Apartado 7.3: Teoría de conjuntos difusos y lógica difusa

1 Introducción

La lógica difusa[1], como su nombre indica, es una lógica alternativa a la lógica clásica que pretende introducir un grado de vaguedad en las cosas que califica. En el mundo real existe mucho conocimiento no-perfecto, es decir, conocimiento vago, impreciso, incierto, ambiguo, inexacto, o probabilístico por naturaleza. El razonamiento y pensamiento humano frecuentemente conlleva información de este tipo, probablemente originada de la inexactitud inherente de los conceptos humanos y del razonamiento basado en experiencias similares pero no idénticas a experiencias anteriores.

El problema principal surge de la poca capacidad de expresión de la lógica clásica. Supongamos por ejemplo que tenemos un conjunto de personas que intentamos agrupar según su altura, clasificándolas en *altas* o *bajas*. La solución que presenta la lógica clásica es definir un umbral de pertenencia (por ejemplo, un valor que todo el mundo considera que de ser alcanzado o superado, la persona en cuestión puede llamarse *alta*). Si dicho umbral es 1.80, todas las personas que midan 1.80 o más serán *altas*, mientras que las otras serán *bajas*. Según esta manera de pensar, alguien que mida 1.79 será tratado igual que otro que mida 1.50, ya que ambos han merecido el calificativo de *bajas*. Sin embargo, si dispusiéramos de una herramienta para caracterizar las alturas de forma que las transiciones fueran suaves, estaríamos reproduciendo la realidad mucho más fielmente.

Asimismo, no hay un valor cuantitativo que defina el término joven. Para alguna gente, 25 años es joven, mientras que para otros, 35 es joven. Incluso el concepto puede ser relativo al contexto. Un presidente de gobierno o de 35 años es joven, mientras que un futbolista no lo es. Hay sin embargo cosas que están claras: una persona de 1 año es joven, mientras que una de 100 años no lo es. Pero una persona de 35 años tiene algunas posibilidades de ser joven (que normalmente dependen del contexto). Para representar este hecho, definiremos el conjunto joven de modo que cada uno de sus elementos pertenezca a él con cierto grado (posibilidad). De un modo más formal, un conjunto difuso A se caracteriza por una función de pertenencia:

$$\mu_A : U \rightarrow [0,1]$$

que asocia a cada elemento x de U un número $\mu_A(x)$ del intervalo [0,1], que representa el grado de pertenencia de x al conjunto difuso A. A U se le llama *universo de discurso*. Por ejemplo, el término difuso *joven* puede definirse mediante el conjunto difuso siguiente:

Edad	Grado de Pertenencia
≤25	1.0
30	0.8
35	0.6
40	0.4
45	0.2

[1] A la hora de traducir el término inglés fuzzy, se barajaron principalmente dos alternativas: borroso y difuso. Aunque en alguna bibliografía se habla aún de Lógica Borrosa o Teoría de los conjuntos borrosos, se utiliza más el término difuso.

≥50	0

Es decir, la función de pertenencia del conjunto difuso *joven* viene dada por:

$\mu_A(x) = 1$ si $x \leq 25$, $\mu_A(30) = 0.8$, …., $\mu_A(x) = 0$ si $x \geq 50$.

Que podemos representar en la siguiente gráfica:

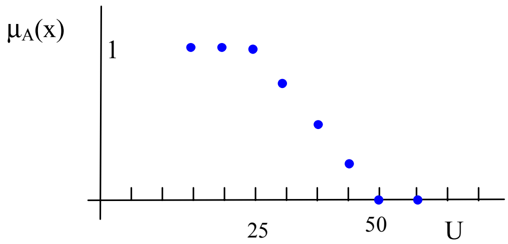


Figura 1. Función de pertenencia del conjunto difuso joven

Si el universo de discurso es continuo, tendremos funciones de pertenencia continuas:

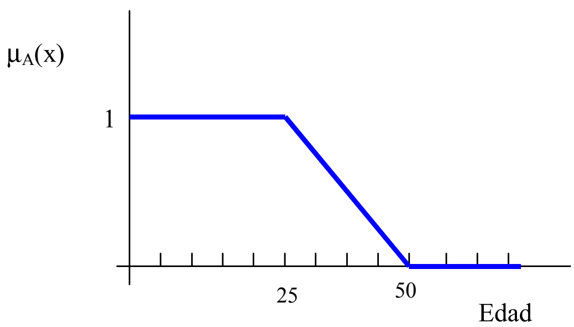


Figura 2. Función de pertenencia de joven si U es continuo

En general, si una función de pertenencia se da especificando los valores correspondientes a un conjunto discreto de elementos del universo de discurso, el valor asociado al resto de los elementos se obtiene por interpolación (utilizando la ecuación de la recta que une los dos puntos[2]).

El origen del interés actual por la teoría de conjuntos difusos se debe a un artículo publicado por Lofti Zadeh en 1.965. En la actualidad es un campo de investigación muy importante, tanto por sus implicaciones matemáticas o teóricas como por sus aplicaciones prácticas. Prueba de esta importancia es el gran número de revistas internacionales (Fuzzy Sets and Systems, IEEE Transactions on Fuzzy Systems..) congresos (FUZZ-IEEE, IPMU, EUSFLAT, ESTYLF...) y libros (Kruse, 1994), (McNeill, 1994), (Mohammd, 1993), (Pedrycz, 1998) dedicados al tema.

¿En qué situaciones es útil aplicar la lógica difusa?

- En procesos complejos, si no existe un modelo de solución sencillo.
- Cuando haya que introducir la experiencia de un operador “experto” que se base en conceptos imprecisos.
- Cuando ciertas partes del sistema a controlar son desconocidas y no pueden medirse de forma fiable (con errores posibles).
- Cuando el ajuste de una variable puede producir el desajuste de otras.

2 Así se hace en FuzzyCLIPS

- En general, cuando se quieran representar y operar con conceptos que tengan imprecisión o incertidumbre

Algunas aplicaciones importantes de la lógica difusa son:

- *Control de sistemas*: Control de tráfico, control de vehículos (helicópteros...), control de compuertas en plantas hidroeléctricas, centrales térmicas, control en máquinas lavadoras, control de metros (mejora de su conducción, precisión en las paradas y ahorro de energía), ascensores...
- *Predicción y optimización*: Predicción de terremotos, optimizar horarios...
- *Reconocimiento de patrones y Visión por ordenador*: Seguimiento de objetos con cámara, reconocimiento de escritura manuscrita, reconocimiento de objetos, compensación de vibraciones en la cámara, sistemas de enfoque automático...
- *Sistemas de información o conocimiento*: Bases de datos, sistemas expertos...

2 Teoría de conjuntos difusos

2.1 Teoría de conjuntos clásica (conjuntos nítidos)

Los Conjuntos Clásicos (nítidos en la terminología de lógica difusa) surgen de forma natural por la necesidad del ser humano de clasificar objetos y conceptos. Por ejemplo, si pensamos en los productos de alimentación, podemos hacer varios conjuntos:

Frutas: Manzana, Pera, plátano, etc.

Verduras: Calabacín, Espinaca, ...

Carnes: ...

Pescados:...

...

