar la imagen de un número a través de una función” al permitir encontrar la imagen de un número difuso a través de una relación difusa. Por otro lado, es la expresión de la *Regla Composicional de Inferencia*, que es el mecanismo principal de razonamiento en lógica difusa. Cada uno de estos conceptos se trata en secciones posteriores.

Como ejemplo de lo que acaba de exponerse, en la Figura 2.19 se muestra gráficamente el proceso de encontrar la imagen de un conjunto difuso $A = \text{"aproximadamente 9"}$ definido en X , a través de la relación difusa expresada por "y es aproximadamente raíz cuadrada de x " cuya función de pertenencia se ha obtenido de forma heurística siguiendo un razonamiento similar al expuesto para la relación "y es aproximadamente x^2 " descrita en la sección 2.5.4.

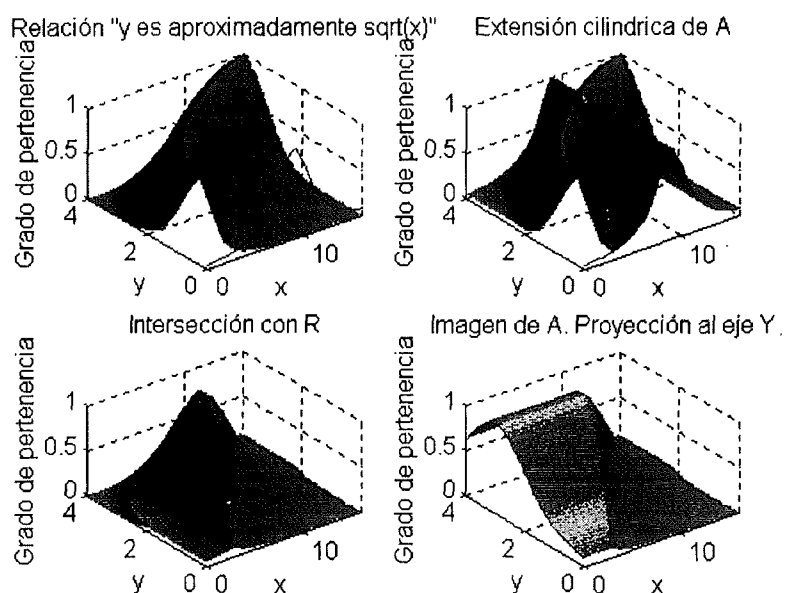


Figura 2.19: Proceso de composición $B = A \circ R$

El conjunto imagen es, como cabría esperar, un número difuso B en Y , el cual, por su forma, podría corresponder a el número difuso "aproximadamente 3". Cabe señalar por último que en el cálculo de la composición se han utilizado el máximo y el mínimo como t-conorma y t-norma respectivamente.

2.6 Funciones.

2.6.1 Función clásica. Generalización del concepto.

El concepto de función es uno de los más familiares de la matemática, por lo que aquí se introduce con el único propósito de servir de base a la discusión de la sección siguiente sobre funciones cuyos argumentos son conjuntos difusos. En la acepción clásica del término, una función f definida sobre el universo X y con imágenes en el universo Y es una *aplicación* entre X e Y , es decir, una **relación nítida** entre X e Y en la

que todo elemento del conjunto origen X está relacionado con un único elemento del conjunto imagen Y :

$$f: X \rightarrow Y \quad (2.64)$$

Los elementos del conjunto origen son los *argumentos* de la función f . Cuando el conjunto origen está formado por el producto cartesiano de varios conjuntos, hablamos de función de varias variables. En el caso de que el conjunto llegada sea el producto cartesiano de varios conjuntos, estaremos ante una función vectorial. Terminológicamente usamos el nombre general de *correspondencia* para denotar a las aplicaciones establecidas entre cualquier tipo de conjuntos, reservando el nombre de *función* para el caso en que los elementos de los universos origen y destino sean numéricos.

Quizá la operación que se efectúa más frecuentemente cuando se manipulan funciones es la de determinar la imagen de un cierto número a través de la función. Si pretendemos generalizar esta operación mediante la introducción de conceptos difusos podemos encontrarnos con 9 situaciones distintas, según como tomemos el número o la propia función. Estas situaciones se describen en la Tabla 2-4:

Tabla 2-4: Distintas posibilidades de generalización del concepto de función

<i>Imagen de un</i>	1) número nítido	<i>a través de una</i>	a) función
	2) conjunto nítido		b) relación nítida
	3) conjunto difuso		c) relación difusa

El caso 1a) es el habitual. En él, encontrar la imagen de un cierto $x_0 \in X$ por una función $f(x)$ no presenta en este caso mayores problemas que el de evaluar la función en ese punto, o sea encontrar $y_0 = f(x_0)$. Gráficamente el proceso se ilustra en la Figura 2.20.

