# Incertidumbre y razonamiento aproximado

## Sistemas difusos

PID_00267998

Vicenç Torra i Reventós
Lluís Godo i Lacasa

Revisión a cargo de
David Isern Alarcón

Tiempo mínimo de dedicación recomendado: 4 horas

Universitat
Oberta
de Catalunya

**Vicenç Torra i Reventós**

**Lluís Godo i Lacasa**

**David Isern Alarcón**

La revisión de este recurso de aprendizaje UOC ha sido coordinada por el profesor: Carles Ventura Royo (2019)

Primera edición: septiembre 2019

Av. Tibidabo, 39-43, 08035 Barcelona
Realización editorial: FUOC

# Índice

# Introducción

Como se ha comentado en el módulo dedicado a los sistemas basados en el conocimiento, se han desarrollado métodos específicos de razonamiento para tratar los diferentes tipos de situaciones que un sistema puede encontrar. Así, se han construido herramientas para razonar sobre el tiempo, para razonar sobre el espacio o para poder actuar cuando la información de la que se dispone no es completa.

En este módulo nos centraremos en una de estas cuestiones: la de razonar con información incompleta (esto es, cuando la información de la que se dispone no es perfecta). De entre las alternativas existentes para tratar este problema, nos centraremos en los sistemas difusos (razonamiento con información imprecisa).

Para estudiar estas herramientas, el módulo se ha dividido en dos apartados:

**1)** El primero es una introducción donde se plantea el problema del razonamiento con información incompleta. Se introduce aquí la necesidad de medir la incertidumbre, y se dan tipos de medida.

**2)** El segundo está dedicado a los sistemas difusos. Se estudia que el conocimiento se representa mediante reglas en las que aparecen los llamados *conjuntos difusos*. Después de introducir estos conjuntos y de ver la forma que tienen las reglas, se describe cómo se realiza la inferencia y se da un ejemplo del funcionamiento de estos sistemas.

# Objetivos

Este módulo didáctico pretende que alcancéis los objetivos siguientes:

**1.** Ser capaces de situar el problema de la información incompleta.

**2.** Introducir las medidas de incertidumbre.

**3.** Conocer formas de razonar cuando la información es incompleta.

**4.** Entender los sistemas difusos.

# 1. Razonamiento con información incompleta: incertidumbre e imprecisión

**Lluís Godo i Lacasa**

Muy pocas veces la información o conocimiento disponible sobre un dominio concreto es completa y precisa, bien al contrario, muy probablemente es parcial, incierta, imprecisa o, en general, incompleta. Sin embargo, aun así, los humanos somos capaces de desarrollar razonamientos y de extraer conclusiones en muchas de estas situaciones. Por lo tanto, si queremos dotar un sistema basado en el conocimiento de capacidades de razonamiento parecidas a las de un experto humano, este sistema ha de ser capaz de representar y razonar con información imperfecta. Esta imperfección se debe principalmente a:

**1)** La poca fiabilidad de la información (apreciaciones subjetivas, limitaciones físicas de los sensores, etc.).

**2)** La información incompleta.

**3)** La imprecisión en la información inherente al uso del lenguaje natural.

**4)** La agregación de información conflictiva, contradictoria o redundante.

En este punto, y para evitar confusiones posteriores, conviene recordar que incertidumbre e imprecisión son dos conceptos de naturaleza diferente, incluso si excluimos la vaguedad como caso particular de imprecisión.

**Incertidumbre e imprecisión**

Si nuestra única información (imprecisa) es I1 = "El termómetro está por encima de los 20 °C", la proposición P1 = "La temperatura exterior es 25 °C", a pesar de ser un enunciado totalmente preciso, es a la vez una proposición totalmente incierta (saber que la temperatura está por encima de 20 °C no es suficiente para afirmar o negar que la temperatura sea de 25 °C). En cambio, la proposición P2 = "La temperatura exterior es superior a los 15 °C", aunque mucho más imprecisa, es totalmente cierta (en efecto, si la temperatura es superior a 20 °C, también es superior a 15 °C). Por otra parte, si la información de la cual disponemos es mucho más precisa, por ejemplo I2 = "El termómetro marca exactamente 24 °C", entonces desaparece toda incertidumbre: obviamente la proposición P2 continúa siendo cierta, pero ahora podemos afirmar que P1 es falsa, y así podemos determinar la certeza o falsedad de cualquier otra proposición en lo referente a la temperatura actual, siempre y cuando no queramos precisar más allá de los grados.

Por lo tanto, la idea que subyace es que la imprecisión es algo inherente al lenguaje de representación, es decir, pueden existir enunciados más o menos precisos y enunciados más o menos imprecisos. En cambio, la incertidumbre en la verdad o falsedad de un enunciado (preciso o impreciso) aparece tan pronto como la información disponible para la evaluación de este enunciado no es suficientemente completa o precisa. En cualquier caso, podemos observar que cuanta más imprecisión hay en la información, más incertidumbre en las conclusiones (y cuanta más precisión tenemos en la información, menos incertidumbre en las conclusiones). Por otra parte, cuanto más imprecisos son los enunciados, menos incertidumbre tenemos sobre su validez.

### 1.1. Medidas de incertidumbre

La manera más usual de tratar la incertidumbre es "midiéndola", es decir, utilizando una medida (en el sentido matemático del término) sobre un conjunto de acontecimientos, proposiciones, conjuntos de valores de las variables, etc. que queremos representar y sobre los que queremos razonar. Este conjunto no puede ser un conjunto cualquiera sino que debe poseer una estructura algebraica que permita manipular y/o combinar sus elementos. La estructura más habitual la constituye el álgebra de Boole.

**Ejemplos de álgebra de Boole**

**1)** El álgebra de subconjuntos de un cierto conjunto (conjunto de las partes) con las operaciones de intersección, unión y complementación, y con el conjunto vacío y el total como cero y unidad.

**2)** El álgebra de proposiciones de un lenguaje lógico con la semántica de la lógica clásica (es decir, con dos valores de verdad, cierto y falso) con los conectivos conjunción, disyunción y negación como operaciones y las constantes "false" y "true" como elementos cero y unidad.

**Álgebra de Boole**

Recordemos que es una estructura $(A, \wedge, \vee, \neg, 0, 1)$, donde $A$ es un conjunto, $\wedge$ y $\vee$ son operaciones binarias sobre $A$, $\neg$ es una operación unaria, y 0, 1 son dos elementos de $A$, tal que se verifican las propiedades siguientes:

**1)** $\wedge$, $\vee$ son operaciones conmutativas, asociativas e idempotentes.

**2)** 0, 1 son elementos cero y unidad respectivamente: $a \wedge 0 = 0$, $a \wedge 1 = a$, $a \vee 0 = a$, $a \vee 1 = 1$.

**3)** absorción: $(a \wedge b) \vee b = b$, $a \wedge (a \vee b) = a$.

**4)** distributivas: $a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c)$, $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$.

**5)** complementación: $a \wedge \neg a = 0$, $a \vee \neg a = 1$.

Aunque equivalentes, a lo largo de este apartado adoptaremos por comodidad un marco conjuntístico en lugar de un marco lógico, así hablaremos de conjuntos y no de proposiciones.

La noción más general de medida de incertidumbre sobre el álgebra de Boole de las partes de un universo $W$, que denotaremos $2^W$, es la de una función $\mu$: $2^W \rightarrow [0, 1]$ tal que se verifican las propiedades siguientes:

**1)** $\mu(\emptyset) = 0$

**2)** $\mu(W) = 1$

**3)** si $A \subseteq B$ entonces $\mu(A) \leq \mu(B)$

El intervalo real [0, 1] se toma como escala, donde el 0 representa la ausencia total de certeza, y el 1 representa la certeza total (otras escalas lineales también podrían ser utilizadas en principio). Si interpretamos $\mu(A)$ como el grado de creencia en que el valor de una variable $X$ está en el conjunto $A$, entonces la propiedad (1) nos dice que la variable efectivamente toma un solo valor (¡el vacío indica que la variable no toma ningún valor!); la (2), que este valor se encuentra en el conjunto total $W$, y la (3), que la medida es monótona (no estricta) con respecto a la inclusión, es decir, la certeza de que la variable adquiera valor en un conjunto mayor será también mayor, o como mínimo igual, pero en ningún caso menor, que en un conjunto más pequeño.

De hecho, las propiedades 1), 2) y 3) son las propiedades mínimas que se requieren en una medida m con el fin de representar la incertidumbre. A partir de aquí, los diferentes modelos de incertidumbre que se pueden encontrar en la literatura se caracterizan por añadir algunas propiedades adicionales.

**Medidas de probabilidad**

Las medidas de probabilidad sobre un universo finito, que son las usuales en las aplicaciones, demandan una cuarta propiedad, la *aditividad*:

$$\text{4-P) } \mu(A \cup B) = \mu(A) + \mu(B), \text{ si } A \cap B = \varnothing$$

**Notación**

La cuarta propiedad tiene diferentes acepciones, notaremos:

4-P) Propiedad 4 – probabilidad.

que también puede ser expresada de manera equivalente como:

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$$

Esta importante propiedad permite reconstruir la medida de cualquier conjunto a partir de la medida de sus elementos, también conocida como *distribución de probabilidad*. En efecto, si conocemos la función $p$: $W \rightarrow [0, 1]$ definida por $p(x) = \mu(\{x\})$ para todo $x$ de $W$, entonces la aditividad asegura que, para cualquier subconjunto $A$, podemos calcular su probabilidad como:

$$\mu(A) = \Sigma_{x \in A} p(x)$$

En particular, la función de distribución cumple que: $1 = \mu(W) = \Sigma_{x \in W} p(x)$.

Otra consecuencia destacable de la aditividad consiste en la conocida relación entre la probabilidad de un conjunto y la de su complementario. En efecto, si $A^c$ denota el complementario de $A$ con respecto a $W$, entonces tenemos que:

$$1 = \mu(W) = \mu(A \cup A^C) = \mu(A) + \mu(A^C)$$

y por lo tanto:

$$\mu(A^C) = 1 - \mu(A)$$

Aparte de las probabilidades, existen otros tipos de medidas de incertidumbre que se han estudiado y que han dado lugar a diferentes modelos para el tratamiento de la incertidumbre. Dos clases de medidas alternativas son: las de posibilidad y necesidad, y las funciones de credibilidad y plausibilidad.

### Medidas de posibilidad y necesidad

Las llamadas *medidas de posibilidad* se caracterizan por la propiedad siguiente:

**Notación**

4-∏) Propiedad 4 – posibilidad.

$$4\text{-}\prod)\ \mu(A \cup B) = max(\mu(A), \mu(B)),$$

que permite descomponer la medida de una unión por la función *máximo*, mientras que las medidas de necesidad se caracterizan por la propiedad de descomposición de la intersección con respecto a la función *mínimo*:

$$4\text{-N})\ \mu(A \cap B) = min(\mu(A), \mu(B)).$$

Las medidas de posibilidad y de necesidad son duales, en el sentido de que si μ es una medida de necesidad entonces la medida μ* definida por:

$$\mu^*(A) = 1 - \mu(A^C),$$

es una medida de posibilidad, y al revés.

Las medidas de posibilidad también permiten calcular la medida de un conjunto cualquiera a partir de la medida de sus elementos. Efectivamente, si μ es una medida de posibilidad sobre un $W$ finito, y definimos $\pi(x) = \mu(\{x\})$ para todo elemento $x$ de $W$, entonces tenemos que:

$$\mu(A) = max_{x \in A}\pi(x),$$

y para la medida de necesidad dual μ* tenemos que:

$$\mu^*(A) = min_{x \notin A}1 - \pi(x).$$

A la función $\pi$: W → [0, 1] antes definida se la denomina *distribución de posibilidad* y cumple que $1 = \mu(W) = \max\{\pi(x) \mid x \in W\}$.

**El modelo posibilístico**

El modelo posibilístico se debe a trabajos de Zadeh (que inventó los Fuzzy Sets) y de Dubois y Prade.

Estas medidas son la base del *modelo posibilístico* para el tratamiento de la incertidumbre, relacionado con el uso de conjuntos borrosos (*fuzzy sets*) para modelizar la imprecisión.

Los sistemas difusos, que se describen en el capítulo segundo de este módulo, están basados en el modelo posibilístico.

### Medidas de credibilidad y de plausibilidad

Otra familia de medidas de incertidumbre duales, en el sentido anterior, son las llamadas *funciones de credibilidad* (*belief*) y de *plausibilidad*. Las funciones de credibilidad cumplen la propiedad siguiente:

**Notación**

4-B) Propiedad 4 – belief.

4-B)

$$\mu(A_1 \cup \ldots \cup A_n) \geq \Sigma_i\mu(A_i) - \Sigma_{ij}\mu(A_i \cap A_j) +$$

$$+ \Sigma_{i,j,k}\mu(A_i \cap A_j \cap A_k) \ldots \pm \mu(A_1 \cap \ldots \cap A_n)$$

para todo $n$. En particular para $n = 2$, tenemos la llamada *superaditividad*:

$$\mu(A \cup B) \geq \mu(A) + \mu(B), \text{ si } A \cap B = \varnothing.$$

Las medidas de plausibilidad cumplen la propiedad dual:

**Notación**

4-Pl) Propiedad 4 – plausibilidad.

4-Pl)

$$\mu(A_1 \cup \ldots \cup A_n) \leq \Sigma_i\mu(A_i) - \Sigma_{i,j}\mu(A_i \cap A_j) +$$

$$+ \Sigma_{i,j,k}\mu(A_i \cap A_j \cap A_k)\ldots \pm \mu(A_1 \cap \ldots \cap A_n)$$

y por lo tanto son medidas llamadas *subaditivas*. Esto es, verifican:

$$\mu(A \cup B) \leq \mu(A) + \mu(B), \text{ si } A \cap B = \varnothing.$$

La característica más importante de las funciones de credibilidad y de plausibilidad la constituye su representación en términos de las llamadas *asignaciones de masas*. Una asignación masa $m$ es una función $m$: $2^W \rightarrow [0, 1]$ tal que:

$$\Sigma_{A \subseteq W} m(A) = 1,$$

es decir, una masa no es más que una distribución de probabilidad sobre el conjunto de las partes. Entonces, μ constituye una función de credibilidad si, y sólo si, existe una asignación de masas $m$ tal que $m(\varnothing) = 0$ y:

$$\mu(A) = \Sigma_{B \subseteq A} m(B)$$

Análogamente, μ constituye una función de plausibilidad si, y sólo si, existe una asignación de masas $m$ tal que $m(\varnothing) = 0$ y,

$$\mu(A) = \Sigma_{B \cap A \neq \varnothing} m(B)$$

Los subconjuntos A tales que $m$(A) > 0 se denominan *elementos focales*.

Las medidas de credibilidad y plausibilidad constituyen la base del llamado *modelo evidencial de Dempster-Shafer* para modelizar la incertidumbre, que resulta ser un modelo más general que el probabilístico y el posibilístico. En efecto, una medida de probabilidad constituye al mismo tiempo un caso particular de función de credibilidad y de plausibilidad, mientras que las de necesidad y posibilidad son casos particulares de medidas de credibilidad y de plausibilidad respectivamente. En términos de masas, una probabilidad viene definida por una masa tal que sus elementos focales son *singletones* (conjuntos de un solo elemento), mientras que las medidas de posibilidad y necesidad lo son por aquellas masas tales que sus elementos focales forman un conjunto anidado.

**Lectura complementaria**

El modelo evidencial está basado en los trabajos de Shafer y de Demspter. G. Shafer (1976). *A Mathematical Theory of Evidence*. Princeton University Press, da una descripción de ello.

### Medidas de incertidumbre: intensionalidad frente a extensionalidad

Es muy importante remarcar que las medidas de incertidumbre son "intensionales" (o no son "extensionales" o "funcionales", equivalentemente) en el sentido de que no extienden las operaciones del álgebra de conjuntos a operaciones sobre el intervalo [0, 1] excepto si trivializamos la medida. Dicho de otra manera, para toda medida $\mu: 2^W \to [0, 1]$ de manera que para algún C de $2^W$ se tenga $0 < \mu(C) < 1$, no existen operaciones binarias $T$, $S$ y unaria $N$ sobre [0, 1] tales que:

$$\mu(A \cap B) = T(\mu(A), \mu(B))$$

$$\mu(A \cup B) = S(\mu(A), \mu(B))$$

$$\mu(A^C) = N(\mu(A))$$

para todo $A$, $B$ de $2^W$. Es decir, la única manera de tener funcionalidad total de la medida es que para todo $A$, o bien $\mu(A) = 1$, o bien $\mu(A) = 0$, es decir, que de hecho no tengamos incertidumbre.