Los conjuntos nítidos pueden definirse de varias formas:

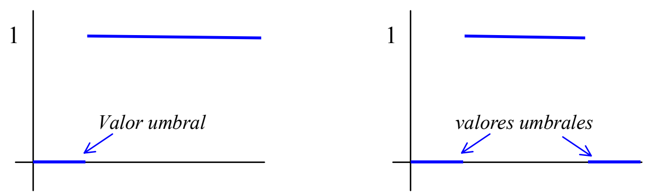
- *Mediante un listado de sus elementos*: *Frutas* = {manzana, pera,...}
- *Mediante una función de pertenencia* μ que toma valores 0 o 1 definida sobre el universo de discurso U (todos los elementos que pueden o no pertenecer al conjunto):

  Ejemplo: sea U el conjunto de todos los alimentos. Entonces *Frutas* es un conjunto tal que μ(manzana)=1, μ(pargo)=0, etc.

  De este modo, para definir un conjunto nítido A podemos utilizar la función de pertenencia dada por:

$$\mu_A(x) = \begin{cases} 0 & \text{si } x \notin A \\ 1 & \text{si } x \in A \end{cases}$$

  Es decir, una función tipo escalón centrada en el valor/valores umbral/umbrales de decisión.



Si se utilizan funciones de pertenencia, la forma de representar el vacío y el conjunto universo será:

- El vacío ∅ es un conjunto tal que para todo x de U, μ(x)=0

- El conjunto universo es tal para todo x de U, $\mu(x)=1$

- *Dando una característica que defina sus elementos*. Esto se puede hacer de varias formas:

  - Dando directamente la definición:

    Fruto = Producto del desarrollo del ovario de una flor después de la fecundación.

  - Como un subjconjunto de un conjunto ya definido:

    Frutas = Fruto comestible

2.2 Conjuntos Difusos

En los conjuntos difusos relajamos la restricción de que la función de pertenencia valga ó 0 ó 1, y dejamos que tome valores en el intervalo [0,1]. La necesidad de trabajar con conjuntos difusos surge de un hecho: hay conceptos que no tienen límites claros. Por ejemplo:

¿Una persona que mide 1.80 es *alta*? ¿Una temperatura de 15 grados es *baja*?. Vemos que, a diferencia de lo que ocurre en el caso de las frutas (no hay vaguedad, un alimento o bien es una fruta o bien no lo es), en otras situaciones nos vemos obligados a tratar con ella.

Veamos algunas definiciones útiles:

- Llamaremos *variable lingüística* a aquella noción o concepto que vamos a calificar de forma difusa. Por ejemplo: la altura, la edad, el error, la variación del error... Le aplicamos el adjetivo "lingüística" porque definiremos sus características mediante el lenguaje hablado.

- Llamaremos *universo de discurso* al rango de valores que pueden tomar los elementos que poseen la propiedad expresada por la variable lingüística. En el caso de la variable lingüística 'altura de una persona normal', sería el conjunto de valores comprendido entre 1.4 y 2.3 m.

- Llamamos *valor lingüístico* a las diferentes clasificaciones que efectuamos sobre la variable lingüística: en el caso de la altura, podríamos dividir el universo de discurso en los diferentes valores lingüísticos: por ejemplo *bajo*, *mediano* y *alto*.

- Llamaremos conjunto difuso a un valor lingüístico junto a una función de pertenencia. El valor lingüístico es el “nombre” del conjunto, y la función de pertenencia se define como aquella aplicación que asocia a cada elemento del universo de discurso el grado con que pertenece al conjunto difuso. Decimos que un conjunto es *nítido* si su función de pertenencia toma valores en {0,1}, y *difuso* si toma valores en [0,1].

- Dado un conjunto difuso A, se define como *alfa-corte* de A, al conjunto de elementos que pertenecen al conjunto difuso A con grado mayor o igual que alfa, es decir:

  $$A_\alpha = \{x \in X / \mu_A(x) \geq \alpha\}$$

- Se define como *alfa corte estricto* al conjunto de elementos con grado de pertenencia estrictamente mayor que alfa, es decir:

  $$A_{\overline{\alpha}} = \{x \in X / \mu_A(x) > \alpha\}$$

- Se define como *soporte* de un conjunto difuso A, al conjunto nítido de elementos que tienen grado de pertenencia estrictamente mayor que 0, o sea, al alfa-corte estricto de nivel 0.

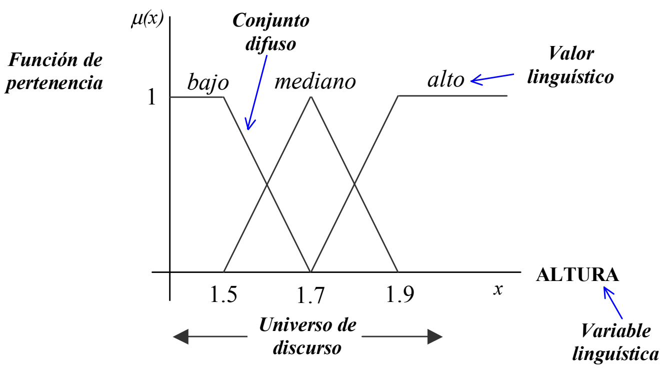
  $\text{Soporte}(A) = \{ x \in X \ / \ \mu_A(x) > 0 \}$
- Se define como *núcleo* de un conjunto difuso A, al conjunto nítido de elementos que tienen grado de pertenencia 1. (alfa-corte de nivel 1)

  $\text{Núcleo}(A) = \{ x \in X \ / \ \mu_A(x) = 1 \}$
- Se define la *altura* de un conjunto difuso A como el valor más grande de su función de pertenencia.
- Se dice que un conjunto difuso está *normalizado* si y solo si su núcleo contiene algún elemento (o alternativamente, si su altura es 1), es decir:

  $\exists\, x \in X \qquad \mu_A(x) = 1$
- El elemento x de U para el cual $\mu_F(x) = 0.5$ se llama el *punto de cruce*.
- Un conjunto difuso cuyo soporte es un único punto x de U y tal que la función de pertenencia de x es 1 (es decir, el soporte coincide con el núcleo y tienen un único punto) se llama un *conjunto difuso unitario (singleton)*.

Ejemplo 1

Consideremos la variable lingüística "Altura de los seres humanos", que toma valores en el universo de discurso U = [1.4, 2.50]. Vamos a hacer una clasificación difusa de los seres humanos en tres conjuntos difusos (o valores lingüísticos): *bajos*, *medianos* y *altos*.



En esta ilustración hemos dibujado 3 conjuntos difusos sobre la variable lingüística altura, cuyos valores lingüísticos asociados son *bajo*, *mediano* y *alto* respectivamente. Las funciones de pertenencia son de tipo L para *bajo*, Lambda o Triángulo para el *mediano* y Gamma para el *alto*. Más adelante aclararemos porqué usamos estos nombres, que únicamente determinan qué forma tendrán las funciones de pertenencia. De este modo si Luis mide 1.80 metros, la lógica difusa nos dice que es un 0.2 mediano y un 0.8 alto. De este modo expresamos que mientras un elemento puede estar dentro de un determinado conjunto, puede no cumplir las especificaciones de dicho conjunto al cien por cien (por ejemplo, en el caso de Luis, a la vista del resultado podríamos afirmar que es *poco mediano* y *más bien alto*.