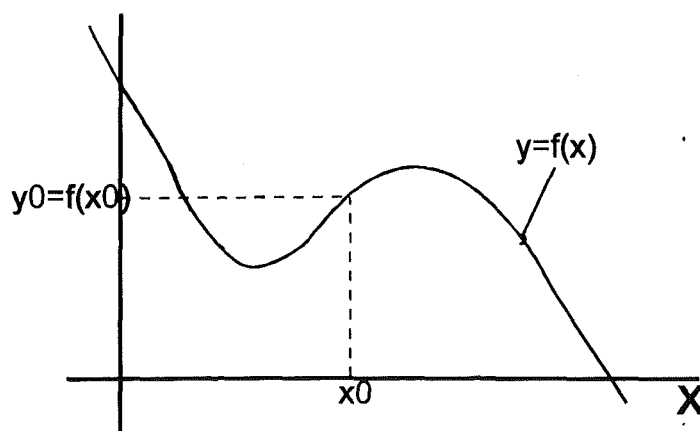


Figura 2.20: Imagen de un punto por la función $f(x)$

Además del 1a), para nosotros tendrán un interés especial los casos 3a) y 3c). El primero de ellos estudia cómo debe encontrarse la imagen de un conjunto difuso a través de una función, dando lugar al denominado *principio de extensión*, que se trata en la sección 2.6.2. El segundo estudia el problema de determinar la imagen de un conjunto difuso a través de una relación difusa cualquiera, dando lugar a la *regla composicional de inferencia*, que se trata en profundidad en la sección 3.4, para la que se ha dado su expresión compacta en la ecuación (2.63).

Los 6 casos restantes, aún no mencionados, carecen de mayor interés para nuestros objetivos, y pueden considerarse meros pasos intermedios. Además, puesto que el caso 3c) es el más general, el resto puede obtenerse mediante particularizaciones convenientes de las expresiones de la regla composicional de inferencia dadas en (2.56)-(2.57).

Puesto que, como se ha visto, una función no es más que un caso particular de relación nítida, y ésta es a su vez un caso particular de relación difusa, es posible ver toda función como relación difusa, y representarla mediante su diagrama tridimensional de función de pertenencia. Tal función será siempre cero excepto en los pares $(x, f(x))$, en los que tomará el valor 1. La Figura 2.21 muestra dicha gráfica para la función $y=\sin(x)$ en el intervalo $x \in [0, 10]$:

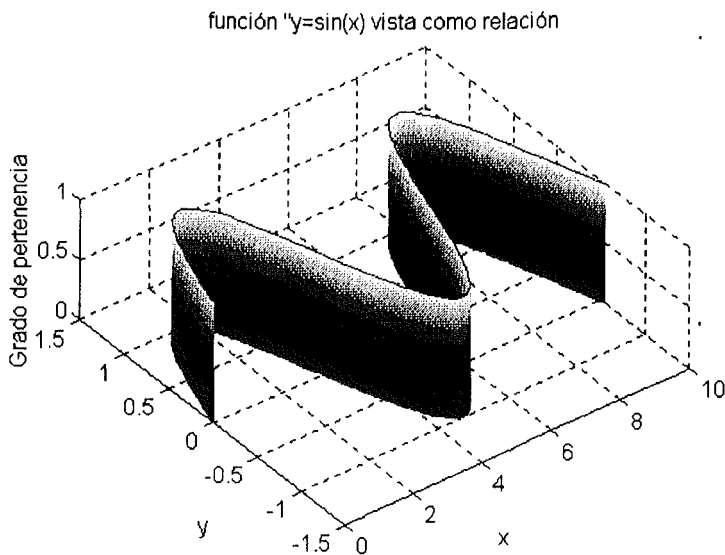


Figura 2.21: Función vista como relación

Tal punto de vista va a ser de utilidad en las secciones venideras.