#### Ejercicio de funcionalidad de la medida

Ejercicio: Demostrad que la única manera de tener total funcionalidad de la medida es que para todo $A$, o bien $\mu(A) = 1$ o bien $\mu(A) = 0$.

Solución: De entrada como la $\cap$ y la $\cup$ son operaciones conmutativas, asociativas y monótonas (con respecto a la inclusión), también lo deben ser la $T$ y la $S$. Además, como $A \cap \varnothing = \varnothing$ y $A \cap W = A$, entonces $T$ ha de ser tal que $T(0, x) = 0$ y $T(1, x) = x$. Análogamente, por razonamiento parecido, $S(0, x) = x$ y $S(1, x) = 1$. Como $A \cap A = A \cup A = A$, entonces $T$ y $S$ deben ser idempotentes, esto es, $T(x, x) = S(x, x) = x$, para todo $x$ de [0, 1]. De estas propiedades se deduce que forzosamente $T(x, y) = \min(x, y)$ y $S(x, y) = \max(x, y)$. En efecto, supongamos $x \le y$; entonces tenemos $x = T(x, x) \le T(x, y) \le T(x, 1) = x$, y de manera parecida tenemos $y = S(y, y) \ge S(y, x) \ge S(y, 0) = y$. Finalmente, tenemos que para todo $A$, $1 = \mu(W) = \mu(A \cup A^c) = \max(\mu(A), \mu(A^c))$, y a la vez $0 = \mu(\varnothing) = \mu(A \cap A^c) = \min(\mu(A), \mu(A^c))$, de donde las únicas soluciones posibles son $\mu(A) = 0$ (y $\mu(A^c) = 1$) o $\mu(A) = 1$ (y $\mu(A^c) = 0$).

Fijaos en que para las medidas antes mencionadas sólo tenemos, en general, funcionalidades parciales: las probabilidades sólo son funcionales con la complementación $\mu(A^*) = 1 - \mu(A)$, pero no lo son ni para la unión, ni para la intersección; las medidas de posibilidad son funcionales para la unión pero no lo son ni para la intersección ni para la complementación; y las medidas de necesidad son funcionales para la intersección pero no lo son ni para la unión ni para la complementación.

La no funcionalidad de las medidas de incertidumbre presenta importantes consecuencias desde el punto de vista computacional, dado que se puede perder la eficiencia de sistemas de razonamiento, como los basados en reglas cuando se les aumenta el poder expresivo con mecanismos genuinos de tratamiento de la incertidumbre. Por ejemplo, en el caso de las probabilidades, supongamos que en una base de conocimiento tenemos una regla del tipo:

si E entonces H

que resulta parcialmente cierta; sólo se da H cuando se observa E en un α% de los casos. Entonces podemos representar esta regla incierta por el par (regla, grado de certeza):

si E entonces H, α

donde α representa la probabilidad $P(H \mid E) = P(E, H) / P(E)$. Si de la información disponible ($E'$) podemos saber por ejemplo que $P(E \mid E') = P(E, E') / P(E) = \beta$, se esperaría poder decir algo sobre la probabilidad de $H$ a partir de α y de β, mediante una función $f$ de combinación como $f(\alpha, \beta)$. Pues, desgraciadamente no es así. Se necesita más información, una información que no es evidente que resulte necesaria a primera vista, como por ejemplo cuál es la probabilidad $P(H \mid E, E')$ y la probabilidad $P(H \mid \neg E, E')$.

**Ejercicio de cálculo de probabilidades**

**Exercici:** Calculad $P(H \mid E')$ en función de $P(H \mid E) = \alpha$, $P(E \mid E') = \beta$, $P(H \mid E, E') = \gamma$ y de $P(H \mid E^c, E') = \delta$.

**Solución:** Por la regla de las probabilidades totales, $P(H \mid E') = P(H, E \mid E') + P(H, E^c \mid E')$, y aplicando la regla de Bayes a cada factor tenemos $P(H \mid E') = P(H \mid E, E') \cdot P(E \mid E') + P(H \mid E^c, E') \cdot P(E^c \mid E') = \gamma \cdot \beta + \delta \cdot (1 - \beta)$.

Precisamente, estas dificultades fueron las causantes de la aparición en los años ochenta de sistemas expertos basados en reglas como PROSPECTOR y MYCIN como intentos de encontrar soluciones más o menos ingeniosas al problema de la no funcionalidad de las probabilidades, es decir, intentando hacer funcionales los cálculos de probabilidades bajo condiciones particulares. De esta manera, estos tipos de sistemas, denominados *extensionales*, eran computacionalmente eficientes, pero el precio que se debía pagar era un uso inadecuado del modelo probabilístico y por lo tanto la posible inexactitud de los resultados. Cabe señalar, sin embargo, que en aplicaciones particulares, los errores a causa del modelo utilizado se podían solventar mediante soluciones parciales y *ad hoc* en cada caso, ya sea retocando los pesos, las reglas, o bien mediante mecanismos de control propios del sistema, ajenos al propio modelo de representación. Ahora bien, estas soluciones no son nunca generales.

Por el contrario, los llamados *sistemas intensionales* son los que realizan un uso correcto de la semántica intensional de las diferentes medidas de incertidumbre. En general, y excepto contadas pero importantes excepciones, el precio que normalmente pagan estos sistemas es el de un alto coste computacional. En particular, los sistemas basados puramente en cálculo de probabilidades que no incorporan ningún tipo de mecanismo especial, como por ejemplo la lógica probabilística de Nilsson, son computacionalmente impracticables cuando el número de variables que involucran es grande. De hecho, los sistemas basados en probabilidad sólo han tenido éxito dentro de la comunidad de inteligencia artificial a partir de la aparición de las llamadas *redes bayesianas* o *redes causales* a finales de los años ochenta; las cuales, gracias a explotar las relaciones de independencia entre variables que quedan explícitas en el formalismo gráfico del modelo, permiten hacer operativo y eficiente razonar con probabilidades.

**Lectura complementaria**

Esta lógica es descrita en:
**N. J. Nilsson** (1986). "Probalistic Logic". *Artificial Intelligence* (núm. 28, págs. 71-87).

# 2. Sistemas difusos

**Vicenç Torra i Reventós**
**David Isern Alarcón**

Los sistemas difusos son sistemas basados en reglas en los que éstas se definen en términos de los llamados *conjuntos difusos*. Estos sistemas se utilizan principalmente en tareas de control o de modelización. En el primer caso, queremos controlar el comportamiento de un aparato. Esto significa que construimos un sistema a fin de que el comportamiento del aparato siga un esquema determinado. Son ejemplos de este tipo de aplicación los sistemas para controlar la temperatura de un horno, para controlar el enfoque de una máquina de fotografiar o la velocidad de un tren. En el segundo caso, el sistema debe perseguir o predecir una determinada variable. Son ejemplos de este tipo de sistema los estudios para la predicción del comportamiento de los mercados financieros.

De acuerdo con lo que explicamos en el módulo 3, el formato de las reglas de un sistema basado en reglas sigue la estructura siguiente:

**si** <premisa> **entonces** <conclusión>

Aquí, las afirmaciones sobre hechos que aparecen tanto en las premisas, como en las conclusiones corresponderán a conceptos *vagos*.

Para mostrar todos los elementos que constituyen los sistemas difusos, consideramos un sistema para controlar la duración de un programa de lavado de una lavadora. La duración del lavado tendrá en cuenta el nivel de suciedad de la ropa y el nivel de grasa. Se supone que el sistema, una vez ha interpretado el nivel de suciedad y el nivel de grasa, deduce la duración del programa según un conjunto de reglas que "gobiernan" el sistema de control. El objetivo de todo ello es calcular la duración efectiva del programa de lavado en minutos que tendría que implementar el sistema de control.

Siguiendo con el ejemplo, nuestro sistema se basará en reglas del estilo:

**si** *nivel_suciedad es poco* **y** *nivel_grasa es alto* **entonces** *duración es larga*

Como se verá durante el módulo, hay que describir cómo se definen las variables (*nivel de suciedad*, *nivel de grasa* o *duración*), qué significan y cómo se definen términos, como por ejemplo, *bajo*, *alto*, *larga*, asociados a cada una de las variables y cómo se interpreta esta regla dentro de un conjunto de reglas. Se verá cómo funciona un sistema experto difuso y, finalmente, cómo se interpreta la salida final del sistema.

El apartado empieza con la descripción de los conjuntos difusos y de las operaciones que en éstos se definen. Esto es necesario porque a cada elemento de las reglas le haremos corresponder un conjunto difuso evaluado para determinar si una regla se aplica o no. Después de ello, veremos cómo se construye un sistema difuso. Examinaremos las reglas y cómo estas definen el sistema. A continuación, mostraremos de qué manera se aplica una regla y cómo, a partir de la aplicación de todas las reglas, obtenemos la conclusión de todo el sistema para un estado concreto.

## 2.1. Definición de los conjuntos difusos y operaciones

Los conjuntos difusos generalizan los conjuntos clásicos (para diferenciar ambos conjuntos, a partir de ahora llamaremos *conjuntos nítidos* a éstos últimos). Por esta razón, muchas de las operaciones que se aplican a los conjuntos difusos no constituyen más que generalizaciones de las que aplicamos a los conjuntos nítidos. En esta sección, veremos cómo se definen los conjuntos difusos y algunas de las operaciones que se han definido. Veremos cómo estas operaciones corresponden a las operaciones que ya conocemos de los conjuntos nítidos.

**Generalización**

Hablamos de generalización porque los conjuntos nítidos (los booleanos) son un caso particular de los difusos.

Para representar los conjuntos, existen varias maneras (representación extensional, función característica, elementos que satisfacen una propiedad), de entre éstas la que nos resulta más adecuada en este momento es la que utiliza la función característica. Esta función declara para cada elemento de un dominio de referencia qué elementos son del conjunto y cuáles no lo son. Esto es, se define una función booleana que cuando se aplica a los elementos del dominio, devuelve su pertenencia (o no pertenencia) al conjunto.

Para poder pasar fácilmente al caso difuso, utilizamos una función característica que retorna 0 ó 1: 0 cuando el elemento no pertenece al conjunto y 1 cuando sí que pertenece. De esta manera, formalmente tenemos que si $D$ es el dominio de referencia donde se define el conjunto $A$ mediante una función característica $\chi_A = A$, aleshores tenim que $\chi_A$ es una función de $D$ en el conjunto {0, 1} (esto es, $\chi_A$: $D \rightarrow \{0, 1\}$). Al conjunto $D$ lo denominamos *dominio de A* o bien *universo de discurso de A*.

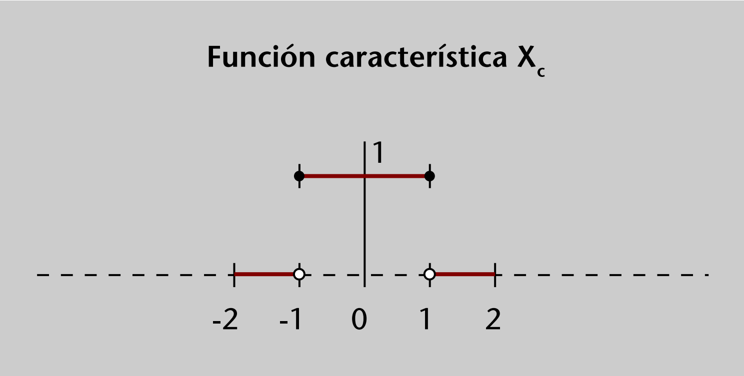
**Ejemplo de función característica**

Tomemos el conjunto definido por los puntos del intervalo [-1, 1]. Entonces, la función característica de este conjunto (que denominamos $C$ –de cero–) es:

$$X_C(x)\begin{cases} 1 & \text{si } x \in [-1, 1] \\ 0 & \text{si } x \notin [-1, 1] \end{cases}$$

Dado un elemento del dominio (un real r) podemos ver si pertenece o no al conjunto aplicando la función $\chi_C$. Así, cuando $\chi_C(r) = 1$ querrá decir que pertenece al conjunto y cuando $\chi_C(r) = 0$ querrá decir que no pertenece a éste. Claro está que 0,5 pertenece (porque $\chi_C$ (0,5) = 1) y, en cambio, 2,5 no pertenece ($\chi_C$(2,5) = 0). En la figura 1 se da la representación gráfica de la función característica de este conjunto.

Figura 1




En los conjuntos difusos, en lugar de tener que un elemento pertenece a un conjunto o no pertenece a éste lo que tenemos es una pertenencia graduada. Así, mientras que en el ejemplo anterior tenemos que se pasa de un elemento que pertenece a un conjunto (el valor 1 pertenece al conjunto) a uno que no (el valor 1,001 no pertenece) de manera abrupta, en un conjunto difuso este cambio se puede efectuar suavemente.

Para poder modelizar este cambio suave tenemos que la función en lugar de tomar valores en {0, 1}, los tomará en [0, 1] (en lugar de tener que los elementos o están o no están en el conjunto, ahora los elementos podrán estar en un cierto grado). De este modo, tendremos que si un elemento toma el valor 1 pertenece completamente al conjunto, si tomamos el valor 0,9 pertenece pero menos, si toma el valor 0,1 casi no pertenece y si toma el valor 0 no pertenece en absoluto al conjunto.

**Función de pertenencia**

La función que define un conjunto difuso se denomina *función de pertenencia*[1] (y la denotaremos por μ). Esta función, dado un elemento del dominio, retornará el correspondiente valor en [0, 1]. Por lo tanto, tenemos que una función de pertenencia μ sobre el dominio D es de la forma μ: D → [0, 1].

(1)Función de pertenencia proviene del inglés *membership function.*

Los conjuntos difusos permiten definir conceptos en los que no resulta fácil (o no es posible) definir claramente un punto donde separar entre aquellos elementos que son del concepto y aquellos que no lo son. Algunos ejemplos clásicos son los relacionados con la temperatura (frío o caliente), la altura (alto o bajo), la longitud (corto, largo o extralargo) o cualquier variable que requiera valores ordenados entre ellos (muy bajo, bajo, medio, alto, muy alto).

**Ejemplo de conjunto difuso**

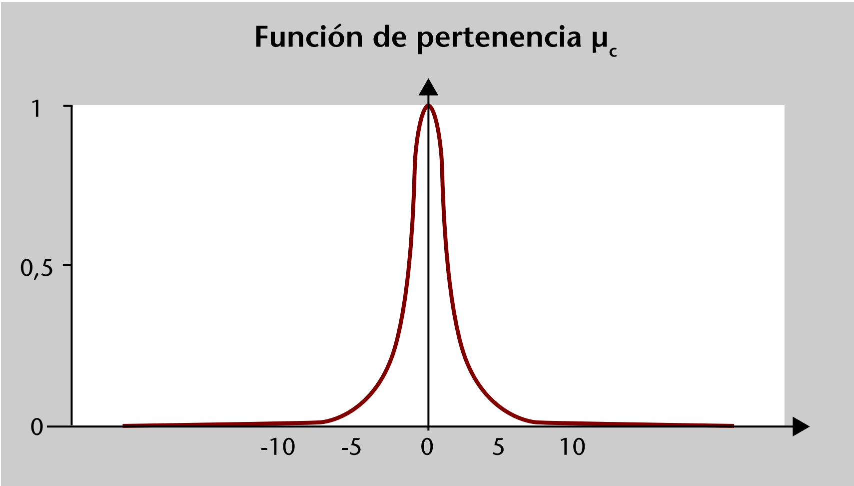
Definimos el conjunto difuso “aproximadamente cero” mediante la función de pertenencia $\mu_c$ siguiente:

$$\mu_C(x) = 1/(1 + x^2)$$

Dado un elemento del dominio (un real r) podemos ver si pertenece al conjunto o no aplicando la función $\mu_C$. El valor de 0 pertenece plenamente al conjunto (observad que $\mu_C(0) = 1$) y a medida que nos alejamos del 0, la pertenencia es cada vez menor (por ejemplo, $\mu_C(0{,}5) = 0{,}8$ y $\mu_C(1) = 1/(1 + 1) = 0{,}5$).