En este ejemplo, dado el conjunto difuso mediano tenemos que:

- El alfa-corte 0.5 es el intervalo [1.6,1.8]
- El alfa corte estricto 0.5 es el intervalo (1.6, 1.8)
- El soporte es (1.5, 1.9)
- El núcleo es 1.7
- Es un conjunto difuso normalizado
- Tiene dos puntos de cruce: 1.6 y 1.8

La notación habitual para los conjuntos difusos es la definida por Lofti Zadeh, que es la siguiente: sea *A* un conjunto difuso definido sobre el universo U:

$$A=\{(x,\mu_A(x)) \ / \ x \in U\}$$

que indica que A está formado por todos los pares ordenados x y el resultado de la función de pertenencia para todo elemento u dentro del universo de discurso U. Para denotar el conjunto difuso A:

si el universo es discreto: $\sum_U \mu_A(x)/x$

si el universo es continuo: $F = \int_u \mu_A(x)/x$

¡Cuidado con esta notación! El sumatorio o la integral pierden su significado habitual, En lógica difusa quieren simbolizar una mera enumeración de tuplas. La barrra tampoco indica una fracción sino que simplemente separa los dos elementos de la tupla. Así por ejemplo el conjunto difuso discreto "Tirada alta del dado" podría definirse como:

$$F = \{ 0/1 + 0/2 + 0.3/3 + 0.6/4 + 0.9/5 + 1/6\}$$

La parte derecha de la tupla indica el elemento y la parte izquierda el grado de pertenencia.

Los conjuntos difusos y las funciones de pertenencia pueden emplearse de dos formas posibles:

a) Para estimar grados de pertenencia a un conjunto. Por ejemplo, si nos dicen que una persona mide 170 cm, ¿en qué grado es una persona alta?

b) Para expresar *posibilidades* en una situación en la que se dispone de información incompleta. Por ejemplo, si nos dicen que una persona es mediana, ¿cuál será su altura? En este caso la función de pertenencia μ puede interpretarse como una *distribución de posibilidad* que nos indica la preferencia sobre los valores que una variable de valor desconocido puede tomar.

De este modo vemos que la principal diferencia entre la teoría de conjuntos clásica y la difusa es que mientras que los valores de la función de pertenencia de un conjunto nítido son siempre 0 o 1, la función de pertenencia de un conjunto difuso toma valores en todo el intervalo [0,1][3]. De este modo vemos que, al contrario de los conjuntos nítidos, que pueden definirse de varias formas, los conjuntos difusos vienen siempre

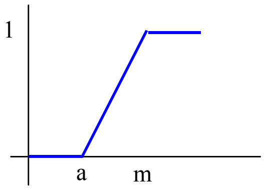
[3] Se suele normalizar el grado de pertenencia máximo a 1.

definidos por su función de pertenencia. Veamos qué tipos de funciones de pertenencia se usan más habitualmente en la lógica difusa.

2.2.1 Funciones de pertenencia

Aunque en principio cualquier función sería válida para definir conjuntos difusos, en la práctica hay ciertas funciones típicas que siempre se suelen usar, tanto por la facilidad de computación que su uso conlleva como por su estructura lógica para definir su valor lingüístico asociado. Las funciones más comunes son:

- Función GAMMA (Γ):

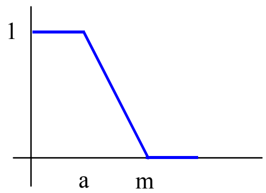


$$\mu(x) = \begin{cases} 0 & \text{si} \quad x \le a \\ (x-a)/(m-a) & \text{si} \quad x \in (a,m) \\ 1 & \text{si} \quad x \ge m \end{cases}$$

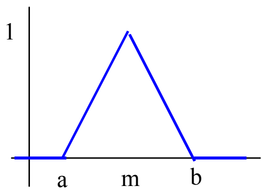
$$\mu(x) = \begin{cases} 0 & \text{para } \ x \le a \\ \dfrac{x-a}{m-a} & \text{para } a < x < m \\ 1 & \text{para } x \ge m \end{cases}$$

- Función L:

Puede definirse simplemente como 1 menos la función GAMMA



- Función LAMBDA o triangular:

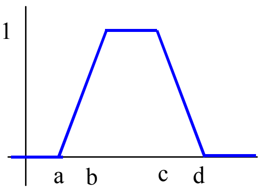


$$\mu(x) = \begin{cases} 0 & \text{si } x \le a \\ (x-a)/(m-a) & \text{si } x \in (a,m] \\ (b-x)/(b-m) & \text{si } x \in (m,b) \\ 0 & \text{si } x \ge b \end{cases}$$

$$\mu(x) = \begin{cases} 0 & \text{para } \ x \le a \\ \dfrac{x-a}{m-a} & \text{para } \ a < x \le m \\ \dfrac{b-x}{b-m} & \text{para } \ m < x \le b \\ 0 & \text{para } \ x > b \end{cases}$$

- Función PI o trapezoidal:

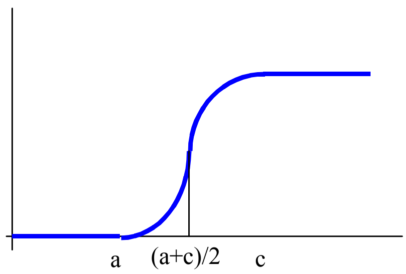
$$\mu(x)=\begin{cases} 0 & \text{para } x \le a \\ \dfrac{x-a}{b-a} & \text{para } a < x \le b \\ 1 & \text{para } b < x \le c \\ \dfrac{d-x}{b-c} & \text{para } c < x \le d \\ 0 & \text{para } x > d \end{cases}$$



Las funciones L y GAMMA se usan para calificar valores lingüísticos extremos, tales como *bebé* o *anciano*, respectivamente. Las funciones PI y LAMBDA se usan para describir valores intermedios (como *joven*, de *mediana edad*, *maduro*). Su principal diferencia reside en que la función PI implica un margen de tolerancia alrededor del valor que se toma como más representativo del valor lingüístico asociado al conjunto difuso.
También se pueden utilizar otras funciones que no sean lineales a trozos. Por ejemplo, en FuzzyCLIPS se utilizan las siguientes funciones:

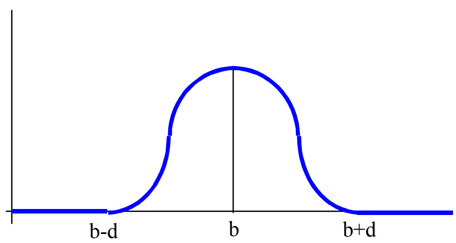
- Función s, definida mediante:

$$\mu_S(x) = \begin{cases} 0 & \text{para } x \le a \\ 2\left(\dfrac{x-a}{c-a}\right)^2, & \text{para } a \le x \le \dfrac{a+c}{2} \\ 1-2\left(\dfrac{x-a}{c-a}\right)^2, & \text{para } \dfrac{a+c}{2} \le x \le c \\ 1 & \text{para } x \ge c \end{cases}$$