2.6.2 Función de un conjunto difuso. Principio de extensión.

Principio de extensión para funciones unidimensionales

El principio de extensión es una útil herramienta para generalizar conceptos matemáticos nítidos al ámbito de los conjuntos difusos. Aparte de su importancia dentro de esta teoría, este principio tiene una especial relevancia en la presente tesis, pues es una de las bases teóricas en la que se sustenta el trabajo. Brevemente, consiste en el principio matemático que nos permite extender el concepto clásico de función para que, además de números, pueda aceptar también argumentos que sean conjuntos difusos.

Una situación intermedia entre encontrar la imagen de un punto y la de un conjunto difuso consiste en determinar la imagen de un *intervalo* de valores de X, puesto que un intervalo no es más que un conjunto nítido en X. Esta situación corresponde al caso 2a) de la Tabla 2-4, y está bien estudiada en la teoría clásica de conjuntos: básicamente, la imagen de un intervalo en X a través de una función f es otro intervalo en Y, como muestra la gráfica de la Figura 2.22:

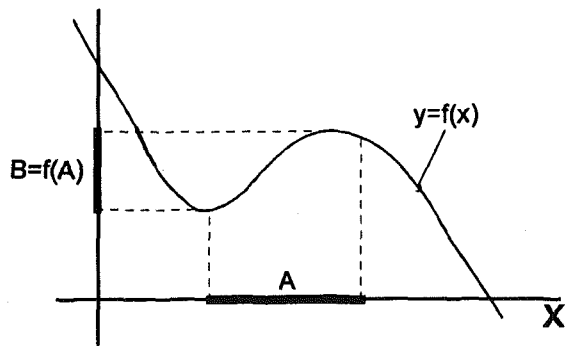


Figura 2.22: Imagen del intervalo A a través de la función $f(x)$

Formalmente, para un conjunto nítido A definido en X , su imagen B definido en Y a través de la función f es el conjunto de todos los elementos $y \in Y$ tales que sean imagen por f de algún $x \in X$. Matemáticamente el procedimiento se sintetiza en la expresión

$$B = f(A) = \{y \mid y = f(x) \quad \forall x \in A\} \quad (2.65)$$

que puede expresarse en una forma más compacta haciendo uso de la función característica, en la forma

$$\chi_B(y) = \chi_{f(A)}(y) = \max_{x=f^{-1}(y)} \{\chi_A(x)\} \quad (2.66)$$

que nos indica que el elemento genérico $y \in Y$ pertenece al intervalo imagen B si tiene alguna antiimagen por f en el intervalo A . El uso de la función *máximo* es una forma conveniente de tratar los casos en que un mismo elemento $y \in Y$ tenga varias antiimágenes en el universo X . En ese caso, basta con que una de las antiimágenes pertenezca al intervalo origen A para que el elemento y pertenezca a B . Las funciones que resultan de esta formulación son de hecho, funciones cuyos argumentos son elementos del conjunto potencia¹⁰ de X , es decir

$$f: \wp(X) \rightarrow Y \quad (2.67)$$

La extensión del concepto anterior para el caso de que el argumento sea un conjunto difuso es inmediata, simplemente haciendo que la función tome argumentos en el

conjunto potencia difusa¹¹ de X , $\mathcal{F}(X)$, en lugar de en $\wp(X)$. En este caso, el conjunto A es difuso, con función de pertenencia $\mu_A(x)$, y el conjunto imagen resulta ser también un conjunto difuso B cuyos elementos y son las imágenes de x por f , con un grado de pertenencia $\mu_B(y)$ igual a

$$\mu_B(y) = \mu_{f(A)}(y) = \max_{x=f^{-1}(y)} \{ \mu_A(x) \} \quad (2.68)$$

en donde debe substituirse el operador *máximo* por el *supremo* en el caso de que se trabaje con universos continuos.

La ecuación (2.68) indica que si varios elementos x_1, \dots, x_n de A tienen la misma imagen y , el grado de pertenencia de y a B corresponde al mayor de los grados de pertenencia de x_1, \dots, x_n a A . Como se ha visto en la descripción de otros operadores, la función *máximo* no es la única elección posible, aunque es con mucho la más extendida. En síntesis, el proceso de encontrar la imagen de un conjunto difuso es similar al de determinar la imagen de un intervalo, completado en esta caso con la asignación de un grado de pertenencia a cada elemento, de la forma mostrada en la Figura 2.23.