La función de pertenencia correspondiente a este conjunto se muestra en la figura 2.

Figura 2




Hay diferentes formas de definir las funciones de pertenencia. Como se ha visto en el ejemplo anterior, se ha utilizado una función cuadrática para definir el conjunto “aproximadamente cero”. La forma de las funciones de pertenencia define el comportamiento que después tendrá el sistema, por lo que normalmente se utilizan formas sencillas en vez de complejas, pues tampoco aportan mucha más precisión al sistema.

Las funciones de pertenencia más utilizadas son las funciones triangulares, las trapezoidales y las gaussianas (ved figura 3).

Una función triangular se define de la forma siguiente:

$$\mu_A(x) = \begin{cases} 0, & x \leq a \\ (x-a)/(m-a) & a < x \leq m \\ (b-x)/(b-m), & m < x < b \\ 0, & x \geq b \end{cases}$$

Donde $a$ es el umbral inferior, $b$ el umbral superior y $m$ el valor donde la función vale 1 y se cumple que $a < m < b$.

Una función trapezoidal se define por cuatro puntos (*a*, *b*, *c*, *d*): un umbral inferior *a*, un umbral superior *d* y un rango de apoyo entre dos puntos *b* y *c* donde la función de pertenencia tiene un valor 1. Considerando que $a < b < c < d$ tenemos:

$$\mu_A(x) = \begin{cases} 0, & x < a \\ (x-a)/(b-a) & a \le x \le b \\ 1, & b \le x \le c \\ (d-x)/(d-c), & c \le x \le d \\ 0, & x > d \end{cases}$$

Dentro de las funciones trapezoidales, hay un par de casos particulares, denominados *funciones R*[2] y *funciones L*[3].

(2)Del inglés *R-functions*.

(3)Del inglés *L-functions*.

Las funciones *R* consideran los valores $a = b = -\infty$, y se definen como:

$$\mu_A(x) = \begin{cases} 0, & x > d \\ (d-x)/(d-c), & c \le x \le d \\ 1, & x < c \end{cases}$$

Las funciones *L* consideran los valores $c = d = \infty$, y se definen como:

$$\mu_A(x) = \begin{cases} 0, & x < a \\ (x-a)/(b-a), & a \le x \le b \\ 1, & x > b \end{cases}$$

Las funciones campana[4] se definen de la forma siguiente:

(4)Del inglés *bell function*.

$$\mu(x; a, b, c) = \frac{1}{1 + \left|\frac{x-c}{a}\right|^{2b}}$$

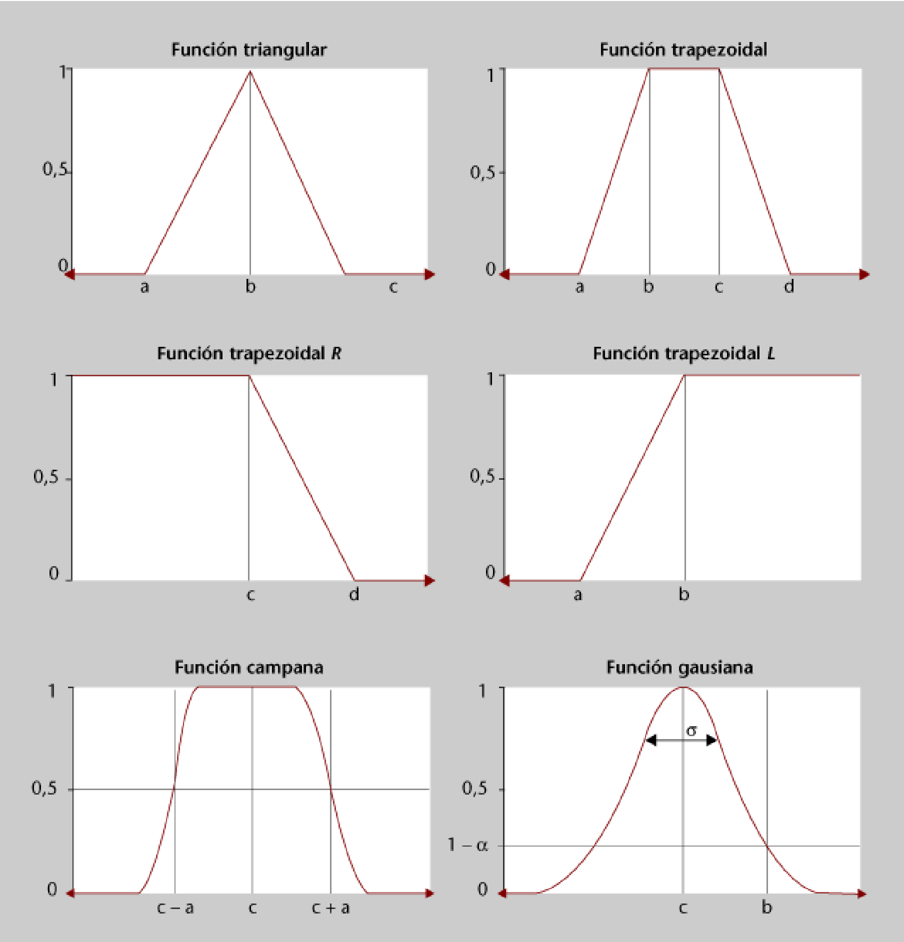
Donde *c* marca el punto central de la función, *a* representa la anchura y *b* (junto con *a*) controla la pendiente definida como $-b/2a$.

Las funciones gaussianas se definen de la forma seguiente:

$$\mu(x; c, \sigma) = e^{-\frac{(x-c)^2}{2\sigma 2}}$$

Donde $\sigma$ representa la desviación estándard i *c* representa la media.

Figura 3




**Ejemplo de conjunto difuso: altura**

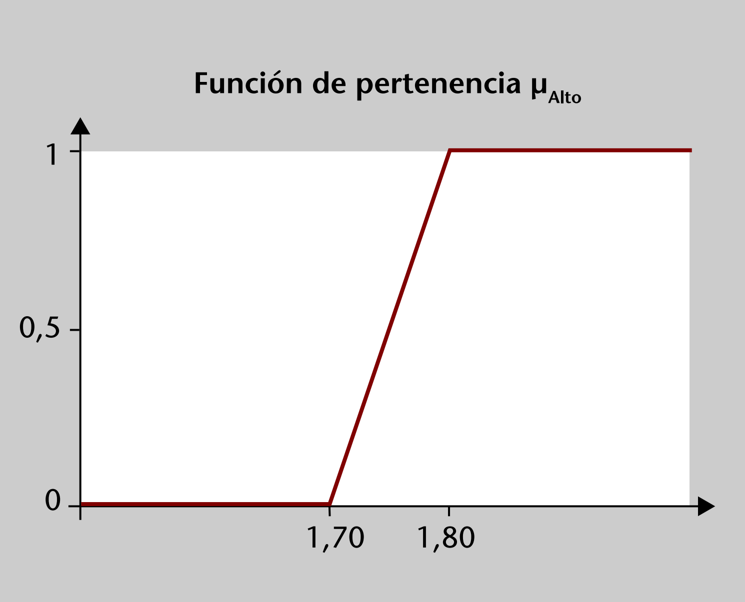
Si queremos definir el concepto alto podemos afirmar que a partir de una cierta altura –por ejemplo 1,80 m– una persona es alta (satisfará plenamente el concepto de alto) pero esto no significa que si mide 1,79 m entonces no sea alta en absoluto. Mediante un conjunto difuso, podemos modelizar el concepto alto de manera que el paso a ser alto es graduado (así podemos decir que una persona que mide 1,79 m es casi alta).

Definimos el concepto alto con una función trapezoidal *L* de la forma siguiente:

$$\mu_{\text{Alto}}(x) = \begin{cases} 0, & x < 1{,}70 \\ (x - 1{,}70)\,/\,(1{,}80 - 1{,}70), & 1{,}70 \le x \le 1{,}80 \\ 1, & x > 1{,}80 \end{cases}$$

En este caso, $\mu_{\text{Alto}}(1{,}79) = 0{,}9$. Por lo tanto, una persona de 1,79 casi satisfará completamente el concepto de alto. La figura presenta la función de pertenencia de este conjunto.

Figura 4




Notad que el hecho de que una persona que mida 1,79 m sea alta constituye una cuestión de grado (en este caso 0,9 –es decir, es cierto que es alta–) y es diferente del concepto de probabilidad. Es diferente considerar que una persona tiene una probabilidad de ser alto de 0,9. Si de una persona que no sabemos la altura se nos dice que la probabilidad de que sea alta es de 0,9 puede ocurrir que al conocerla encontremos que mide 1,50. Éste sería el caso si sabemos que la persona pertenece a un grupo de un cursillo de baloncesto donde se han inscrito diez alumnos y nueve miden más de 1,90 m y uno mide 1,50 m. Si, en cambio, se nos dice que la persona pertenece al concepto alto con grado 0,9 no nos podemos encontrar que mida 1,50 m (1,50 m pertenece con grado 0 al concepto alto).

### Operaciones

Resulta bien conocido que sobre los conjuntos existen varias operaciones. Así, si A y B son conjuntos se puede calcular la unión ($A \cup B$), la intersección ($A \cap B$), el complemento ($\neg A$) o la cardinalidad ($|A|$). También podemos comprobar si un conjunto es un subconjunto de otro ($A \subset B$) o si un elemento a pertenece a un conjunto ($a \in A$).

**Generalización**

Hablamos de generalizar las operaciones nítidas porque cuando las aplicamos a conjuntos nítidos nos devolverán lo mismo que si aplicáramos directamente operaciones nítidas. La diferencia consiste en que ahora dispondremos de operaciones para los conjuntos difusos.

Todas estas operaciones se pueden generalizar para trabajar con conjuntos difusos. Así, tendremos unión, intersección, etc. para conjuntos difusos. En esta sección, veremos algunas de estas operaciones. A la hora de definirlas se ha de tener en cuenta que para que las operaciones resulten correctas es necesario que cuando se aplican a conjuntos nítidos su comportamiento corresponda al de las operaciones correspondientes de los conjuntos nítidos.

**¿Qué quiere decir aquí “correcto”?**

Con “correcto” queremos decir que su significado sea coherente con el hecho de que los conjuntos difusos generalizan los conjuntos nítidos.

### Complementación

Empecemos considerando la operación de complementación. Esta función dado un conjunto difuso en un dominio D nos permite construir su complementario. En general, suponemos que si $\mu_A$ es un conjunto difuso, su complementario será $\mu_{no\text{-}A}$.

**Ejemplo de complemento de conjunto nítido**

Cuando $C$ corresponde al conjunto de puntos en el intervalo [-1, 1] que hemos visto antes y $X_C$ corresponde a su función característica, entonces $X_{no\text{-}C}$ se debe definir como sigue (notad que ésta es la función que corresponde a no pertenece al intervalo [-1, 1]):

$$\chi_{\text{no-C}}(x) = \begin{cases} 1, & x \notin [-1,\ 1] \\ 0, & x \in [-1,\ 1] \end{cases}$$

Para construir el complementario de un conjunto difuso cualquier $\mu_A$ sobre D tendremos en cuenta, evidentemente, la función de pertenencia $\mu_A$. De hecho, para definir la pertenencia de un elemento $x$ cualquiera del dominio al conjunto complementario, consideraremos sólo la pertenencia de $x$ al conjunto inicial (es decir, es $\mu_A(x)$). Así, si $\mu_A(x) = 1$, entonces tendremos que este elemento $x$ no ha de pertenecer al conjunto complementario (por lo tanto, debemos tener que $\mu_{no\text{-}A}(x) = 0$). De manera análoga, si $\mu_A(x) = 0$, entonces tendremos que este elemento $x$ sí que ha de pertenecer al complementario ($\mu_{no\text{-}A}(x) = 1$).

De acuerdo con esto podemos escribir que la pertenencia del elemento $x$ al complementario ($\mu_{no\text{-}A}(x)$) es una función (que denominamos $N$ *–de función de negación–*) del valor $\mu_A(x)$. Esta función, cuando $\mu_A(x) = 1$ retornará 0 y cuando $\mu_A(x) = 0$ devolverá 1. Por lo tanto, $N(1) = 0$ y $N(0) = 1$.

**Conjuntos y lógica estrechamente relacionados**

Se llama de negación porque igual que en el caso nítido, los conjuntos y la lógica se encuentran estrechamente relacionados. Así, el complemento está relacionado con la negación lógica.

Así, podemos decir:

$$\mu_{no\text{–}A}(x) = N(\mu_A(x)),$$

con $N(1) = 0$ y $N(0) = 1$. Estas dos igualdades (condiciones sobre la función $N$) son las llamadas *condiciones de contorno de N*.

Sin embargo, con esto no tenemos suficiente. Mientras que en el caso nítido esto resulta suficiente porque la pertenencia del conjunto es booleana, en el caso difuso la función $\mu_A$ puede recorrer todo el intervalo [0, 1]. Por lo tanto, hemos de considerar $N$ como una función que toma todos los valores del intervalo [0, 1]. Además, como la pertenencia a un conjunto es graduada, podemos considerar que la pertenencia al conjunto complementario también debe ser graduada. De acuerdo con esto, la función $N$ ha de ser una función que tome valores en [0, 1] y que los devuelva en [0, 1]. Además, para que cambios suaves de la pertenencia en el conjunto original no provoquen cambios bruscos en la pertenencia en el conjunto complementario, a menudo, se exige la continuidad de la función $N$. Todo esto lo escribiremos diciendo que $N$: $[0, 1] \rightarrow [0, 1]$ y que es continua.

Además de las condiciones que han aparecido, existen otras condiciones que también son razonables para la función $N$. Una de éstas es la condición de monotonía. Esta condición se expresa como:

para todo $a$, $b$ en [0, 1], si $a \leq b$ entonces $N(a) \geq N(b)$.

**Ejemplo de condición de monotonía: alturas**

Para interpretar esta condición, consideremos las alturas de dos chicas, Alba y Blanca. Supongamos que Alba mide 1,72 y que Blanca mide 1,75. Entonces tenemos que, de acuerdo con la definición del concepto de alto que hemos dado más arriba, $\mu_{Alto}$(Alba) = 0,2 y $\mu_{Alto}$(Blanca) = 0,5.

Si definimos $a = \mu_{Alto}$(Alba) = 0,2 y $b = \mu_{Alto}$(Blanca) = 0,5, podemos ver que la condición anterior nos dice que ya que $a \leq b$, entonces $N(a) \geq N(b)$. Esto es equivalente a decir $N(\mu_{Alto}$(Alba)) $\geq N(\mu_{Alto}$(Blanca)) que significa que al definir el complementario de alto (lo podemos interpretar como una aproximación tosca de bajo), Alba será más baja que Blanca.

Si consideramos el conjunto complementario del conjunto complementario, podemos incluir una nueva condición para la función $N$: podemos exigir que el conjunto complementario del conjunto complementario sea el mismo conjunto original. En este caso necesitamos que si se aplica la función N dos veces retorne el valor original:

**Igual que en los conjuntos nítidos**

Notad que esto es lo que ocurre con los conjuntos nítidos. El complementario de [-1, 1] es $(-\infty, -1) \cup (1, \infty)$ y el complementario de este último conjunto es [-1, 1].

$N(N(a)) = a$ para cualquier valor de a en [0, 1].

Tened en cuenta que si el complementario del conjunto $\mu_A$ es $\mu_{no\text{-}A}(x) = N(\mu_A(x))$, su conjunto complementario será $\mu_{no\text{-}no\text{-}A}(x) = N(\mu_{no\text{-}A}(x)) = N(N(\mu_A(x)))$. Por lo tanto, si queremos que $\mu_{no\text{-}no\text{-}A}(x) = N(N(\mu_A(x))) = \mu_A(x)$, necesitaremos que para todo $a = \mu_A(x)$ se cumpla $N(N(a)) = a$. A esta condición se la denomina *involución*.