- Función z, que es la función opuesta,
  $\mu_Z(x) = 1- \mu_S(x)$
- Función Π , definida mediante

$$\mu_\Pi(x) = \begin{cases} \mu_S(x) & \text{para } x \le b \\ \mu_Z(x) & \text{para } x > b \end{cases}$$



2.2.2 Etiquetas lingüísticas

Tradicionalmente se han utilizado modificadores de los conjuntos difusos a los que llamamos etiquetas lingüísticas, equivalentes a lo que en lenguaje natural serían los adverbios. La interpretación en el modelo difuso de estos enunciados consiste en la composición de la función de pertenencia con una operación aritmética simple. Por ejemplo, es habitual considerar como interpretación del adverbio *muy* el cuadrado de la función de pertenencia original, Es decir, "Juan es muy alto" se interpretaría como:

$$\mu_{MUY\ ALTO}(x) = (\mu_{ALTO}(x))^2$$

Existe una amplia bibliografía sobre el tema, véase por ejemplo [ESH81], de la que podemos entresacar algunas interpretaciones clásicas de los adverbios, aunque no siempre serán las más adecuadas a un problema de representación del conocimiento concreto:

MUY $\mu_{MUY\ A}(x) = (\mu_A(x))^2$

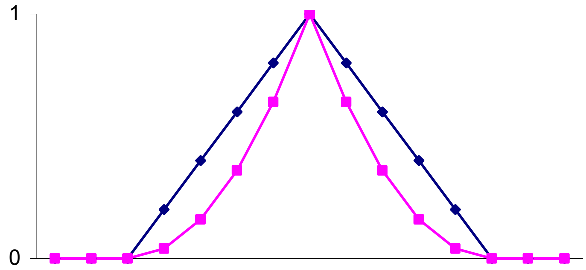
De este modo, si el grado de pertenencia de una persona a la clase *alto* es 0.5, el grado de pertenencia a la clase muy alto es sólo 0.25.

ALGO $\mu_{ALGO\ A}(x) = \sqrt{\mu_A(x)}$

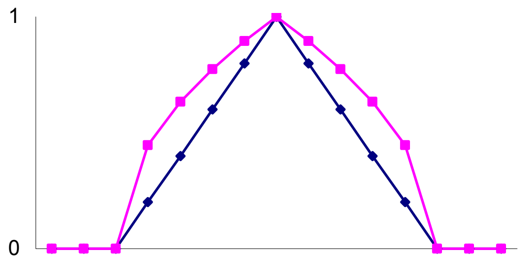
Así, si el grado de pertenencia de una persona a la clase alto es 0.5, el grado de pertenencia a la clase algo alto es de 0.707.

Existe todo un catálogo de posibles adverbios y sus modificadores asociados, pero las modificaciones que más usualmente se aplican a un conjunto difuso son las siguientes:

- *Normalización*, al convertir un conjunto difuso no normalizado en uno normalizado (dividiendo por la altura del conjunto).
- *Concentración*, al componer con una función tipo $f(y)=y^p$, con $p>1$. El efecto es que la función de pertenencia toma valores más pequeños, centrándose en los valores mayores.
  El efecto de aplicar la concentración puede verse en la siguiente figura (la función de pertenencia base es la azul, y la modificada la rosa):

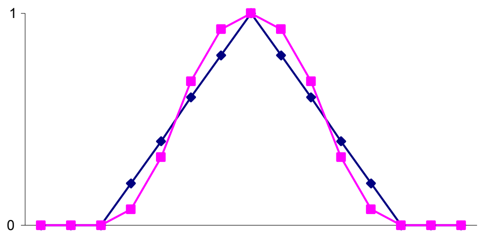


- *Dilatación*, al componer con una función tipo $f(y)=y^p$ con $0<p<1$ (o también con $2y-y^2$). El efecto es el contrario a la concentración.



- *Intensificación del contraste*. Se disminuyen los valores menores a 1/2 y se aumentan los mayores. Componemos con una función del tipo: $f(y) = \begin{cases} 2^{p-1}y^p & \text{para } y \le 0.5 \\ 1-2^{p-1}(1-y)^p & \text{en otro caso} \end{cases}$
  donde $p>1$. Normalmente se suele poner $p=2$ (a mayor p, mayor intensificación).

  El efecto es:

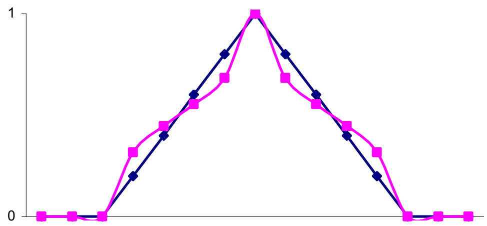


- *Difuminación*. Efecto contrario al anterior.

  Se compone con la función:

$$f(y) = \begin{cases} \sqrt{y/2} & \text{para } y \leq 0.5 \\ 1 - \sqrt{(1-y)/2} & \text{en otro caso} \end{cases}$$

Cuyo efecto es:



Los adverbios o modificadores pueden componerse entre sí, obteniendo múltiples combinaciones para representar enunciados complejos como "Juan es mucho más que alto". FuzzyCLIPS tiene varios modificadores predefinidos:

Nombre del modificador	Descripción del modificador
not	1-y
very (muy)	y^2
somewhat (algo)	$y^{1/3}$
more-or-less (más o menos)	$y^{1/2}$
extremely (extremadamente)	y^3

y también admite la definición de nuevos modificadores por el usuario.

2.3 Operaciones elementales con conjuntos difusos

Al igual que en la teoría clásica de conjuntos, sobre los conjuntos difusos podemos definir las operaciones de unión, intersección, complementario, etc.

2.3.1 Complementario

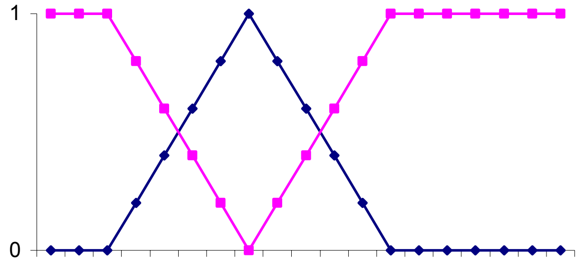
Dado un conjunto A, el conjunto complementario de A está formado por los elementos del universo que no pertenecen a A. En el caso difuso, este conjunto vendrá definido por una función de pertenencia que se calcula para cada elemento a partir de su pertenencia al conjunto A. Es decir:

$$\mu_{\overline{A}}(x) = c(\mu_A(x))$$

siendo c una función c: [0,1] → [0,1] que, dado el grado de pertenencia al conjunto A, nos da el grado de pertenencia al conjunto complementario de A. A esta función c desde un punto de vista intuitivo deben exigírseles las siguientes características:

- **c1.** concordancia con el caso nítido $c(1) = 0$ y $c(0) = 1$
- **c2.** estrictamente decreciente $\forall \alpha,\beta \in [0,1]\ \alpha>\beta \Rightarrow c(\alpha) < c(\beta)$
- **c3.** involución $\forall \alpha \in [0,1]\ c(c(\alpha)) = \alpha$