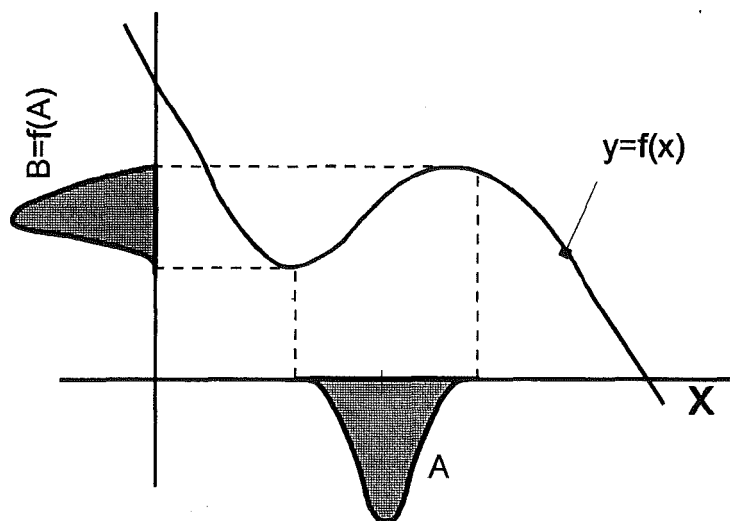


Figura 2.23: Imagen del conjunto difuso A a través de la función $f(x)$

¹⁰ Véase sección 2.3.1

¹¹ Véase sección 2.3.2.

Dado que una función no es más que cierto tipo de relación nítida, y que todo conjunto difuso puede verse como relación, una forma conveniente de expresar el principio de extensión hace uso de la composición max-min o max-prod, según se muestra en la ecuación (2.69)

$$(\mu_B(y_1) \cdots \mu_B(y_m)) = (\mu_A(x_1) \cdots \mu_A(x_n)) \otimes \begin{pmatrix} \chi_F(x_1, y_1) & \cdots & \chi_F(x_1, y_m) \\ \vdots & \ddots & \vdots \\ \chi_F(x_n, y_1) & \cdots & \chi_F(x_n, y_m) \end{pmatrix} \quad (2.69)$$

o, más escuetamente

$$B = A \circ F$$

lo que muestra que el principio de extensión es en definitiva un caso particular de la composición de relaciones. A modo de ejemplo, sea el universo $X = \{0, 1, 2, 3, 4, 5\}$ en el que se ha definido el conjunto difuso $A = \{0, 0.2, 0.7, 1, 0.7, 0.2\}$. La imagen de A a través de la función $f(x) = x^2 - 4x + 5$ es un conjunto difuso B cuyo soporte Y lo constituyen todos los elementos imagen de X por f , a saber, $Y = \{1, 2, 5, 10\}$, y que vale $B = \{0.7, 1, 0.7, 0.2\}$, según muestran la ecuación (2.70) y la Figura 2.24.

$$(0.7 \ 1 \ 0.7 \ 0.2) = (0 \ 0.2 \ 0.7 \ 1 \ 0.7 \ 0.2) \otimes \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.70)$$