Si consideramos todas las propiedades que hemos ido planteando podemos decir que una función $N$: [0, 1] → [0, 1] continua que satisface las condiciones:

**1)** $N(0) = 1$ y $N(1) = 0$ (condiciones de contorno).

**2)** Para todo $a$, $b$ en [0, 1], si $a \leq b$ entonces $N(a) \geq N(b)$ (monotonía).

**3)** Para todo $a$ en [0, 1], $N(N(a)) = a$ (involución), es una función de complementación.

No existe una única función que satisfaga todas estas condiciones, sino que hay infinitas. Esto provoca que mientras para los conjuntos nítidos existe una única manera de construir el complementario, en los conjuntos difusos hay tantas como funciones $N$ que satisfacen las condiciones anteriores. Para cada función que elijamos, tendremos una complementación.

**Funciones de complementación**

Demos dos familias de funciones de complementación:

**1)** $N_\lambda(a) = (1 - a)/(1 + \lambda a)$ para cualquier $\lambda > -1$ (familia de Sugeno).

**2)** $N_w(a) = (1 - a^w)^{1/w}$ para cualquier $w > 0$ (familia de Yager).

La primera es la llamada *familia de Sugeno* y la segunda es la llamada *familia de Yager*. Con $\lambda = 0$ en el primer caso y $w = 1$ en el segundo caso se obtiene la negación $N(a) = 1 - a$ que es la que se utiliza de manera más frecuente. Notad que en estas funciones siempre se cumple $N(1) = 0$ y $N(0) = 1$. Por lo tanto, se generaliza el caso nítido.

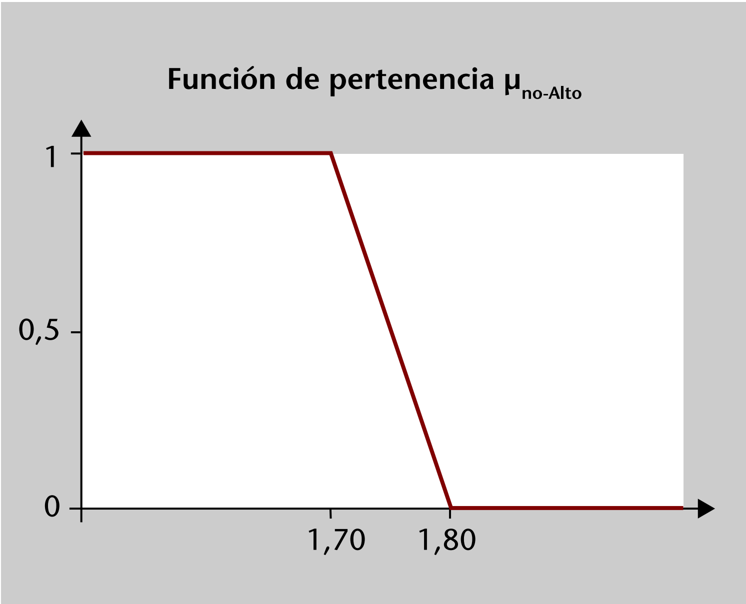
**Ejemplo de las alturas**

Utilizando la complementación definida por Yager con $w = 1$ o por Sugeno con $\lambda = 0$, tenemos $N(a) = 1 - a$. Así, obtenemos que el complementario del conjunto $\mu_{Alto}$ será:

$$\chi_{\text{no-Alto}}(x) = \begin{cases} 1, & \text{si } x < 1{,}70 \\ 1 - ((x - 1{,}70) / (1{,}80 - 1{,}70)) & \text{si } 1{,}70 \leq x \leq 1{,}80 \\ 0, & \text{si } x > 1{,}80 \end{cases}$$

La figura 5 ofrece la representación gráfica de esta función de pertenencia.

Figura 5




## Unión de conjuntos de conjuntos difusos

Pasemos ahora a la unión de conjuntos (empecemos considerando la unión de sólo dos conjuntos). El procedimiento para definir la operación para conjuntos difusos será análogo al caso de la complementación teniendo en cuenta que, al tener dos conjuntos (los llamamos A y B), habrá dos funciones de pertenencia (las denominamos $\mu_A$ y $\mu_B$). Entonces, construimos la función de pertenencia del conjunto unión, que denominamos $\mu_{A \cup B}$, a partir de las dos funciones $\mu_A$ y $\mu_B$. De manera parecida a lo que ocurre en el caso nítido, para la definición de la pertenencia de un elemento $x$ al conjunto unión sólo consideraremos la pertenencia o no de este elemento a los conjuntos que unimos. Así, cuando o bien $\mu_A(x)$ sea 1 o bien sea $\mu_B(x)$, diremos que $\mu_{A \cup B}(x) = 1$. De acuerdo con esto, y de manera parecida al caso de la complementación, podemos decir que $\mu_{A \cup B}(x)$ es una función de $\mu_A(x)$ y de $\mu_B(x)$. Denotando esta función para $S$ (esta función se llama *t-conorma*) podemos escribir:

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

Como tanto $\mu_A(x)$ como $\mu_B(x)$ corresponden a valores en el intervalo [0, 1] ($\mu_A$ y $\mu_B$ son funciones de pertenencia) y como $\mu_{A\cup B}(x)$ debe ser también un valor en este intervalo ($\mu_{A\cup B}$ es también una función de pertenencia), tendremos que $S$ es una función que dados dos valores en el intervalo [0, 1] retorna otro en el mismo intervalo. Por lo tanto, $S$: $[0, 1] \times [0, 1] \rightarrow [0, 1]$.

A esta función se le exigen habitualmente las cuatro condiciones siguientes:

**1)** $S(a, b) = S(b, a)$ (conmutatividad).

**2)** $S(a, S(b, c)) = S(S(a, b), c)$ (asociatividad).

**3)** Si $a \leq c$ y $b \leq d$ entonces $S(a, b) \leq S(c, d)$ (monotonía).

**4)** $S(a, 0) = a$ (elemento neutro).

donde $a$, $b$, $c$ y $d$ son valores cualesquiera del intervalo [0, 1].

Teniendo en cuenta que los valores $a$, $b$, $c$ y $d$ corresponden a valores de funciones de pertenencia, estas condiciones se pueden interpretar de la manera siguiente:

**1)** Se debe cumplir $S(\mu_A(x), \mu_B(x)) = S(\mu_B(x), \mu_A(x))$ porque la unión es conmutativa:

$$A \cup B = B \cup A.$$

**2)** Si consideramos la unión de tres conjuntos difusos $\mu_A(x)$, $\mu_B(x)$ y $\mu_C(x)$, el resultado no ha de depender del orden de cómo hacemos las uniones dos a dos. La unión es asociativa, por lo tanto debe cumplir:

$$S(\mu_A(x), S(\mu_B(x), \mu_C(x))) = S(S(\mu_A(x), \mu_B(x)), \mu_C(x))$$

Como al exigir la asociatividad, el orden en el que agrupamos los elementos no afecta al resultado, cuando haya $n$ conjuntos difusos para unir escribiremos sin agrupar $S(a_1, ..., a_n)$.

**3)** Cuanto mayor sea el grado de certeza correspondiente a los dos conjuntos que unimos, mayor será el grado de certeza del conjunto unión.

**4)** La unión de un conjunto $\mu_A$ con el conjunto vacío (el que satisface $\mu_\varnothing(x) = 0$ para todo $x$) es $\mu_A$. Por lo tanto, se ha de cumplir:

$$S(\mu_A(x), \mu_\varnothing(x)) = S(\mu_A(x), 0) = \mu_A(x)$$

Se puede comprobar que las cuatro condiciones dadas más arriba sobre la *t*-conorma *S* implican a las condiciones que se han de cumplir de acuerdo con las propiedades de la unión de conjuntos:

**Deducción**

Podéis ver que $S(0, 0) = 0$ y $S(1, 0) = 1$ se deducen de la cuarta condición haciendo $a = 0$ y $a = 1$, respectivamente. $S(0, 1) = 1$ se obtiene por conmutatividad en $S(1, 0) = 1$. La igualdad $S(1, 1) = 1$ es implicada por la condición de monotonía y porque sabemos que $S(1, 0) = 1$ (no puede ser $S(1, 1) = 0$ si $S(1, 0) = 1$).

$$S(0, 0) = 0,\ S(1, 0) = 1,\ S(0, 1) = 1 \text{ y } S(1, 1) = 1.$$

Como ya ocurría en el caso de la función *N* para construir el conjunto complementario, no existe una única *t*-conorma *S* que satisfaga las condiciones anteriores. Como ejemplo damos una familia de *t*-conormas (la familia de Yager) y otras tres *t*-conormas muy conocidas:

**1)** $S(a, b) = \min(1, (a^w + b^w)^{1/w})$ para $w > 0$ (familia de Yager).

**2)** $S(a, b) = \max(a, b)$ (máximo).

**3)** $S(a, b) = a + b - ab$ (suma algebraica).

**4)** $S(a, b) = \min(1, a + b)$ (Lukasiewitz o suma acotada).

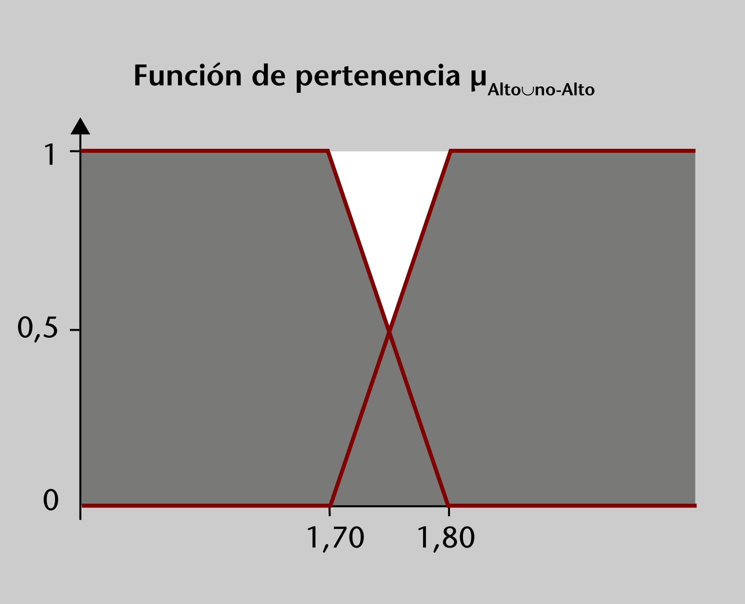
**Ejemplo de unión**

Damos la resultante de los conjuntos difusos $\mu_{\text{Alto}}$ y $\mu_{\text{no-Alto}}$ cuando la *t*-conorma elegida es el máximo:

$$\chi_{\text{Alto}\cup\text{no-Alto}}(x) = \begin{cases} 1, & \text{si } x < 1{,}70 \\ 1 - ((x - 1{,}70) / (1{,}80 - 1{,}70)) & \text{si } 1{,}70 \le x \le 1{,}75 \\ (x - 1{,}70) / (1{,}80 - 1{,}70) & \text{si } 1{,}75 \le x \le 1{,}80 \\ 1, & \text{si } x > 1{,}80 \end{cases}$$

la representación gráfica de $\mu_{\text{Alto} \cup \text{no-Alto}}$ está en la figura.

Figura 6




**Intersección de conjuntos difusos**

La definición de la intersección de conjuntos difusos es similar a la definición de la unión. Este caso se basa en la definición de una función denominada *t-norma* y que denotaremos *T*. Con esta función, si disponemos de dos funciones de pertenencia $\mu_A(x)$ y $\mu_B(x)$, podemos construir su intersección haciendo:

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

La función *T* será una función $T$: $[0, 1] \times [0, 1] \rightarrow [0, 1]$ que satisface, igual que la *t*-conorma, la conmutatividad, la asociatividad y la monotonía. Además, satisface la condición del elemento neutro pero ahora el elemento neutro es el 1. Por lo tanto, tenemos que las *t*-normes satisfacen:

**1)** $T(a, b) = T(b, a)$ (conmutatividad).

**2)** $T(a, T(b, c)) = T(T(a, b), c)$ (asociatividad).

**3)** Si $a \leq c$ y $b \leq d$ llavors $T(a, b) \leq T(c, d)$ (monotonía).

**4)** $T(a, 1) = a$ (elemento neutro).

Donde *a*, *b*, *c* y *d* son valores cualesquiera del intervalo [0, 1].

La condición del uno como elemento neutro se obtiene si consideramos que la intersección de un conjunto $\mu_A$ con aquél que contiene todos los elementos del dominio es el mismo conjunto $\mu_A$. En este caso, si el conjunto que contiene todos elementos del dominio D lo denotamos por $\mu_D(x) = 1$, tendremos que exigir $\mu_{A \cap D} = \mu_A$ corresponde a pedir: $T(\mu_A(x), \mu_D(x)) = T(\mu_A(x), 1) = \mu_A(x)$.

**El orden no afecta al resultado**

De manera análoga al caso de la *t*-conorma, como la asociatividad de la *t*-norma implica que el orden en el que agrupamos los elementos no afecta al resultado, cuando haya *n* conjuntos difusos para intersecionar escribiremos $T(a_1, ..., a_n)$.

Las otras condiciones son iguales que en el caso de la unión de conjuntos difusos.

En este caso, también tenemos más de una función que satisface las propiedades de las *t*-normas. Como ejemplo, consideramos éstas:

**1)** $T(a, b) = 1 - \min(1, [(1 - a)^w + (1 - b)^w]^{1/w})$ para $w > 0$ (familia de Yager).

**2)** $T(a, b) = \min(a, b)$ (mínimo).

**3)** $T(a, b) = ab$ (producto algebraico).

**4)** $T(a, b) = \max(0, a + b - 1)$ (Lukasiewitz o diferencia acotada).

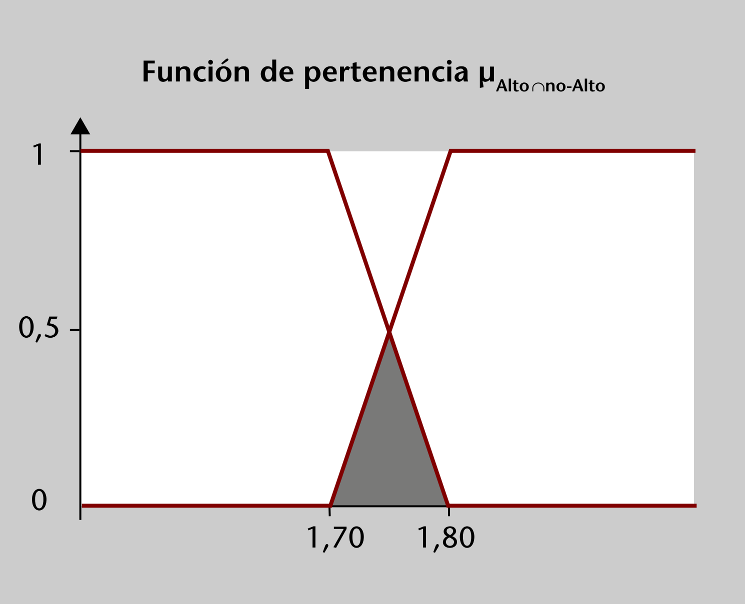
**Ejemplo de intersección**

Consideremos la intersección del conjunto difuso $\mu_{\text{Alto}}$ y el conjunto $\mu_{\text{no-Alto}}$. En este caso, usando la *t*-norma mínima obtenemos:

$$\chi_{\text{Alto} \cup \text{no-Alto}}(x) = \begin{cases} 0, & \text{si } x < 1{,}70 \\ (x - 1{,}70) / (1{,}80 - 1{,}70) & \text{si } 1{,}70 \le x \le 1{,}75 \\ 1 - ((x - 1{,}70) / (1{,}80 - 1{,}70)) & \text{si } 1{,}75 \le x \le 1{,}80 \\ 0, & \text{si } x > 1{,}80 \end{cases}$$

la representación gráfica de $\mu_{\text{Alto} \cap \text{no-Alto}}$ está en la figura 7).