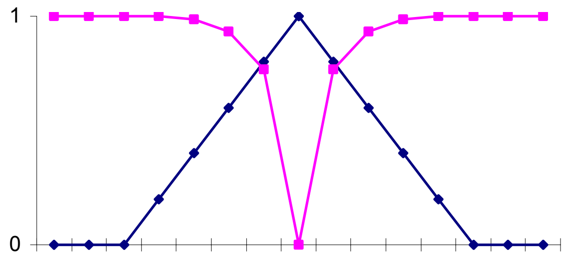
En general se considera como función del complementario a $c(\alpha) = 1 - \alpha$. Así, para el conjunto difuso definido por una función triangular (por ejemplo, el conjunto difuso *mediano*) su complemento sería:



aunque también existen otras variantes que cumplen las propiedades antes citadas como:

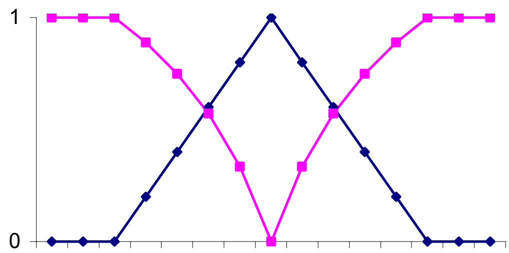
- Complementario de Yager $c_w(\alpha) = (1 - \alpha^w)^{1/w}$ $w \in [0, \infty]$

Para una función triangular y con w=2, tendríamos:



- clase de complementarios de Sugeno $c_\lambda(\alpha) = \dfrac{1-\alpha}{1-\lambda\alpha}$ $\lambda \in [0,1]$

para $\lambda = 1/2$:



2.3.2 Intersección

En teoría de conjuntos clásica, se considera que un elemento pertenece al conjunto intersección de dos conjuntos si pertenece a ambos. En el caso difuso el problema consiste en determinar el grado de pertenencia al conjunto intersección, conocido el grado de pertenencia a cada uno de los conjuntos originales. Supongamos:

$\mu_{A \cap B}(x) \;=\; i(\mu_A(x), \mu_B(x))$

donde:

$i: \; [0,1] \times [0,1] \;\rightarrow\; [0,1]$

análogamente al caso anterior, imponemos las siguientes condiciones:

$\forall \alpha, \beta, \gamma \in [0,1]$

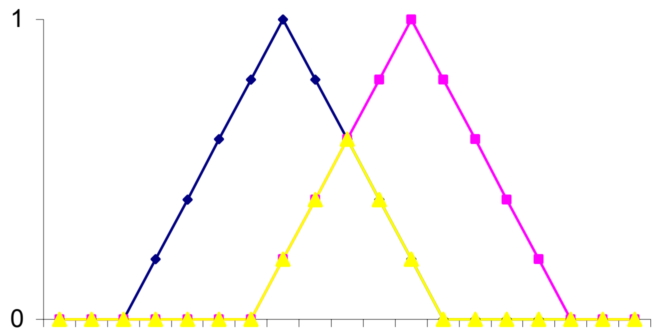
- **i1.** concordancia con el caso nítido $\quad i(0,1) = i(0,0) = i(1,0) = 0; \quad i(1,1) = 1$
- **i2.** conmutatividad $\quad i(\alpha,\beta) = i(\beta,\alpha)$
- **i3.** asociatividad $\quad i(\alpha,i(\beta,\gamma)) = i(i(\alpha,\beta),\gamma)$
- **i4.** identidad $\quad i(\alpha,1) = \alpha$
- **i5.** monotonía $\quad$ si $\alpha \leq \alpha'$ $\;\beta \leq \beta'$, entonces $i(\alpha,\beta) \leq i(\alpha', \beta')$

Si se verifican los axiomas anteriores ([0,1],i) tiene estructura de semigrupo abeliano con elemento neutro. Las funciones i que verifican esta propiedad se llaman dentro de la teoría de conjuntos difusos *normas triangulares* (t-normas).

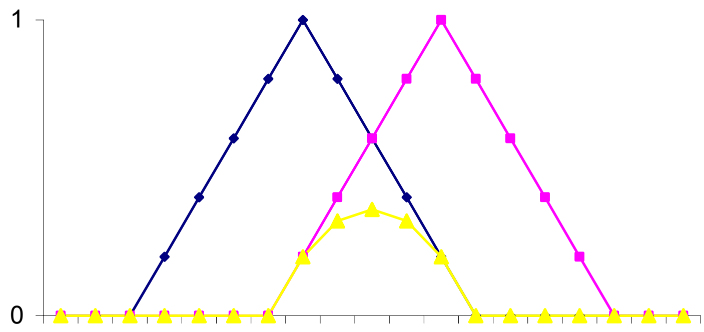
Las t-normas usadas más habitualmente son las siguientes:

- t-norma del mínimo $\quad i_{min}(\alpha,\beta) \;= \min(\alpha,\beta)$

Por ejemplo si consideramos dos funciones tipo triangular (*niño*, *adolescente*), la t-norma del mínimo sería:

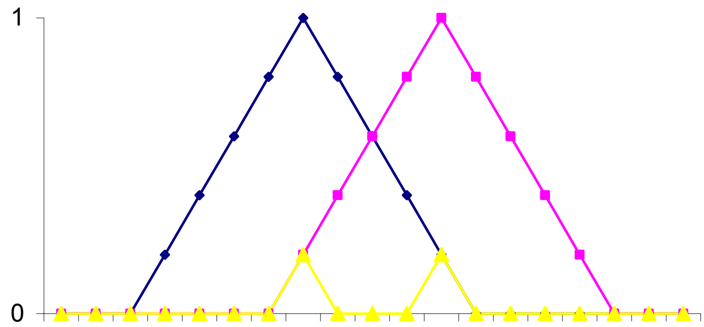


- t norma del producto $\quad i_{*}(\alpha,\beta) \;= \alpha * \beta$



- t-norma del producto drástico

$$i_{inf}(\alpha,\beta) \;=\; \begin{cases} \alpha & \text{si } \beta = 1 \\ \beta & \text{si } \alpha = 1 \\ 0 & \text{en otro caso} \end{cases}$$



Aunque no siempre se puede decir que una t-norma es mayor que otra, se puede demostrar que toda t-norma verifica las siguientes desigualdades:

$$\forall \alpha,\beta \in [0,1] \qquad i_{inf}(\alpha,\beta) \leq i(\alpha,\beta) \leq i_{min}(\alpha,\beta) ,$$

es decir, que la menor t-norma es la t-norma del producto drástico y la mayor t-norma es la norma del mínimo.