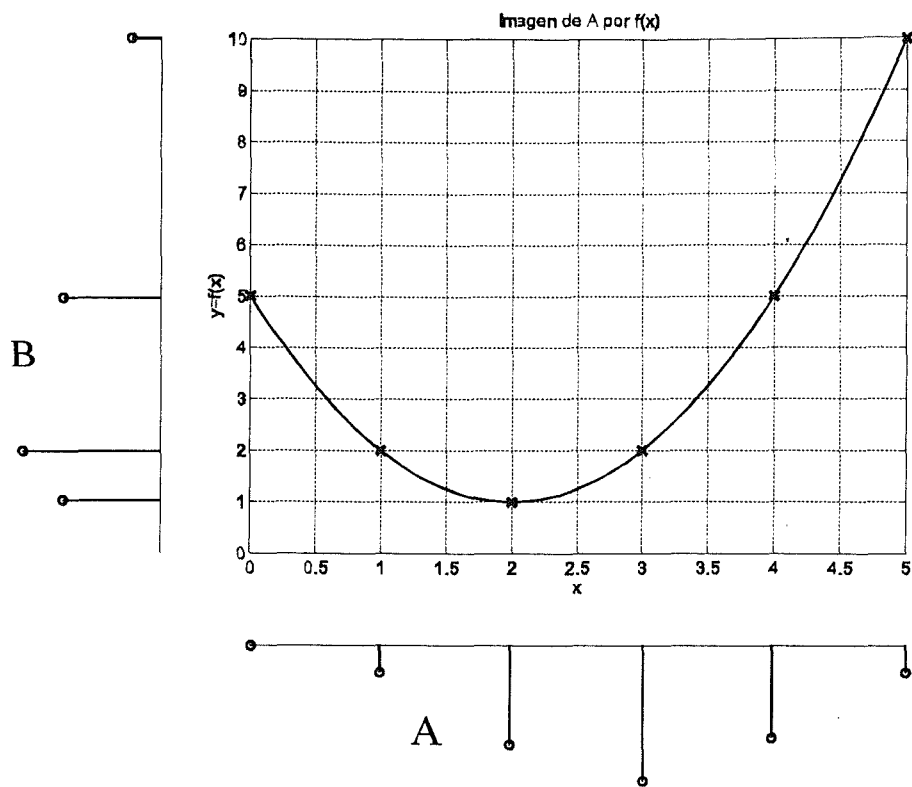


Figura 2.24: Aplicación del principio de extensión

Principio de extensión para funciones de varias variables

Cuando la función bajo consideración es de varias variables, su extensión al manejo de conjuntos difusos requiere generalizar previamente el concepto de producto cartesiano.

Sea X el producto cartesiano de n universos $X = X_1 \times X_2 \times \dots \times X_n$, y sean A_1, \dots, A_n n conjuntos difusos definidos respectivamente en X_1, X_2, \dots, X_n . El *producto Cartesiano difuso* $A_1 \times \dots \times A_n$ se define como un subconjunto difuso de X , siendo el grado de pertenencia asociado a cada n -pla de $A_1 \times \dots \times A_n$ el dado por la expresión

$$\mu_{A_1 \times \dots \times A_n}(x_1, \dots, x_n) = \min\{\mu_{A_1}(x_1), \dots, \mu_{A_n}(x_n)\} \quad (2.71)$$

Con esta definición, puede observarse fácilmente que el producto cartesiano difuso no es más que un tipo particular de relación difusa en X , obtenida mediante el uso de la función mínimo.

Se trata ahora de establecer qué se entiende por función de varios conjuntos difusos. Para ello, si f es una función clásica definida en $X_1 \times X_2 \times \dots \times X_n$ con imágenes en el universo Y de forma que $y = f(x_1, \dots, x_n)$, generalizando el principio de extensión dado en (2.68) e incorporándole el resultado (2.71) se obtiene la expresión

$$\mu_B(y) = \mu_{f(A_1 \times \dots \times A_n)}(y) = \max_{(x_1, \dots, x_n) = f^{-1}(y)} \min\{\mu_{A_1}(x_1), \dots, \mu_{A_n}(x_n)\} \quad (2.72)$$

que puede verse también como un caso especial de composición de relaciones difusas, tal y como ocurría para el principio de extensión de funciones de una variable.

2.7 Teoría de la posibilidad

2.7.1 Introducción.

La teoría de la posibilidad se enmarca en el contexto de la presente tesis al constituir la base matemática sobre la que se define el modelo temporal difuso utilizado más adelante para expresar nuestros conceptos temporales. Por ello, a continuación se exponen aquellos aspectos más relevantes de esta teoría, que serán de utilidad en los desarrollos posteriores.

Fué Zadeh quien, en 1978, introdujo las primeras ideas de esta teoría, relacionándola estrechamente con la teoría de conjuntos difusos [Zadeh 1978]. Su motivación principal fué la constatación de que, en cualquier transmisión de información, si el interés principal de la misma reside en su *significado*, más que en su medida, entonces el marco correcto para analizarla no era la teoría de la probabilidad, sino otro enfoque distinto que denominó *teoría de la posibilidad*, término acuñado unos años antes por Gaines y Kohout, según él mismo describe. Esta teoría, además, debía poder tratar correctamente aquellas formas de incertidumbre que no eran modelables mediante conjuntos difusos.

Parafraseando a Klir y Yuan [Klir 1995], considérese la situación en que los miembros de un jurado deben decidir sobre la culpabilidad o inocencia de un acusado de homicidio. Evidentemente en este caso no procede describir los términos “inocente” o “culpable” como conjuntos difusos y emitir un veredicto del tipo “es culpable en un grado 0.8 e inocente en un grado 0.2”. Por el contrario, el jurado se enfrenta al problema

de *clasificar* al acusado en el conjunto **nítido** de las personas culpables o en el de las inocentes, y hacerlo basándose en la información imprecisa que proporcionan las pruebas. Se llega entonces mentalmente a valorar de forma subjetiva la posibilidad de que el acusado sea culpable o inocente asignando un valor o *medida* a cada uno de estos conjuntos que refleje esa posibilidad. Nótese que en el caso de no tener dudas respecto al veredicto, se ubicará al acusado tranquilamente en el conjunto que le corresponda. Esto es muy distinto de lo que ocurre en la teoría de conjuntos difusos, en la que un conocimiento perfecto de, por ejemplo, el peso de una persona, puede ser perfectamente compatible con una pertenencia parcial y simultánea a más de un conjunto, como “obesos” y “normales”.