Figura 7




Aquí, hemos definido las funciones *N*, *T* y *S* para construir el conjunto difuso complementario, el conjunto difuso intersección y el conjunto difuso unión. Sin embargo, el uso de estas funciones es más general. La función *N* corresponde a la función de negación booleana (¬), la *t*-norma *T* corresponde a una conjunción (^) y la *t*-conorma *S* orresponde a una disyunción (v) cuando los valores de certeza no son sólo cierto y falso, sino que tenemos que la certeza es difusa. Esto liga con la correspondencia que existe entre proposiciones y conjuntos. De hecho, las condiciones que hemos exigido para *N*, *S* y *T* se pueden entender también desde este punto de vista.

Una vez sabemos cómo definir conjuntos difusos y tenemos las operaciones de complementación, unión e intersección (y, por lo tanto, la negación, disyunción y conjunción), podemos pasar a ver cómo se construye un sistema difuso.

## 2.2. Construcción de un sistema difuso

Cuando se construye un sistema difuso (o, en general, cualquier sistema basado en el conocimiento), se ha de tener en cuenta que el conjunto de reglas que hay en el sistema debe cubrir todo dominio que queremos controlar o modelizar. Por ejemplo, siguiendo con el ejemplo del módulo de un sistema de control de la duración del programa de lavado de una lavadora, debemos definir las reglas que tendremos en cuenta en el sistema de control. Ahora bien, puede ser que tengamos claras todas las posibilidades o puede que se trate de un subconjunto. En el caso de que dadas unas entradas no se active ninguna regla, el sistema no dará ningún resultado. La zona definida por una combinación de las variables de entrada (partición) que tenga al menos 1 regla la denominamos *dominio de aplicación*.

**Ved también**

Recordad que en el apartado 1.1.1 del módulo "Sistemas basados en el conocimiento" se dijo que la base de conocimientos debe contener todo el conocimiento que el sistema necesita para resolver los problemas previstos. En este caso necesitamos que haya suficientes reglas para tratar todas las situaciones que se preveen.

Una manera de asegurarnos de que el sistema trata todos los posibles casos es definir una partición del dominio que se va a tratar y asociar, como mínimo, una regla a cada elemento de la partición.

En los sistemas difusos, construiremos las reglas de forma que su premisa se cumpla sólo en la región que le corresponde.

Si seguimos esta estructura, lo que tendremos, en general, en la premisa es un conjunto de variables de entrada $x_1$, ..., $x_n$ donde cada $x_i$ está relacionada con un término $t_{i,a}$ que corresponde a una parte del dominio de la variable $x_i$. Esto es, tenemos una regla de la forma:

**si** $x_1$ es $t_{1,a}$ **y** $x_2$ és $t_{2,b}$ **y** $x_3$ es $t_{3,c}$ **y** ... **y** $x_n$ es $t_{n,z}$ **entonces** $Y$ es $t_{Y,o}$

De forma general, las reglas pueden contener conjunciones, disyunciones o negaciones. En todos los casos se irán aplicando los operadores de los conjuntos difusos que se han ido viendo. En el caso de las conjunciones, hablaremos de intersección; en el caso de la disyunción, hablaremos de unión y, en el caso de la negación, hablaremos del complementario.

Como la construcción de los sistemas difusos recae en el concepto de variable (lingüística), pasamos ahora a su definición. Veremos que para cada variable tendremos asociado un dominio (un universo de discurso) y un término que tiene asociado un subconjunto dentro de este dominio. Para que las reglas del sistema se puedan aplicar según lo que se ha expuesto más arriba, es necesario que los subconjuntos asociados a los términos definan una partición del dominio de las variables.

### Variables lingüísticas

Informalmente, podemos decir que una variable lingüística (por ejemplo, nivel de suciedad, de altura o de temperatura) es la que toma como valores palabras en lenguaje natural (por ejemplo, frío, caliente, alto, bajo) que se interpretan más tarde en un dominio de referencia mediante conjuntos difusos (por ejemplo, caliente corresponde a temperaturas superiores a unos 40 grados).

> "Linguistic variable: a variable whose values are words or sentences in a natural or artificial language."
>
> L. A. Zadeh (1975). "The concept of a linguistic variable and its application to approximate reasoning". *Information Science* (núm. 8, págs. 199-249 y 301-357).

Formalmente, una variable lingüística viene definida por una tupla de la forma $<X, L_X, U_X, S_X>$, donde los elementos que aparecen corresponden a:

X: el nombre de la variable lingüística. Por ejemplo, edad, altura, temperatura.

$L_X$: los valores lingüísticos que puede tomar la variable X. Por ejemplo, si tomamos como variable lingüística la temperatura podemos definir $L_{temperatura}$ como el conjunto {frío, tibio, caliente}. A cada elemento del conjunto $L_X$ se lo denomina *término lingüístico*, y, por lo tanto, $L_X$ es el conjunto de términos lingüísticos para la variable X. Así, frío es un término lingüístico de la variable lingüística temperatura y, siguiendo con el ejemplo, $L_{temperatura}$ = {frío, tibio, caliente}.

$U_X$: es el universo de discurso sobre el cual toma valores la variable lingüística. De hecho, corresponde al dominio en el que podemos interpretar los términos lingüísticos. Por ejemplo, cuando hablamos de temperatura, los términos frío o caliente se pueden interpretar como ciertos valores numéricos en el dominio de los números reales. Así, en este caso, y considerando la temperatura en grados centígrados, utilizaremos $U_{temperatura}$ como un subconjunto de números reales. En muchos sistemas, el dominio de las variables corresponde a un subconjunto de los números reales.

$S_X$: es la función semántica que da significado (interpreta) a cada uno de los términos lingüísticos. Así, a cada valor de $L_X$ se le asigna un conjunto difuso sobre $U_X$. Por ejemplo, al término caliente le asignaremos las temperaturas a partir de 40 grados.

**Definición completa para la variable temperatura**

La definición de la variable temperatura necesita los cuatro elementos descritos anteriormente. Así, tenemos que la variable lingüística temperatura se define como la tupla:

<temperatura, $L_{temperatura}$ = {frío, tibio, caliente}, $U_{temperatura}$ = [-50, 50], $S_{temperatura}$>

donde la función $S_{temperatura}$ és:

$S_{temperatura}$(frío) = $\mu_{frío}$, $S_{temperatura}$(tibio) = $\mu_{tibio}$ y $S_{temperatura}$(caliente) = $\mu_{caliente}$

con las funciones de pertenencia $\mu_{frío}$, $\mu_{tibio}$ y $\mu_{caliente}$ definidas como:

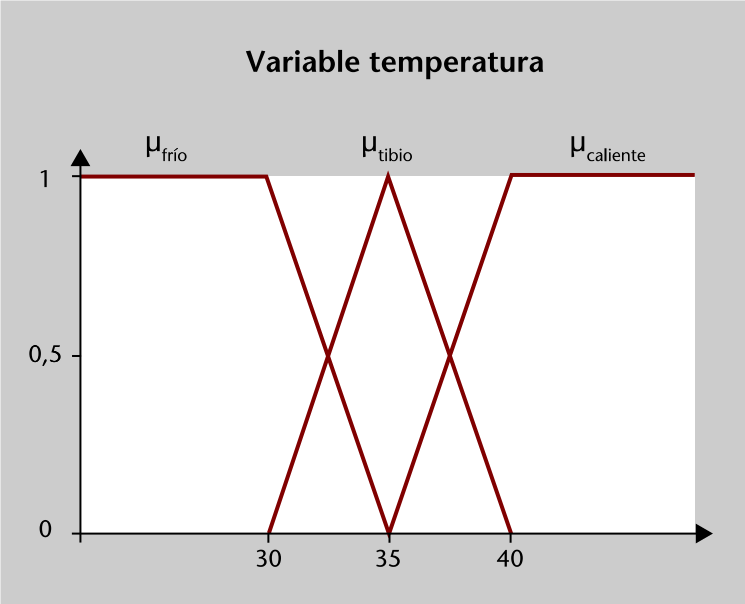
$$\mu_{\text{frío}}(x) = \begin{cases} 1, & \text{si } x < 30 \\ 1 - ((x-30)/(35-30)) & \text{si } 30 \le x \le 35 \\ 0, & \text{si } x > 35 \end{cases}$$

$$\mu_{\text{tibio}}(x) = \begin{cases} 0, & \text{si } x < 30 \\ (x-30)/(35-30) & \text{si } 30 \le x \le 35 \\ 1 - ((x-35)/(40-35)) & \text{si } 35 \le x \le 40 \\ 0, & \text{si } x > 40 \end{cases}$$

$$\mu_{\text{caliente}}(x) = \begin{cases} 0, & \text{si } x < 30 \\ (x-35)/(40-35) & \text{si } 35 \le x \le 40 \\ 0, & \text{si } x > 40 \end{cases}$$

En la figura 8 se muestra la representación gráfica de los tres términos lingüísticos que conforman la variable lingüística temperatura. Como es frecuente en la literatura de los conjuntos difusos, las funciones de pertenencia se presentan superpuestas. Cada término se puede definir de forma diferente uno del otro, como se ve en la figura, combinando funciones triangulares y trapezoidales.

Figura 8




Es conveniente que el dominio de aplicación se divida en regiones regulares y que podamos construir estas regiones fácilmente cuando disponemos de una partición del dominio de las variables de entrada (de su universo de discurso). Como para cada variable hay definido un conjunto de términos, y para cada término hay asociado un conjunto difuso, podemos exigir que estos conjuntos definan una partición del universo de discurso. Dado que estamos hablando de conjuntos difusos, y no de conjuntos nítidos, lo que exigimos es que la partición sea difusa.

El concepto de partición difusa es una generalización del concepto de partición nítida.

Un conjunto $\mu = \{\mu_1, ..., \mu_n\}$ dde conjuntos difusos define una partición difusa del universo de discurso *D*, cuando $\Sigma_{i=1,n}\mu_i(x) = 1$ para todo $x \in$ D.

Notad que si los conjuntos $\mu_i$ son nítidos, lo que estamos exigiendo corresponde, de hecho, a una partición nítida. En este caso, tenemos que cada *x* del dominio *D* sólo puede pertenecer a un único conjunto. En el caso difuso, tendremos que un elemento podrá pertenecer a más de un conjunto, pero en este caso las pertenencias serán parciales porque han de sumar 1. Como la suma de las pertenencias debe ser 1, si los conjuntos son nítidos tenemos una única $\mu_i$ con $\mu_i(x) = 1$ mientras que para todas las otras la pertenencia es cero.

**Definición completa para la variable temperatura: función de pertenencia**

De acuerdo con todo esto, en nuestro ejemplo exigimos que las funciones de pertenencia asociadas a los términos lingüísticos caliente, tibio y frío de la variable temperatura definan una partición difusa del universo de discurso de esta variable. Se puede comprobar que las funciones de pertenencia $\mu_{frío}$, $\mu_{tibio}$ y $\mu_{caliente}$ definen una partición difusa del universo de discurso $U_{temperatura} = [-50, 50]$ porque:

para todo $x$ del dominio $U_{temperatura} = [-50, 50]$ se cumple $\mu_{frío}(x) + \mu_{tibio}(x) + \mu_{caliente}(x) = 1$.

Podemos ver que en este ejemplo hay elementos del dominio que sólo pertenecen a un conjunto. Por ejemplo, 50 grados sólo pertenece al conjunto difuso caliente (a los otros pertenece con grado de pertenencia cero) mientras que 36 pertenece con grado $\mu_{tibio}(36) = 1 - ((36 - 35)/(40 - 35)) = 0,8$ al conjunto difuso tibio y con grado $\mu_{caliente}(36) = (36 - 35)/(40 - 35) = 0,2$ al conjunto difuso caliente.

### Selección de las variables y construcción de las reglas

Con las indicaciones que hemos suministrado sobre la definición de las variables lingüísticas y la construcción de las reglas, se puede pasar a construir el sistema difuso. Para construirlo, ante todo, debemos seleccionar las variables que utilizaremos, eligiendo las que son de entrada y las que son de salida. Después habremos de definir para cada una de las variables sus términos lingüísticos, el universo de discurso y los conjuntos difusos que corresponden a cada término. Una vez hecho esto, podremos pasar a construir las reglas. A continuación detallamos un poco más cada uno de estos pasos:

**1) La variable de salida**: En primer lugar, hemos de determinar cuál es la variable que nos permite controlar el aparato o la variable que queremos modelizar. Esta variable será la de salida del sistema difuso (supondremos que existe una única variable de salida).

**2) Las variables de entrada**: Una vez determinada la variable de salida, debemos seleccionar las variables de entrada del sistema. Esto es, elegir las variables que hemos de tener en cuenta a la hora de determinar el valor de la variable de salida. El conjunto de variables a tener en cuenta puede ser muy grande ya que puede haber muchas que influyen en el resultado. Sin embargo, sólo se han de seleccionar las que son más relevantes: aquellas variables que son realmente significativas para el proceso que se quiere controlar o para el que se quiere modelizar. De hecho, cuantas menos variables se seleccionen, más sencilla resulta la construcción del sistema.

**Ved también**

Podéis ver algunos comentarios sobre la complejidad de un sistema en relación con el número de variables en el apartado 2.5.1.

**3) Componentes de las variables**: Después de definir el conjunto de variables, hemos de definir para cada una de éstas sus términos lingüísticos, el universo de discurso y los conjuntos difusos que haremos corresponder a cada término. Los conjuntos difusos de las variables de entrada deben definir una partición del universo de discurso de la variable. La definición de las funciones de pertenencia constituye un proceso difícil ya que el comportamiento final del sistema depende de esto. Por ello, se han desarrollado métodos que permiten optimizar el sistema a partir de la modificación de las funciones de pertenencia.

**Ved también**

Algunos de los métodos se basan en los algoritmos genéticos vistos en el apartado 7 del módulo "Resolución de problemas y búsqueda".

**4) Definición de las reglas**: Si tenemos $n$ variables de entrada $X_1, ..., X_n$, construiremos para cada tupla de la forma $(t_{1,a}, t_{2,b}, t_{3,c}, ..., t_{n,z})$ donde $t_{i,j}$ es un término lingüístico de $X_i$ (por lo tanto $t_{i,j} \in L_{Xi}$) una regla de la forma:

**si** $X_1$ es $t_{1,a}$ **y** $X_2$ es $t_{2,b}$ **y** $X_3$ es $t_{3,c}$ **y** ... **y** $X_n$ es $t_{n,z}$ **entonces** $Y$ es $t_{Y,o}$

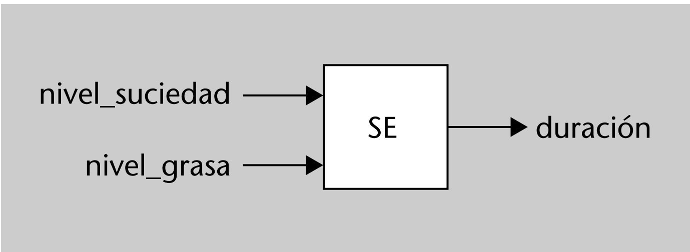
donde $t_{Y,o}$ es un término lingüístico de la variable de salida $Y$ (por lo tanto $t_{Y,o} \in L_Y$). El conjunto de tuplas debe cubrir todo el dominio de la aplicación.

Se debe decir que la construcción de un sistema difuso (tanto si es para un sistema de control, como para uno de modelización) resulta difícil ya que la selección de cuáles son las variables de entrada a menudo no resulta sencilla (que sean pocas y relevantes) y además tampoco resulta fácil construir las reglas (definir cuáles son las conclusiones de cada regla). De hecho, en el proceso de modelización se deben tener en cuenta todos aquellos elementos que se han comentado en el apartado 1.1.1. del módulo “Sistemas basados en el conocimiento”.

**Un ejemplo de sistema de inferencia difuso**

A continuación, completamos los pasos para definir el sistema de control de una lavadora que se ha ido explicando en este módulo.

Como se ha visto hasta ahora, el sistema describe dos variables (nivel de suciedad y nivel de grasa). El sistema de decisión (SE) toma estas dos *variables de entrada* para decidir la duración del programa. Por lo tanto, la duración, es una **variable de salida**.