2.3.3 Unión

Al igual que en el caso anterior podemos declarar una axiomática intuitiva para la unión de dos conjuntos difusos. Sea:

$$\mu_{A \cup B}(x) = u(\mu_A(x), \mu_B(x))$$

en donde:

$u : [0,1] \times [0,1] \to [0,1]$

que debe verificar:

$\forall \alpha,\beta,\gamma \in [0,1]$

- **u1.** concordancia con el caso nítido $\quad u(0,1) = u(1,1) = u(1,0) = 1; \quad u(0,0) = 0$
- **u2.** conmutatividad $\quad u(\alpha,\beta) = u(\beta,\alpha)$
- **u3.** asociatividad $\quad u(\alpha,u(\beta,\gamma)) = u(u(\alpha,\beta),\gamma)$
- **u4.** identidad ($A \cup \varnothing = A$) $\quad u(\alpha,0) = \alpha$
- **u5.** monotonía $\quad$ Si $\alpha \leq \alpha'$ $\beta \leq \beta'$, entonces $u(\alpha,\beta) \leq u(\alpha', \beta')$

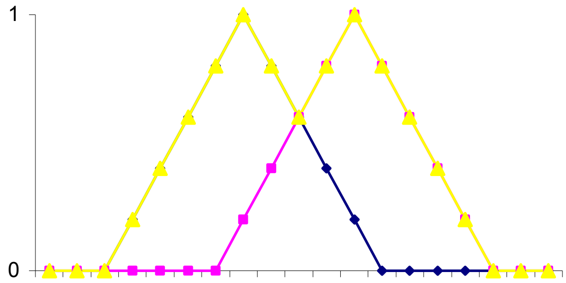
Además, sería deseable que se mantuvieran también las siguientes propiedades:

- **u6.** Leyes de De Morgan $\quad u(\alpha,\beta) = c(i(c(\alpha),c(\beta))$
  $\quad i(\alpha,\beta) = c(u(c(\alpha),c(\beta))$

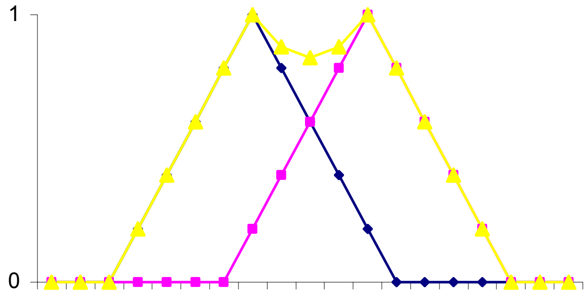
Que nos permiten calcular el grado de la unión en función de los grados del complementario y la intersección. A las funciones que verifiquen estas seis propiedades se las llama *conormas triangulares* (t-conormas).

Considerando la función de complementación $c(\alpha) = 1 - \alpha$, las t-conormas correspondientes a las t-normas anteriores son:

- t-conorma del máximo $\quad u_{max}(\alpha,\beta) = \max(\alpha,\beta)$

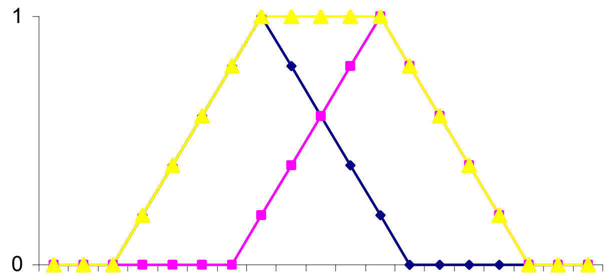


- t-conorma de la suma $u_*(\alpha,\beta) = \alpha + \beta - \alpha * \beta$



- t-conorma de la suma drástica

$$u_{sup}(\alpha,\beta) = \begin{cases} \alpha & \text{si } \beta = 0 \\ \beta & \text{si } \alpha = 0 \\ 1 & \text{en otro caso} \end{cases}$$



análogamente al caso de la intersección, se puede demostrar que cualquier t-conorma verifica las desigualdades:

$\forall \alpha,\beta \in [0,1]$ $\quad u_{max}(\alpha,\beta) \leq u(\alpha,\beta) \leq u_{sup}(\alpha,\beta)$

Es decir, que la menor t-conorma es la del máximo y la mayor t-conorma la suma drástica.

Pero las condiciones que exigimos a la unión y a la intersección no garantizan en general que se cumplan las siguientes condiciones:

$\forall \alpha,\beta,\gamma \in [0,1]$

I1: Idempotencia $(A \cap A = A)$ $\quad i(\alpha,\alpha) = \alpha$

I1: Distributividad $(A \cap (B \cup C)) = ...$ $\quad i(\alpha,u(\beta,\gamma)) = u(i(\alpha,\beta),i(\alpha,\gamma))$

U1 : Idempotencia $(A \cup A = A)$ $\quad u(\alpha,\alpha) = \alpha$

U2 : Distributividad $(A \cup (B \cap C)) = ...$ $\quad u(\alpha,i(\beta,\gamma)) = i(u(\alpha,\beta),u(\alpha,\gamma))$

propiedades que sólo verifica la t-norma del mínimo y su t-conorma del máximo.

Podríamos también definir el conjunto vacío y el conjunto universal. El concepto de conjunto vacío corresponde al de aquel conjunto que no contiene ningún elemento. Por tanto, parece adecuado definirlo en la teoría de conjuntos difusos como:

$$\forall\, x \in X \quad \mu_{\varnothing}(x) = 0$$

y consiguientemente el conjunto universal se definiría como:

$$\forall\, x \in X \quad \mu_{X}(x) = 1$$

Pero asumiendo estas definiciones no se verifican en la teoría de conjuntos difusos algunos famosos teoremas de la teoría de conjuntos clásica, como:

$$A \cap \overline{A} = \varnothing$$

$$A \cup \overline{A} = X$$

que se conocen como el principio de contradicción y del tercio excluso, respectivamente (lógica aristotélica).

Si tomamos por ejemplo el conjunto difuso "joven" es fácil comprobar que no se cumplen ninguno de los dos principios.

Sin embargo es posible definir una t-norma y una t-conorma que satisfagan esto (la t-norma del producto acotado y la t-norma de la suma acotada), aunque entonces no se satisfarán las propiedades I1,I2,U1,U2.

2.4 Razonamiento difuso

La teoría de conjuntos difusos nos permite representar hechos y relaciones vagas (imprecisas). Se entiende por razonamiento difuso el proceso de realizar inferencias a partir de hechos y relaciones difusas, así como la combinación de evidencias difusas y la actualización de la precisión de las creencias.

Una *proposición difusa simple* es aquella que asigna un valor a una variable difusa, por ejemplo: "la estatura de Pepe es *mediana*" o "la velocidad es *normal*". Una proposición difusa tiene por tanto asociado un conjunto difuso A (el valor lingüístico asignado, "mediana" en este caso) y su correspondiente función de pertenencia μ_A definida sobre los elementos del universo de discurso $u \in U$.