Este problema podría intentar abordarse bajo una formulación probabilística, pero para hacerlo debe resolverse previamente el modo de cuantificar unas evidencias que en muchos casos son subjetivas o están expresadas en forma lingüística, y de las cuales muy raramente se dispone de información frecuencial.

Como se discutirá en la sección 2.7.3, la teoría de la posibilidad, al igual que la de la probabilidad, aunque distintas, presentan numerosas semejanzas. El motivo de ello es que ambas se fundamentan en la rama de la matemática conocida como *teoría de la medida*, un campo amplio e interesante del cual también se nutren otras disciplinas como el cálculo integral. No es el objetivo de estas líneas el exponer, ni siquiera brevemente, la formulación general de esta teoría, por lo que a continuación se trata directamente el caso particular de las medidas de posibilidad.

2.7.2 Medida y distribución de posibilidad.

Una medida de posibilidad Π puede definirse verbalmente como una función que toma como argumento cualquier subconjunto **nítido** de un universo X y devuelve un valor entre 0 y 1 que expresa hasta qué punto es posible el “evento” que asociamos al citado subconjunto. Es, por tanto, una función definida sobre el conjunto potencia¹² de X :

¹² Para la discusión del conjunto potencia, refiérase el lector a la sección 2.3.1.

$$\Pi: \wp(X) \rightarrow [0,1] \quad (2.73)$$

debe, además, cumplir las dos propiedades siguientes:

$$1. \Pi(\emptyset) = 0; \Pi(X) = 1 \quad (\text{condiciones de frontera}) \quad (2.74)$$

$$2. \forall A_i \in \wp(X), \quad \Pi\left(\bigcup_{i=1,2,\dots} A_i\right) = \sup_{i=1,2,\dots} \Pi(A_i) \quad (2.75)$$

para el caso de sólo dos conjuntos, la segunda propiedad se reduce a

$$\Pi(A \cup B) = \max(\Pi(A), \Pi(B)) \quad (2.76)$$

que expresa el hecho de que la realización de uno de los dos eventos asociados a los subconjuntos A o B, tomados indistintamente, está afectado del mismo coeficiente de posibilidad que la realización del evento más posible. La primera propiedad nos indica que el evento nulo es imposible y que el conjunto formado por todos los eventos es totalmente posible (se cumple con certeza absoluta).

La diferencia básica con la probabilidad, como se verá, la constituye el segundo supuesto, que para el caso probabilístico debe de sustituirse por la imposición más fuerte de *aditividad*.

De la misma forma que toda medida de probabilidad suele caracterizarse mediante su correspondiente función de distribución o densidad, toda medida de posibilidad queda completamente definida si se especifica el coeficiente de posibilidad asociado a cada posible subconjunto de X (miembro de $\wp(X)$). Más simplemente, puesto que la ecuación (2.75) permite encontrar la posibilidad asociada a cualquier conjunto a partir de la asociada a los subconjuntos que lo forman, la medida de posibilidad quedará completamente determinada especificando la posibilidad asociada a cada *elemento* de X. Definimos de este modo una función denominada *distribución de posibilidad* π , que toma como argumentos los elementos de X y devuelve el grado de posibilidad asociado a ellos como un número entre 0 y 1:

$$\pi: X \rightarrow [0,1] \quad (2.77)$$

Medida y distribución de posibilidad están evidentemente relacionadas. Para todo subconjunto A de X , la posibilidad asociada a él puede encontrarse a partir de la distribución mediante

$$\Pi(A) = \sup_{x \in A} \pi(x) \quad (2.78)$$

situación que se refleja en la Figura 2.25

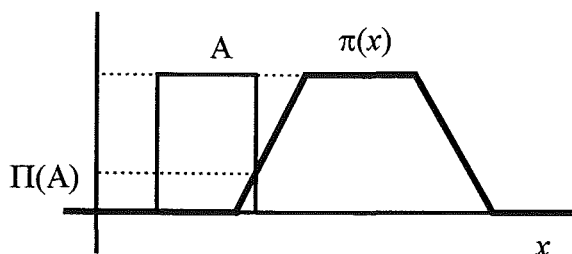


Figura 2.25: Medida de posibilidad a partir de la distribución

Similarmente, el valor que toma la distribución de posibilidad para cualquier elemento x puede encontrarse como la medida de posibilidad asociada al conjunto formado por ese único elemento, es decir