Para cada una de estas variables, tenemos que definir los términos lingüísticos, su universo de discurso y las funciones de pertenencia que se asocian a los términos lingüísticos:

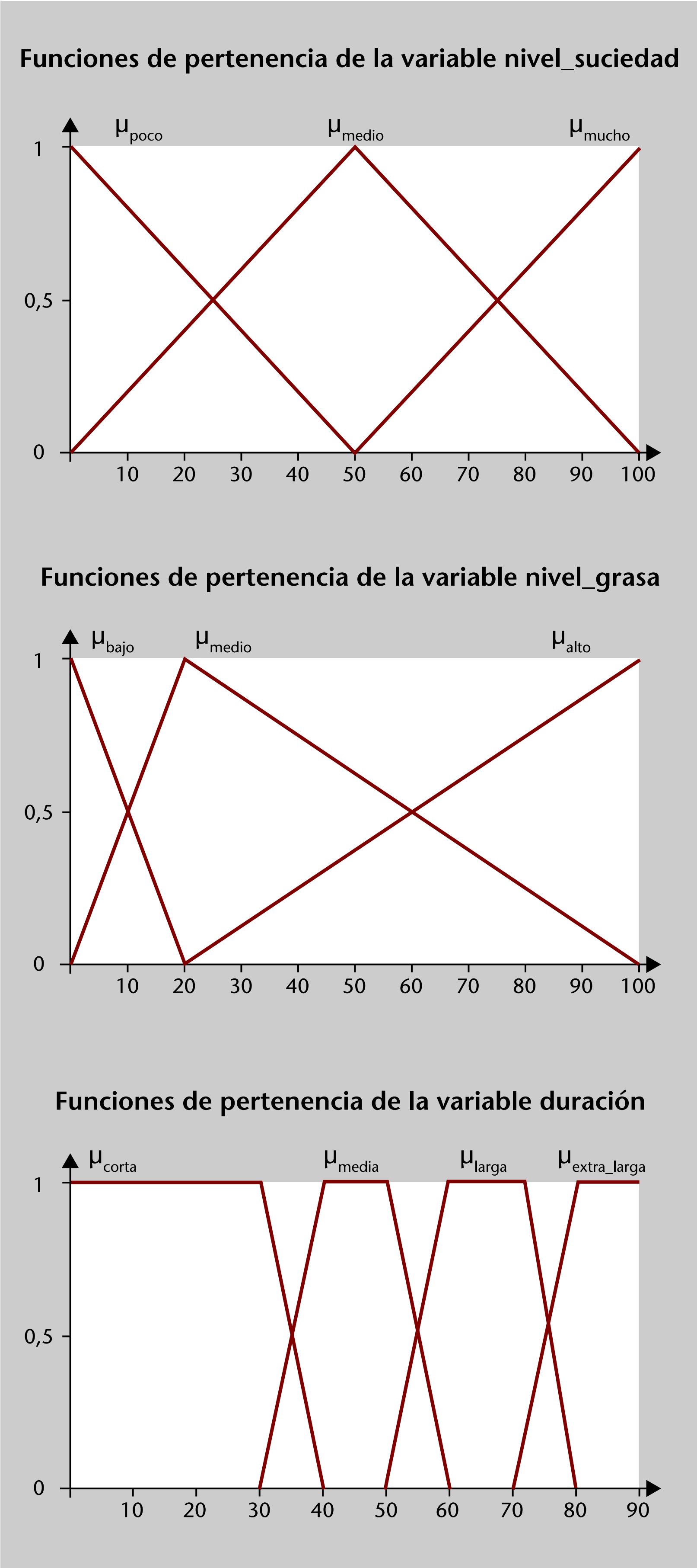
<nivel_suciedad, $L_{nivel\_suciedad}$ = {poco, medio, mucho}, $U_{nivel\_suciedad}$ = [0, 100], $S_{nivel\_suciedad}$>

<nivel_grasa, $L_{nivel\_grasa}$ = {baja, medio, alto}, $U_{nivel\_grasa}$ =[0, 100], $S_{nivel\_grasa}$>

<duración, $L_{duración}$ = {corta, media, larga, extra_larga}, $U_{duración}$ =[0, 90], $S_{duración}$>

Como se ha visto anteriormente, ahora nos queda describir los términos lingüísticos asociados a las tres variables. En este caso, también se ha optado por utilizar funciones de pertenencia lineales, trapezoidales y triagonales (o triangulares). La figura 9 muestra los términos definidos por estas tres variables.

Figura 9


Funciones de pertenencia de la variable nivel_suciedad
μpoco
μmedio
μmucho
1
0,5
0
10 20 30 40 50 60 70 80 90 100
Funciones de pertenencia de la variable nivel_grasa
μbajo
μmedio
μalto
1
0,5
0
10 20 30 40 50 60 70 80 90 100
Funciones de pertenencia de la variable duración
μcorta
μmedia
μlarga
μextra_larga
1
0,5
0
10 20 30 40 50 60 70 80 90

A continuación, describimos las funciones de pertenencia de cada término de cada variable:

Variable nivel_suciedad:

$$\mu_{\text{poco}}(x)=\begin{cases}-0{,}02x+1, & \text{si } 0\leq x<50\\ 0, & \text{si } 50\leq x\leq 100\end{cases}$$
$$\mu_{\text{medio}}(x)=\begin{cases}0{,}02x, & \text{si } 0\leq x<50\\ -0{,}02x+2, & \text{si } 50\leq x\leq 100\end{cases}$$
$$\mu_{\text{mucho}}(x)=\begin{cases}0, & \text{si } 0\leq x<50\\ 0{,}02x-1, & \text{si } 50\leq x\leq 100\end{cases}$$

Variable nivel_grasa:

$$\mu_{\text{bajo}}(x)=\begin{cases}-0{,}05x+1, & \text{si } 0\leq x<20\\ 0, & \text{si } 20\leq x\leq 100\end{cases}$$
$$\mu_{\text{medio}}(x)=\begin{cases}0{,}05x, & \text{si } 0\leq x<20\\ -0{,}0125x+1{,}25, & \text{si } 20\leq x\leq 100\end{cases}$$
$$\mu_{\text{alto}}(x)=\begin{cases}0, & \text{si } 0\leq x<20\\ 0{,}0125x-0{,}25, & \text{si } 20\leq x\leq 100\end{cases}$$

Variable duración:

$$\mu_{\text{corta}}(x)=\begin{cases}1, & \text{si } 0\leq x<30\\ -0{,}1x+4, & \text{si } 30\leq x<40\\ 0, & \text{si } 40\leq x\leq 90\end{cases}$$
$$\mu_{\text{media}}(x)=\begin{cases}0, & \text{si } 0\leq x<30\\ 0{,}1x-3, & \text{si } 30\leq x<40\\ 1, & \text{si } 40\leq x<50\\ -0{,}1x+6, & \text{si } 50\leq x<60\\ 0, & \text{si } 60\leq x\leq 90\end{cases}$$
$$\mu_{\text{larga}}(x)=\begin{cases}0, & \text{si } 0\leq x<50\\ 0{,}1x-5, & \text{si } 50\leq x<60\\ 1, & \text{si } 60\leq x<70\\ -0{,}1x+8, & \text{si } 70\leq x<80\\ 0, & \text{si } 80\leq x\leq 90\end{cases}$$
$$\mu_{\text{extra\_larga}}(x)=\begin{cases}0, & \text{si } 0\leq x<70\\ -0{,}1x-7, & \text{si } 70\leq x<80\\ 1, & \text{si } 80\leq x\leq 90\end{cases}$$

Una vez definidas las variables, nos queda conocer las reglas que gobernarán el SE. De acuerdo con lo que se ha apuntado en el cuarto punto del esquema anterior, habría que definir una regla para cada uno de los pares (*a*, *b*) donde *a* es un término lingüístico del conjunto $L_{\text{nivel\_suciedad}}$ y *b* es un término lingüístico del conjunto $L_{\text{nivel\_grasa}}$. En total, si utilizamos reglas conjuntivas, tenemos un total de nueve reglas posibles.

| **nivel_grasa\nivel_suciedad** | **poco** | **medio** | **mucho** |
|---|---|---|---|
| **bajo** | corta | media | larga |
| **medio** | corta | media | larga |
| **alto** | larga | larga | extra_larga |

Aprovechando que tenemos dos variables, podemos mostrar las reglas en forma de tabla, indicando en cada eje los términos de cada variable.

Así, una de las reglas de la tabla sería la que se describió anteriormente en este módulo:

**si***nivel_suciedad es poco***y***nivel_grasa es alto***entonces***duración es larga*

**Comentarios adicionales sobre las reglas y la construcción del sistema**

En la descripción que se ha hecho, los antecedentes de las reglas aparecen ligados a conjunciones. Aunque esto facilita la construcción del sistema, porque es más fácil recubrir todo el dominio de aplicación, no es un requisito indispensable.

De hecho, en general, las premisas son combinaciones de expresiones de la forma "X es Y", donde X es una variable e Y es un término lingüístico de la variable X, y estas expresiones se pueden combinar mediante conjunciones, disyunciones o se pueden negar con negaciones.

Por ejemplo, podríamos definir reglas del estilo:

**si** *nivel_suciedad es mucho* **y** *nivel_grasa* ***no*** *es alto* **entonces** *duración es larga*

O utilizando una disyunción:

**si** *nivel_suciedad es mucho* **o** *nivel_grasa es alto* **entonces** *duración es larga*

Del mismo modo, es posible tener expresiones en las conclusiones en las que aparezca más de una variable y también haya conjunciones, disyunciones y negaciones. Aquí no consideraremos ninguno de estos casos.

## 2.3. Aplicación de una regla y de un conjunto de reglas

En el apartado anterior hemos visto cómo definir una regla y un sistema difuso. En este apartado veremos cómo aplicar las reglas cuando disponemos de una cierta información correspondiente a las variables de entrada. Empezamos mostrando cómo calcular el grado en el que se satisface el antecedente de una regla, después veremos qué se concluye de una regla cuando ésta sólo se satisface en un cierto grado y, finalmente, veremos qué ocurre cuando en un sistema existe más de una regla que se cumple en un determinado grado.

**Ebraim H. Mamdani**

Básicamente, el método Mamdani se define como un sistema de inferencia donde la *t-norma* es el mínimo y la *t-conorma* es el máximo.

**Lectura recomendada**

**IEEE** (1977). "Application of Fuzzy Logic to Approximate Synthesis using Linguistic Synthesis". *Transaction on Computers* (vol. 12, págs. 1182-1191).

La manera de aplicar las reglas que describimos aquí corresponde al **método de Mamdani** y a un tipo particular de inferencia.

### Grado de satisfacción del antecedente

El grado de satisfacción del antecedente de una regla corresponde al grado en el que la regla se activará. Este grado se calcula viendo la pertenencia de los valores de las variables a los conjuntos difusos que aparecen en el antecedente y, después, combinando las pertenencias de acuerdo con las conjunciones, disyunciones o complementos que aparecen en el antecedente. Como se ha dicho, en nuestro caso sólo consideraremos conjunciones.

Consideramos una regla con $n$ variables de entrada $X_1$, ..., $X_n$ de la forma:

**si** $X_1$ es $t_{1,a}$ **y** $X_2$ es $t_{2,b}$ **y** $X_3$ és $t_{3,c}$ **y** ... **y** $X_n$ es $t_{n,z}$ **entonces** $Y$ es $t_{Y,o}$

donde $t_{i,j}$ es un término lingüístico de $X_i$ (suponemos que el conjunto difuso asociado al término $t_{i,j}$ es $\mu_{i,j}$) y donde $t_{Y,o}$ es un término lingüístico de la variable de salida $Y$.

Entonces, cuando en un instante concreto conocemos el valor de cada variable $X_i$ (denotaremos estos valores por $x_i$), podemos evaluar la certeza de las expresiones "$X_i$ es $t_{i,j}$" para cada una de las variables. Como la función de pertenencia asociada a $t_{i,j}$ es $\mu_{i,j}$, tenemos que la evaluación de la expresión "$X_i$ es $t_{i,j}$" no es más que ver en qué grado el valor $x_i$ satisface el conjunto $\mu_{i,j}$. Esto es, $\mu_{i,j}(x_i)$. Por lo tanto, para la regla anterior podemos calcular:

$$\mu_{1,a}(x_1),\ \mu_{2,b}(x_2),\ \mu_{3,c}(x_3),\ \ldots,\ \mu_{n,z}(x_n)$$

Ahora, como en el antecedente de esta regla las expresiones "$X_i$ es $t_{i,j}$" se encuentran unidas con conjunciones, y la conjunción la modelizamos en el caso difuso con $t$-normas, tenemos que la certeza de todo el antecedente es:

$$T(\mu_{1,a}(x_1),\ \mu_{2,b}(x_2),\ \mu_{3,c}(x_3),\ \ldots,\ \mu_{n,z}(x_n))$$

Como de acuerdo con lo que hemos visto en el subapartado 2.1. no existe una única $t$-norma, a la hora de construir un sistema difuso se debe decidir qué $t$-norma utilizamos para modelizar la conjunción. Evidentemente, $t$-normas diferentes darán resultados diferentes

**Ved también**

En el subapartado 2.5 se comentará un poco más sobre la cuestión de la elección de los parámetros.

**Evaluación del antecedente de una de las reglas del sistema de control usando la definición anterior de las variables**

Evaluaremos el grado de satisfacción de una regla teniendo en cuenta unos valores dados de las variables de entrada. En este caso, tomamos *nivel_suciedad* = 70 y *nivel_grasa* = 50.

Consideramos la regla siguiente:

**si** *nivel_suciedad es medio* **y** *nivel_grasa es alto* **entonces** *duración es larga*

Consideramos un sistema Mamdani con $t$-norma del mínimo.

En este punto, debemos conocer el grado de pertenencia de los valores de entrada en cada uno de los antecedentes de nuestra regla.

En este caso, tenemos que saber qué nivel de pertenencia tiene el valor nivel_suciedad de 70 en el término medio: $\mu_{\text{medio}}(70) = (-0{,}02 * 70 + 2) = 0{,}6$.

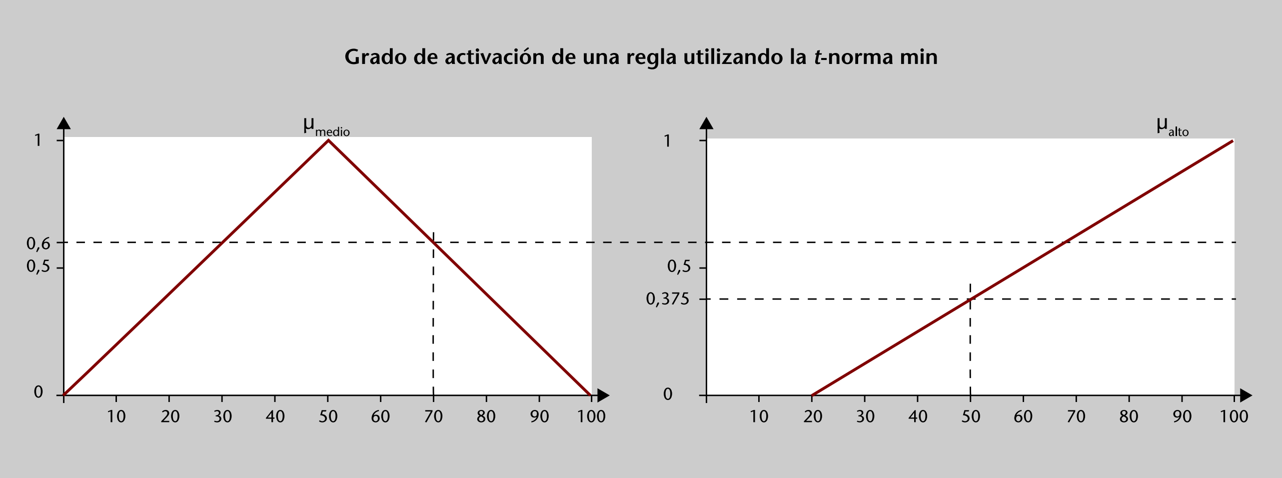
De forma análoga, hacemos lo mismo para conocer el grado de pertenencia del valor de entrada de 50 para la variable *nivel_grasa* y el término alto: $\mu_{\text{alto}}(50) = (0{,}0125 * 50 - 0{,}25) = 0{,}375$.