Una *proposición difusa compuesta* es aquella que se obtiene mediante la agrupación de dos o más proposiciones difusas simples, que pueden haber sido modificadas o no antes de la agrupación. Para agrupar proposiciones difusas simples podemos utilizar las conectivas Y y O, y para modificar una proposición difusa simple podemos utilizar el NO. Así por ejemplo podemos construir proposiciones difusas del tipo:

"la velocidad es *normal*" Y "el objeto está *cerca*"
"la velocidad es *alta*" O "el objeto está *muy cerca*"
"la velocidad NO es *alta*"

Los *operadores lógicos difusos* pueden definirse de forma análoga a como se definieron las operaciones entre conjuntos: sean p y q dos proposiciones difusas, A y B los conjuntos difusos que intervienen en ellas, con funciones de pertenencia μ_A y μ_B definidas respectivamente sobre universos de discurso U y V. Entonces, los operadores lógicos pueden definirse mediante:

- NO ($\neg p$) vendrá definida por una función de pertenencia tipo complemento de A, por ejemplo $\mu_{\neg A}(u) = 1 - \mu_A(u)$
- Y ($p \wedge q$) vendrá definida por una función de pertenencia tipo intersección, por ejemplo $\mu_{A \wedge B}(u,v) = \mu in(\mu_A(u), \mu_B(v))$
- O ($p \vee q$) vendrá definida por una función de pertenencia tipo unión, por ejemplo $\mu_{A \vee B}(u,v) = \mu ax(\mu_A(u), \mu_B(v))$

Ahora tenemos que definir lo que significa una implicación, es decir, tenemos que asignar una función de pertenencia a una agrupación antecedente consecuente del tipo $p \rightarrow q$. Definir el significado de la implicación nos permitirá razonar con reglas del tipo:

SI "la velocidad es *normal*"
ENTONCES "la fuerza de frenado debe ser *moderada*"

Esta función de pertenencia será del tipo:

$$\mu_{p \rightarrow q} : \quad U \times V \rightarrow [0,1]$$
$$\qquad (u,v) \rightarrow \mu_{p \rightarrow q}(u,v)$$

Al definir la relación de implicación surge una cuestión importante ¿qué se quiere representar mediante la relación de implicación? La cuestión es fundamental porque las relaciones de implicación son la base del razonamiento basado en reglas. Existen en principio dos posibilidades diferenciadas:

a) Dar a la implicación el mismo significado que en la lógica clásica. Por ejemplo, en lógica clásica tenemos la equivalencia $p \rightarrow q \equiv \neg p \vee q$. De este modo, la función de pertenencia asociada a la regla "Si A entonces B", donde A y B son conjuntos difusos sería:

$$\mu_{p \rightarrow q}(u,v) = \max(1-\mu_A(u), \mu_B(v))$$

En lógica clásica también tenemos la equivalencia $p \rightarrow q \equiv \sim(p \wedge (\sim q))$, que conduciría a la siguiente definición:

$$\mu_{p \rightarrow q}(u,v) = 1 - \min[\mu_A(u), 1-\mu_B(v)]$$

Ambas funciones son equivalentes.

b) Dar a la implicación el significado de relaciones causa-efecto normalmente utilizadas en los sistemas basados en conocimiento.

El primer caso se corresponde con la implicación lógica, al segundo podemos darle el nombre de "implicación causa-efecto". Aunque la implicación lógica difusa es interesante desde el punto de vista teórico, conduce a una formulación inadecuada para muchas aplicaciones de sistemas basados en conocimiento, que representan las relaciones causa-efecto de un modo no consistente plenamente con la lógica. Queda entonces abierta la cuestión de como formalizar el significado de las relaciones de implicación causa-efecto. La más utilizada actualmente fue propuesta por Mamdani:

IMPLICACIÓN DE MAMDANI: $\quad p \rightarrow q \equiv p \wedge q \quad \Rightarrow \mu_{p \rightarrow q}(u,v) = \min(\mu_A(u), \mu_B(v))$

Para Mamdani, el grado de verdad de p→q es idéntico al de la proposición A y B. Podríamos justificar esto diciendo que, para Mamdani, una condición tan sólo resulta cierta cuando el antecedente es cierto y el consecuente también.

2.4.1 Inferencia difusa

Con la formalización anterior pueden representarse hechos y reglas difusas, y pueden realizarse inferencias aplicando reglas de inferencia. Veremos dos casos: antecedentes nítidos y antecedentes difusos.
2. *Inferencia difusa con antecedentes difusos*. Vamos a suponer que tenemos una regla difusa del tipo:

Si p ENTONCES q
y un valor de entrada difuso p*.

La conclusión será un hecho difuso q*, del cual queremos saber su función de pertenencia.

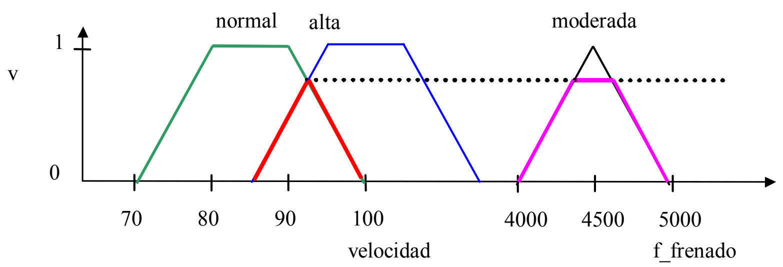
Ejemplo

Sean:
la regla difusa
p→q ≡ "SI la velocidad es *normal*, ENTONCES la fuerza de frenado es *moderada*".
el hecho difuso p*

El resultado de la inferencia será una proposición difusa q* con su correspondiente conjunto difuso B* asociado, que vendrá dado por la función de pertenencia $\mu_B*(v)$. Supongamos que el hecho del que disponemos es p* = "la velocidad es *alta*". Vamos a ver dos tipos diferentes de inferencia:

1) inferencia tipo max-min (implicación de Mamdani):

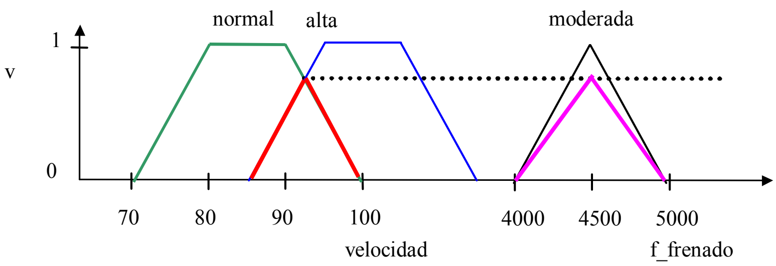
$$\mu_B*(v) = \min(z, \mu_B(v))$$
$$\text{donde } z = \max\left(\min\left(\mu_{A*}(u), \mu_A\,(u)\right)\right)$$



Con este tipo de inferencia estamos considerando la implicación como una implicación de Mamdani.

b) inferencia tipo max-prod:

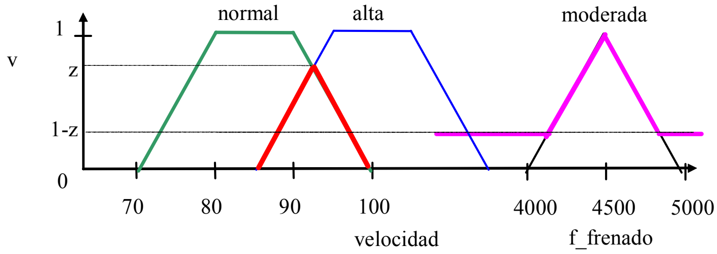
$$\mu_B*(v) = \text{prod}(z, \mu_B(v))$$
$$\text{donde } z = \max\left(\min\left(\mu_{A*}(u), \mu_A\,(u)\right)\right)$$