$$\forall x \in X, \quad \pi(x) = \Pi(\{x\}) \quad (2.79)$$

De forma equivalente a lo que sucede en probabilidad, también es posible aquí definir distribuciones de posibilidad conjuntas, condicionales, marginales, etcétera, que se salen del ámbito de esta breve exposición.

Las medidas de posibilidad que acaban de describirse suministran información sobre la ocurrencia de un cierto evento asociado al subconjunto A de X . Sin embargo, esta medida no basta por sí sola para describir toda la incertidumbre existente sobre este evento. Por ejemplo, el hecho de conocer que un evento A es totalmente posible ($\Pi(A) = 1$) no impide que el evento contrario también lo sea ($\Pi(\bar{A}) = 1$), lo que indicará una total incertidumbre sobre la producción de ese evento. Puede suceder también que el evento contrario sea imposible ($\Pi(\bar{A}) = 0$), lo que pondrá en evidencia que el evento asociado a A , además de posible, se producirá con total seguridad.

Para complementar la medida de posibilidad, pues, es necesario definir otra medida que nos indique el grado en que un evento en cuestión es *cierto*. Este otro tipo de medidas reciben el nombre de *medidas de necesidad*, $N(A)$, y son duales de las de posibilidad en el sentido de que la necesidad de un evento equivale a la posibilidad del evento contrario, es decir

$$N(A) = 1 - \Pi(\bar{A}) \quad (2.80)$$

Aparte de la definición que supone esta última ecuación, también es posible definir las medidas de necesidad partiendo de un esquema axiomático parecido al de las ecuaciones (2.74) y (2.75).

La existencia de la ecuación (2.80) hace que en la práctica se trabaje únicamente con distribuciones de posibilidad y no se defina un concepto equivalente para la necesidad.

La similitud existente entre las formulaciones de la posibilidad y la teoría de conjuntos difusos, sobre todo a lo que se refiere a las expresiones (2.74) y (2.75), suscita rápidamente la cuestión de cuáles son sus puntos de contacto. Esta es una cuestión sutil cuya respuesta puede darse en los siguientes términos: *Ambas teorías están conectadas a través de la equivalencia funcional entre una función de distribución de posibilidad y un conjunto difuso*. En efecto, una cierta distribución de posibilidad puede verse como una restricción difusa que actúa como una cota elástica sobre los valores que pueden ser asignados a una variable. En este supuesto, los valores mencionados pueden considerarse conjuntos difusos y, por lo tanto,

$$\pi(x) = \mu_A(x) \quad (2.81)$$

Esto no implica que una distribución de posibilidad y un conjunto difuso tengan el mismo significado. Que esto no es así ya se ha comentado en la sección 2.7.1. Esta última expresión, sin embargo, permite tratar mediante las mismas herramientas matemáticas ambos tipos de estructuras.

Otra manera de ver esta relación es considerando lo que sucede cuando se pretende encontrar la medida de posibilidad asociada a un conjunto difuso A en lugar de nítido. Esta situación se refleja en la Figura 2.26

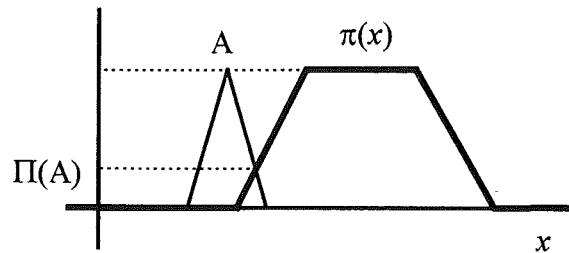


Figura 2.26: Posibilidad de un conjunto difuso

En este caso la ecuación (2.78) debe modificarse haciendo

$$\Pi(A) = \sup_{x \in A} \min\{\mu_A(x), \pi(x)\} \quad (2.82)$$

ecuación que expresa la posibilidad de A restringida a π como el grado de compatibilidad entre estas dos funciones.