Dado que tomamos la $t$-norma del mínimo, el grado de satisfacción de los antecedentes en esta regla es min(0,6, 0,375) = 0,375.

Esto quiere decir que la regla se activa con un grado de certeza de 0,375 y, por lo tanto, tendremos que tenerla en cuenta a la hora de tomar una decisión final.

En la figura 10, se muestra una representación gráfica del cálculo de la satisfacción de la regla.

Figura 10




### Aplicación de una regla

Una vez tenemos la certeza de que se satisface el antecedente, pasamos a obtener la conclusión de la regla. Hemos de tener en cuenta que cuando el antecedente se satisface plenamente, la conclusión será la misma que tenemos en la regla, pero cuando el antecedente no se satisface plenamente, lo que tendremos como conclusión será una modificación de lo que aparece en la regla. De hecho, cuando aplicamos el sistema de Mamdani, el funcionamiento es como sigue.

Si suponemos que el grado en el que se satisface el antecedente es α, la conclusión de la regla será el conjunto difuso que aparece truncado con el valor de α. Esto es, si tenemos la regla:

**Si** $X_1$ es $t_{1,a}$ **y** $X_2$ es $t_{2,b}$ **y** $X_3$ es $t_{3,c}$ **y** ... **y** $X_n$ es $t_{n,z}$ **entonces** $Y$ es $t_{Y,o}$

donde $t_{i,j}$ es un término lingüístico de la variable $X_i$ y $\mu_{i,j}$ es el conjunto difuso asociado al término $t_{i,j}$ y donde $t_{Y,o}$ es un término lingüístico de la variable de salida $Y$ y $\mu_{Y,o}$ s el conjunto asociado a $t_{Y,o}$ entonces, la conclusión de la regla es el conjunto difuso:

$$\mu_{Y,o}(x) = \min(\alpha, \mu_{Y,o}Y(x)),$$

cuando:

$$\alpha = T(\mu_{1,A}(x_1), \mu_{2,A}(x_2), \mu_{3,C}(x_3), \ldots, \mu_{n,Z}(x_n))$$

Se puede comprobar que si el antecedente de una regla se cumple completamente ($\alpha = 1$), entonces la conclusión es $\mu_{Y,0}$. Dicho de otra manera, si el antecedente se cumple plenamente, la conclusión es la que aparece en la regla. Por otra parte, cuando el antecedente de la regla no se cumple ($\alpha = 0$), entonces esta construcción nos dará como conclusión de la regla el conjunto vacío.

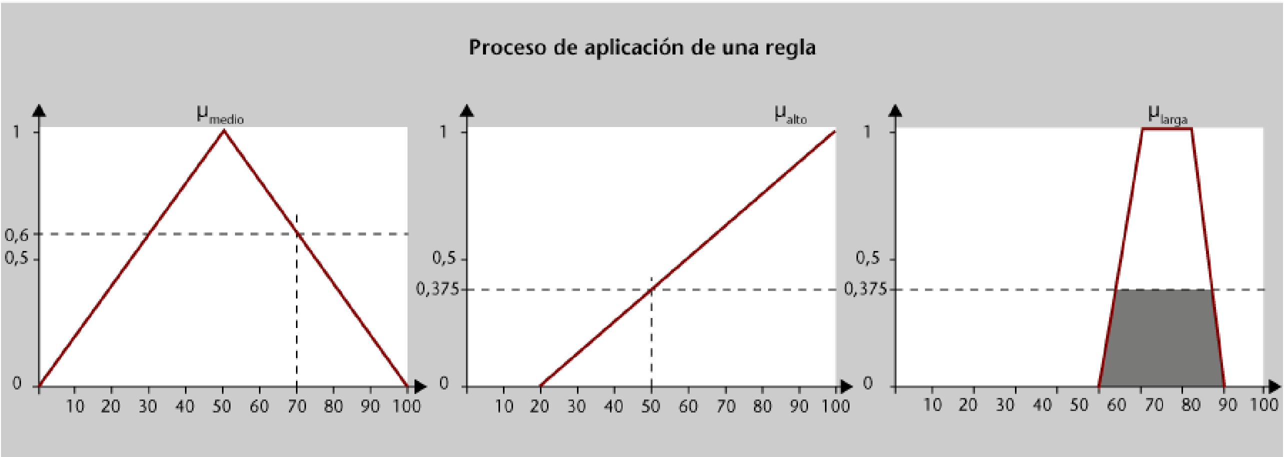
**Continuamos con el ejemplo de la lavadora**

Después de describir cómo se calcula el consecuente aplicando la *t*-norma del mínimo, observamos gráficamente cómo se aplica finalmente la regla que estamos viendo.

**si** *nivel_suciedad es medio* **y** *nivel_grasa es alto* **entonces** *duración es larga*

Cómo se ha visto, el consecuente de la regla queda activado con un nivel de 0,375, resultado de aplicar el mínimo a los antecedentes. Como se muestra en la figura 11, tenemos que la aplicación de la regla nos da que se activa el término *larga* de la variable de salida *duración* con un nivel de 0,375.

Figura 11




### Aplicación de un conjunto de reglas

Cuando en lugar de disponer de una única regla, tenemos un conjunto, lo que haremos es **aplicar el mismo proceso para cada una de las reglas**. Esto nos retornará un conjunto difuso para cada una de las reglas.

La conclusión del conjunto de reglas se define como la **unión de las conclusiones** (en la definición de la unión se usará una *t*-conorma).

Como se ha visto hasta ahora, el resultado de la unión de las conclusiones es un nuevo conjunto difuso. En la práctica, tendremos que calcular la función de pertenencia de dicho conjunto.

Volvemos a subrayar que en este proceso cuando el antecedente de una regla no se cumple, la conclusión será el conjunto vacío. Esto hará que al efectuar la unión de las conclusiones estas reglas no presenten ninguna participación en la conclusión final.

**Proseguimos con el ejemplo aplicando este procedimiento al conjunto de todas las reglas del sistema de control**

Para los valores de entrada dados, se activan unas normas y otras no. Podemos recuperar la tabla con las reglas e indicar el nivel de activación de los antecedentes. Con esto, podemos ver de forma directa qué reglas se activan y cuál es el nivel final.

Recordamos que consideramos los valores de entrada siguientes: *nivel_suciedad* = 70 y *nivel_grasa* = 50. Con estos valores se activan los términos *medio* (grado 0,6) y *mucho* (grado 0,4) de la variable *nivel_suciedad*, y los términos *medio* (grado 0,625) y *alto* (0,375) de la variable *nivel_grasa*.

En la tabla siguiente, en las filas con trama gris y entre paréntesis, mostramos los grados de activación de estos términos lingüísticos. Y en el interior de la tabla, se muestran los grados de satisfacción de los antecedentes (aplicando la *t*-norma del mínimo).

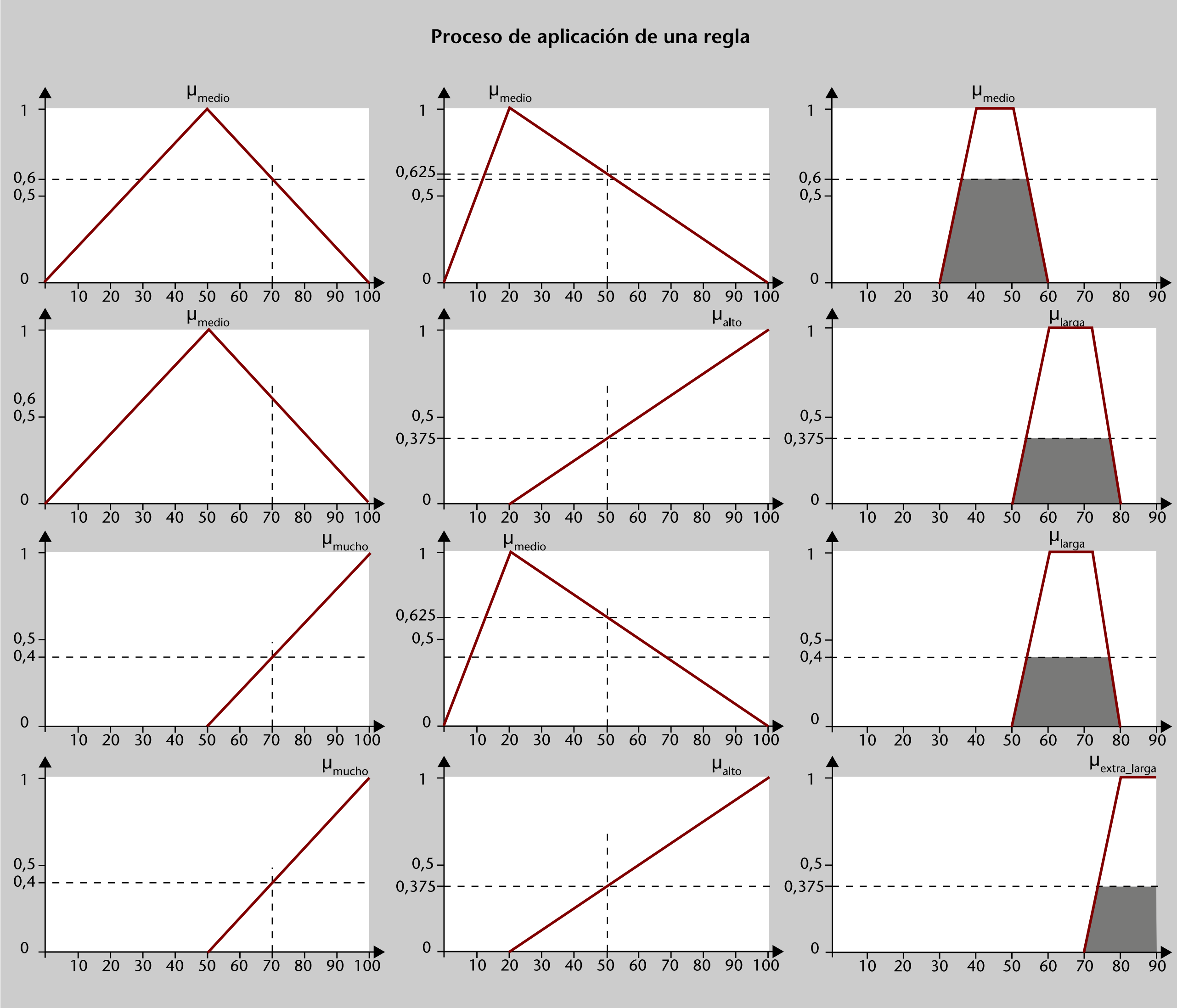
| nivel_grasa\nivel_suciedad | poco | medio (0,6) | mucho (0,4) |
|---|---|---|---|
| **bajo** | corta | media | larga |
| **medio (0,625)** | corta | media (0,6) | larga (0,4) |
| **alto (0,375)** | larga | larga (0,375) | extra_larga (0,375) |

Es decir, con los valores de entrada dados, se activan cuatro reglas (de las nueve que contiene el bloque).

- **si** *nivel_suciedad es medio* **y** *nivel_grasa es medio* **entonces** *duración es media*
- **si** *nivel_suciedad es medio* **y** *nivel_grasa es alto* **entonces** *duración es larga*
- **si** *nivel_suciedad es mucho* **y** *nivel_grasa es medio* **entonces** *duración es larga*
- **si** *nivel_suciedad es mucho* **y** *nivel_grasa es alto* **entonces** *duración es extra_larga*

En la figura 12, se muestra gráficamente la obtención de los consecuentes en estas reglas, donde las gráficas de la primera columna hacen referencia al nivel de suciedad, las gráficas de la segunda columna al nivel de grasa y las de la tercera columna a la variable duración (salida del sistema). Cada fila corresponde a cada una de las cuatro reglas que se activan y que se han especificado anteriormente.

Figura 12



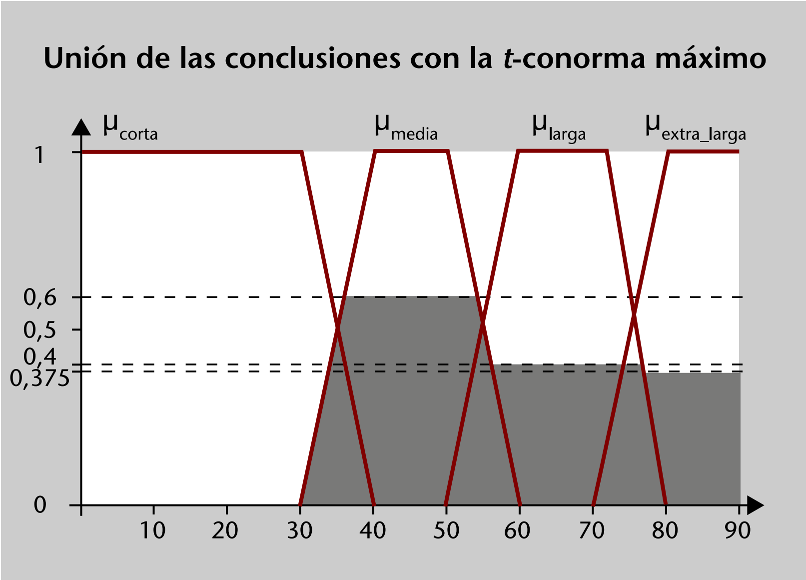

En la figura 13, se observa la unión de las conclusiones aplicando la t-conorma del máximo. Vemos que, en cada término, se aplica el máximo calculado anteriormente.

En el caso de los términos *media* y *extra_larga* no hay ningún problema, porque sólo hay una regla que los active y el resultado final es directo; pero en el caso del término *larga* hay dos reglas que lo activan. En este caso, es donde aplicamos la *t*-conorma del máximo para calcular el grado de activación: max(0,375, 0,4) = 0,4.

Antes de pasar a calcular el valor nítido de la salida, nos queda calcular la función de pertenencia resultante de la unión de los conjuntos difusos ($\mu_{\text{duración}}(x)$).

$$\mu_{\text{duración}}(x) = \begin{cases} 0, & \text{si } 0 \le x < 30 \\ 0,1x - 3, & \text{si } 30 \le x < 36 \\ 0,6, & \text{si } 36 \le x < 54 \\ -0,1x + 6, & \text{si } 54 \le x < 56 \\ 0,4, & \text{si } 56 \le x < 76 \\ -0,1x + 8, & \text{si } 76 \le x < 76,25 \\ 0,375, & \text{si } 76,25 \le x \le 90 \end{cases}$$

Figura 13




**Nitidificación**

En el paso anterior, hemos obtenido como conclusión del sistema un conjunto difuso. Desafortunadamente, una solución de este tipo a menudo no es adecuada porque la variable de salida requiere ser un valor real. Por ejemplo, la intensidad que hemos de suministrar al aparato que controlamos, o la velocidad a la que se ha de desplazar un robot. En estos casos, los sistemas difusos añaden un paso más al proceso de deducción. Este paso se llama de *nitidificación*[5] y consiste en obtener un valor nítido a partir de un conjunto difuso.

(5) En inglés, *defuzzification*.

De hecho, existen varias maneras de realizar el proceso de nitidificación. Una de éstas la constituye el llamado *centro de masas* (o *centro de área*) y se define como una *media* de los valores del dominio, cada uno ponderado según su pertenencia al conjunto que estamos nitidificando.

Así, si tenemos que la función de pertenencia conclusión de un sistema difuso es μ y si suponemos que esta función de pertenencia está definida en un dominio discreto $D$, entonces la nitidificación de μ es:

$$CoM = \frac{\sum_{x \in D} (\mu(x) * x)}{\sum_{x \in D} \mu(x)}$$

Dado que nuestra función de pertenencia es continua en el dominio $D$, también podemos hacer el cálculo analíticamente, sustituyendo el sumatorio por una integral definida de la forma siguiente:

$$CoM = \frac{\int_{x \in D} (\mu(x) * x)dx}{\int_{x \in D} (\mu(x))dx}$$

**Cálculo del valor nítido**

Siguiendo con el ejemplo de la lavadora, en el paso anterior se ha calculado la función de pertenencia de la variable de salida *duración*. Así, pues, ya podemos utilizar esta función para calcular el valor nítido final.