Se puede hacer otro tipo de inferencia interpretando la implicación como una implicación de la lógica en lugar de como una implicación de Mamdani. Para ello, las funciones utilizadas serían:

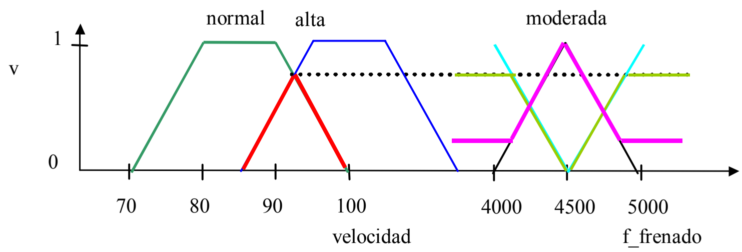
$$\mu_B^*(v) = \max(1-z, \mu_B(v))$$

cuyo resultado es:



o

$$\mu_B^*(v) = 1-\min(z, 1-\mu_B(v))$$



2. *Inferencia difusa con antecedentes nítidos*. Vamos a suponer que tenemos una regla difusa del tipo:

Si p ENTONCES q
y un valor de entrada nítido p*.

La conclusión será un hecho difuso q*, del cual queremos saber su función de pertenencia.

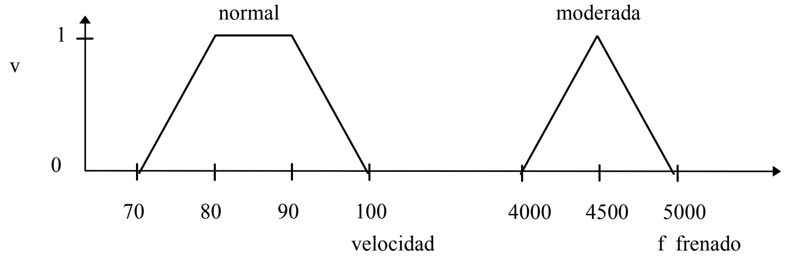
Ejemplo:
Sean:

la regla difusa
$p \rightarrow q \equiv$ "SI la velocidad es normal, ENTONCES la fuerza de frenado es moderada".
el hecho nítido

p* ≡ “la velocidad es 75 km/h”
El hecho p* puede utilizarse para disparar la regla y obtener así un valor difuso para la fuerza de frenado que debe aplicarse:

velocidad = 75
SI velocidad = normal ENTONCES fuerza_frenado = moderada
fuerza_frenado = q*

Supongamos que las funciones de pertenencia de los conjuntos difusos A = *normal* y B= *moderada* son los que se dan en la siguiente figura:



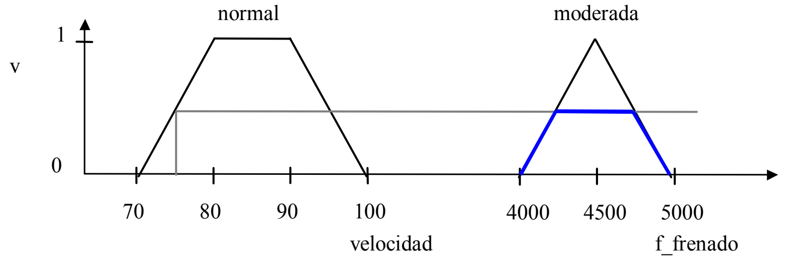
Vemos que el valor de la función de pertenencia para el hecho nítido p es $\mu_A*(75) = 0.5$. La función de pertenencia asociada a la regla es $\mu_{p\rightarrow q}(x,y) = \min(\mu_A(u), \mu_B(v))$. El resultado de la inferencia será una proposición difusa q* con su correspondiente conjunto difuso B* asociado, que vendrá dado por la función de pertenencia $\mu_B*(y)$.

La única diferencia con el caso difuso es la forma de escoger el valor z, que en este caso se calcula simplemente como $z=\mu_A(x)$, donde x es el valor nítido del que dispongamos, en nuestro ejemplo 75. La inferencia se hace entonces con cualquiera de las alternativas vistas en el apartado anterior, por ejemplo:

- Inferencia tipo max-min (implicación de Mamdani):

$$\mu_B*(y) = \min(\mu_A(75), \mu_B(y))$$

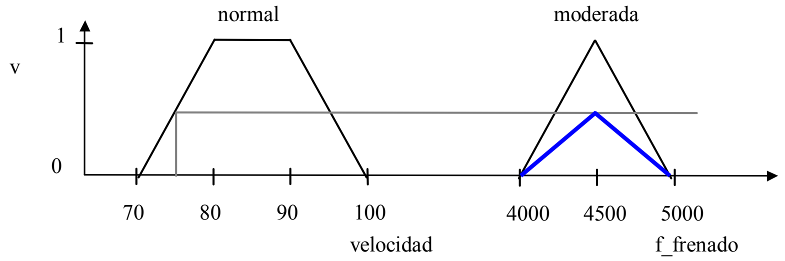
Es decir, que el resultado de la inferencia sería “velocidad es moderada*”, donde la función de pertenencia del conjunto difuso moderada* es la representada en negrita en la siguiente figura:



- Inferencia tipo max-prod:

$$\mu_B*(y) = \text{prod}(\mu_A(75), \mu_B(y))$$

Cuyo resultado es el representado en la siguiente figura:

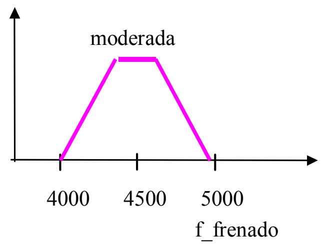


Igual se haría en el caso de que la implicación se interprete como una implicación lógica.

2.4.2 Decodificación (transformación de un conjunto difuso en un valor nítido)

La principal aplicación de los sistemas de razonamiento difuso es el control de dispositivos, que normalmente precisan de una salida nítida (acción de control). Así por ejemplo en el ejercicio anterior podemos querer saber qué fuerza de frenado que debemos aplicar si la velocidad es *alta*. Existen diversas alternativas para transformar un valor difuso en nítido (proceso que en inglés se llama defuzzification y en español podríamos llamar decodificación), las más empleadas son:

a) El valor máximo (es decir, el más posible). Si se producen empates puede seleccionarse el primer valor encontrado o la media (en cuyo caso el método se denomina media de máximos (MOM en FuzzyCLIPS).



En nuestro ejemplo, nos encontramos que la función de pertenencia tiene varios máximos: todos los valores entre 4250 y 4750. La estrategia MOM cogería el valor medio, esto es, 4500, y la estrategia del primer valor máximo cogería 4250.

b) El centroide difuso (o centro de gravedad, COG en FuzzyCLIPS), definido como:

$$y_{centroide} = \frac{\sum_{x \in X} x\, \mu_A(x)}{\sum_{x \in X} \mu_A(x)}$$

En nuestro ejemplo sería: $\dfrac{4000*0+4250*1+4750*1+5000*0}{1+1} = 4500$

Normalmente, si varias reglas tienen el mismo consecuente, lo que se suele hacer para acumular la evidencia es unir los dos conjuntos difusos resultantes y después decodificar el resultado.