2.7.3 Relación entre posibilidad y probabilidad.

La cultura “probabilística” que impregna desde hace décadas los currículos de las más diversas disciplinas científicas provoca la sensación, al que se acerca por primera vez a la teoría de la posibilidad, de que ésta es una forma poco rigurosa de probabilidad. Sin embargo hoy en día está bien establecido que tanto la teoría de la posibilidad como la de la probabilidad constituyen casos particulares de una teoría más general, llamada *teoría de la evidencia*. En el marco de esta última teoría es posible expresar matemáticamente las diferencias y parecidos entre las dos primeras en base a los llamados *cuerpos de evidencia*. Dado que la teoría de la evidencia no se trata en ningún momento en esta tesis, el lector interesado en profundizar sobre esta interpretación deberá remitirse a la bibliografía especializada sobre el tema¹³.

¹³ A tal efecto pueden consultarse las obras de Klir y Yuan [Klir 1995] pág. 177, Ross [Ross 1995] pág. 556, o Bouchon-Meunier [Bouchon-Meunier 1993] pág. 51.

Lo que se hace a continuación es exponer mediante una serie de ejemplos algunas de las consecuencias de las diferencias matemáticas entre probabilidad y posibilidad, en aras de distinguir ambas a nivel funcional.

El primer ejemplo es el siguiente: supóngase que cierto profesor imparte clase en un grupo de 80 alumnos, y desea evaluar la probabilidad y la posibilidad de que un cierto número de ellos escojan la carrera docente en un futuro. Para encontrar la distribución de posibilidad este profesor debería asignar a cada número un valor entre 0 y 1 según su experiencia. Así, dado que él cree posible que al menos 5 alumnos lo hagan, asignaría un grado 1 a los valores 1, 2, 3, 4 y 5. A partir de ahí asignaría valores decrecientes hasta llegar a 10, número que él considera imposible que se produzca y al que por tanto asigna posibilidad cero. Se obtiene así una distribución de la forma

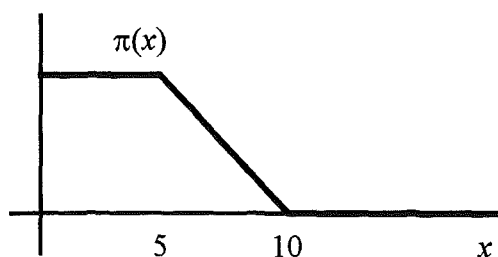


Figura 2.27: Distribución de posibilidad para el ejemplo propuesto

En cambio, desde un punto de vista probabilístico, la función densidad de probabilidad tendría un pico en el valor considerado más probable. Además, la suma de todos los grados de probabilidad debería ser uno:

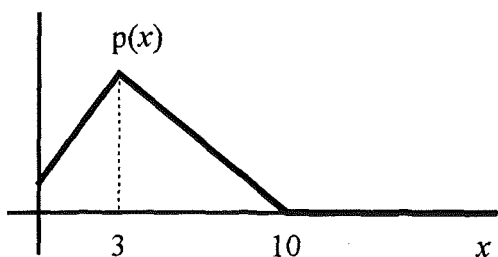


Figura 2.28: Distribución de probabilidad para el ejemplo propuesto

Otro punto en que probabilidad y posibilidad discrepan es en la representación de la *ignorancia total*: Considérese el experimento de lanzar un dado del que no se conoce ninguna información. Ante esta situación la teoría de la posibilidad asume que todos los

valores del 1 al 6 son posibles y asigna por tanto un valor de 1 a su posibilidad. La misma situación se trata desde el punto de vista probabilístico asignando igual probabilidad a cada evento e igual a $1/6$. Sin embargo puede verse fácilmente que esta última modelación no es demasiado adecuada, pues presume ciertas características del dado sin conocer nada de él, como, por ejemplo, si el dado está cargado o no.

Si se da el caso de conocer que el dado está trucado, el probabilista modificaría su distribución para reflejar este hecho, seguramente asignando más probabilidad al valor trucado. Por el contrario, el posibilista no debería modificar su distribución, porque a pesar del trucaje todos los valores siguen siendo igualmente *posibles*.

De estos ejemplos puede inferirse que la posibilidad asigna sus valores sin tener demasiado en cuenta la *frecuencia* en que un evento se produce, sino más bien mediante impresiones subjetivas. Puede perfectamente construirse una distribución de posibilidad a partir de una única observación de un fenómeno. En cambio, para la probabilidad la información frecuencial es esencial para lograr buenos modelos.