**Web recomendada**

Para obtener el valor nítido, podéis utilizar herramientas en línea como CalcMe (http://www.wiris.com/calc/).

Tenemos la función siguiente para cualquier *x* en el dominio [0, 90]:

$$\mu_{\text{duración}}(x) = \begin{cases} 0, & \text{si } 0 \le x < 30 \\ 0,1x - 3, & \text{si } 30 \le x < 36 \\ 0,6, & \text{si } 36 \le x < 54 \\ -0,1x + 6, & \text{si } 54 \le x < 56 \\ 0,4, & \text{si } 56 \le x < 76 \\ -0,1x + 8, & \text{si } 76 \le x < 76,25 \\ 0,375, & \text{si } 76,25 \le x \le 90 \end{cases}$$

Entonces, si calculamos el valor nítido de forma iterativa, el resultado final que obtenemos es:

$$CoM = \frac{\sum_{x=0}^{90}(\mu_{\text{duración}}(x) * x)}{\sum_{x=0}^{90}(\mu_{\text{duración}}(x))} = \frac{78.323,31}{1.342,8745} = 58,325$$

El cálculo de forma iterativa se puede hacer con un pequeño programa en el que se define un bucle y una función que retorne el valor de la función de pertenencia. En nuestro caso, esta función tiene que definir los diferentes tramos y retornará el valor que corresponda.

```
v_numerador = v_denominador = 0;
v_intervalo := 0,02; -- resolución de cálculo
i = 0;
mientras (i<90) hacer
    v_numerador += mu(i)*i;
    v_denominador += mu(i);
    i += v_intervalo;
finmientras;
vuelve (v_numerador/v_denominador);
```

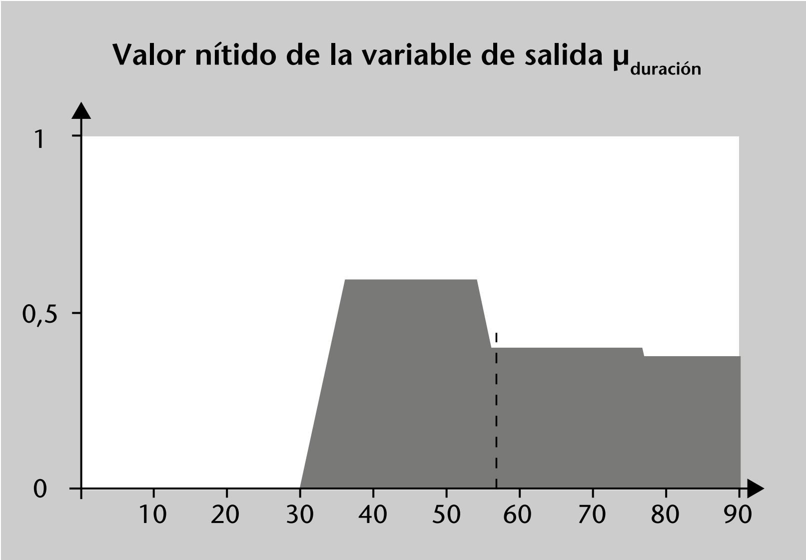
De forma análoga, podemos hacer el cálculo del *CoM* utilizando integrales por partes. Si hacemos los cálculos –sin ningún error de aproximación añadido–, nos da el resultado siguiente:

$$\text{CoM} = \frac{\int_0^{90}(\mu_{\text{duración}}(x) * x)\text{dx}}{\int_0^{90}(\mu_{\text{duración}}(x))\text{dx}} =$$

$$= \frac{\int_{30}^{36}((0,1x-3)x)dx + \int_{36}^{54}(0,6x)dx + \int_{56}^{54}((-0,1x+6)x)dx + \int_{56}^{76}(0,4x)dx + \int_{76}^{76,25}((-0,1x+8)x)dx + \int_{76}^{90}(0,375x)dx}{\int_{30}^{36}(0,1x-3)dx + \int_{36}^{54}(0,6)dx + \int_{56}^{54}(-0,1x+6)dx + \int_{56}^{76}(0,4)dx + \int_{76}^{76,25}(-0,1x+8)dx + \int_{76}^{90}(0,375)dx} =$$

$$= \frac{(61,2+486+54,933+528+7,3745+428,61)}{(1,8+10,8+1+8+0,096875+5,1563)} = 58,322$$

Figura 14




## 2.4. Sistemas con reglas conjuntivas y con reglas disyuntivas. Las implicaciones

El mecanismo de inferencia que se ha descrito no es el único que existe para los sistemas difusos. De hecho, podemos destacar cuatro tipos diferentes (dos de los cuales son equivalentes) de inferencia.

**Ved también**

Podéis ver el mecanismo de interferencia descrito en el subapartado 2.3 de este módulo.

En general, la inferencia se puede ver a partir de la construcción de una relación entre las variables de entrada y las de salida. Así, una regla de la forma:

**si** $X$ está en el intervalo [1, 2] **entonces** $Y$ es 2

indica que los elementos $x$ del intervalo [1, 2] se encuentran relacionados con el valor 2. Entonces, cuando tenemos un valor concreto de $x$, podemos buscar cuáles son los elementos de $y$ que se encuentran relacionados. Este conjunto constituye la conclusión de la regla para un estado dado.

**Elementos relacionados o no relacionados**

En el caso de la regla anterior los elementos del par ($x$ = 1,5, $y$ = 2) se encuentran relacionados, pero, en cambio, los elementos del par ($x$ = 1,5, $y$ = 3) no lo están.

En el caso nítido tenemos que la relación que se construye es nítida. Así, podemos ver si los elementos de un par (entradas, salida) se encuentran relacionados o no. Como es conocido, una relación se puede modelizar mediante una función característica R. En general, y para el caso del ejemplo de dos variables, tenemos que $R(x, y) \in \{0, 1\}$. En la figura 15, se da una representación gráfica de la función característica correspondiente a la regla anterior.

**Modelización de la relación**

Suponiendo que la función característica asociada a la regla se denota R, tenemos que $R(x = 1{,}5, y = 2) = 1$ y que $R(x = 1{,}5, y = 3) = 0$.

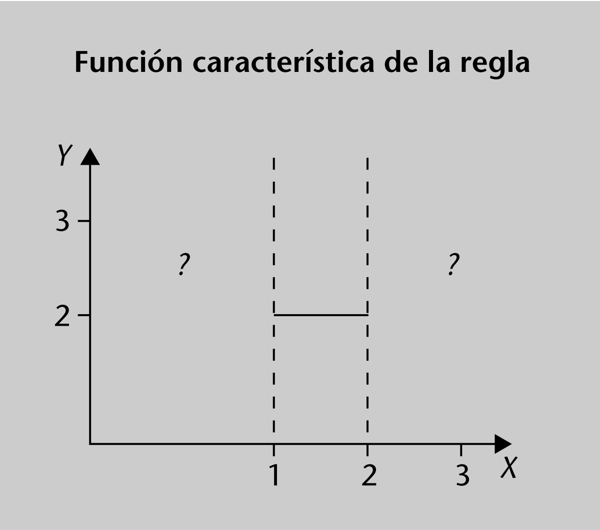
**si** $X$ está en el intervalo [1, 2] **entonces** $Y$ es 2

Los puntos ($x$, $y$) marcados corresponden a valores de $x$ e $y$ que se encuentran relacionados. En el caso de tener reglas difusas, tenemos que la relación que se construye es difusa y, por lo tanto, en lugar de poseer una función característica, lo que tenemos es una función de pertenencia.

**Relación difusa**

En este caso será posible, por ejemplo, tener $\mu(x = 1{,}5,\ y = 2{,}1) = 0{,}9$.

Figura 15




Un elemento importante de cómo efectuar la inferencia es cómo interpretar una regla o, lo que es lo mismo, cómo construir una relación difusa –su función de pertenencia– a partir de una regla. De hecho, esto se puede efectuar de dos maneras diferentes porque existen dos interpretaciones de las reglas:

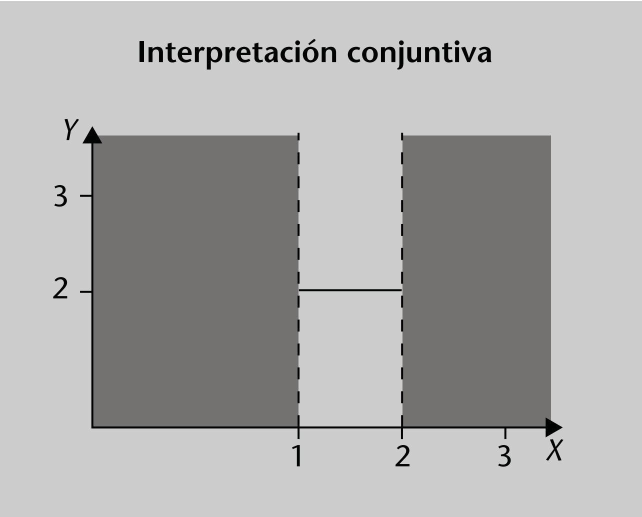
**1) Reglas conjuntivas**: una regla se interpreta como una implicación. El comportamiento de todas las reglas de un sistema se define a partir de la *conjunción* de las implicaciones (conjunción de las reglas). De hecho, si cada regla se ve como una relación, la relación que define el sistema formado por todas las reglas corresponde a la intersección de todas las relaciones

**2) Reglas disyuntivas**: una regla se interpreta como un punto del espacio entradas $x$ salida. El comportamiento de todas las reglas de un sistema es la *unión* de los "puntos". De hecho, si cada regla se ve como una relación, lo que tenemos es que el comportamiento del sistema viene definido por la unión de todas las relaciones.

De manera gráfica, estas dos interpretaciones corresponden a dos maneras de llenar la relación en las zonas que se han marcado con un interrogante en la figura 15.

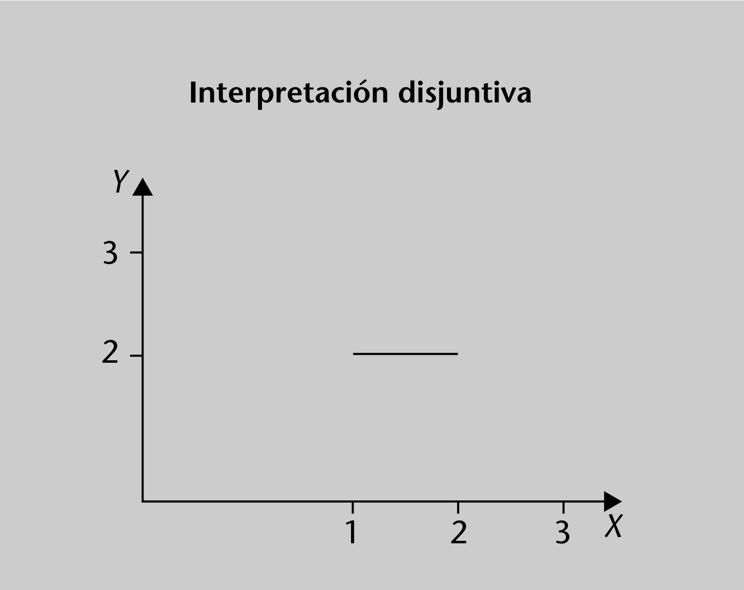
En el caso de las reglas conjuntivas, tenemos que la zona marcada con el interrogante se llena diciendo que los valores sí que pueden estar relacionados (la función característica posee un valor de 1). Esto es así porque, como se ha dicho, una regla se ve como una implicación. Por lo tanto, para calcular la certeza para los pares podemos tener en cuenta cómo se calcula la certeza en una implicación. Recordemos de la asignatura de lógica que la certeza de una implicación cuando el antecedente no se cumple resulta cierta. Esto es, F → T es T y que F → F también es T. Así, para todos los pares con una *x* que no es del intervalo [1, 2] la certeza de la implicación es 1. En la figura 16 se ofrece la representación de la relación que tendríamos de acuerdo con la interpretación conjuntiva.

Figura 16




En el caso de reglas disyuntivas, la regla nos da los valores que sabe relacionados y el resto se consideran como no relacionados (se basan en que, si hay otros puntos relacionados, la misma regla u otra ya nos lo dirá). En este caso, se ve la regla como una conjunción. Un par (*x*, *y*) se encuentra relacionado si la regla nos lo dice. Esto corresponde en el caso anterior a decir: x está en el intervalo [1, 2] y, además, *y* = 2. Esto se consigue modelizando la regla mediante una conjunción (conjunción de la certeza del antecedente y de la certeza del consecuente). Por lo tanto, cuando *x* se encuentra fuera del intervalo [1, 2] la certeza de la relación debe ser cero. La figura 17 da la representación de la relación en este caso.

Figura 17

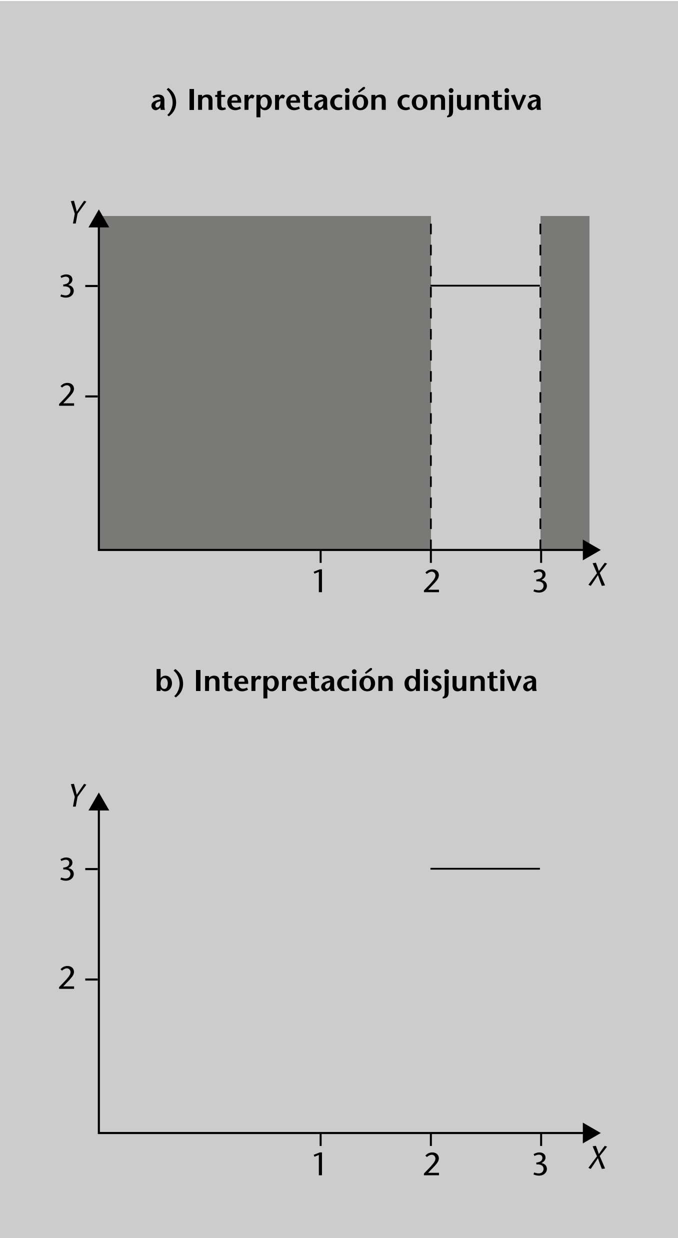



Como ya se ha dicho, los nombres de interpretación conjuntiva y disyuntiva vienen de cuando en lugar de tener una única regla tenemos más. En el primer caso, haremos conjunción de las relaciones (intersección) y en el segundo caso haremos una disyunción de las relaciones (unión). Para comprender cómo se interpreta un conjunto de reglas en cada uno de los casos tomamos la regla anterior y la nueva regla:

**si** *X* está en el intervalo [2, 3] **entonces** *Y* es 3

Las relaciones correspondientes a esta regla de acuerdo con las interpretaciones conjuntiva y disyuntiva se dan en la figura 18.

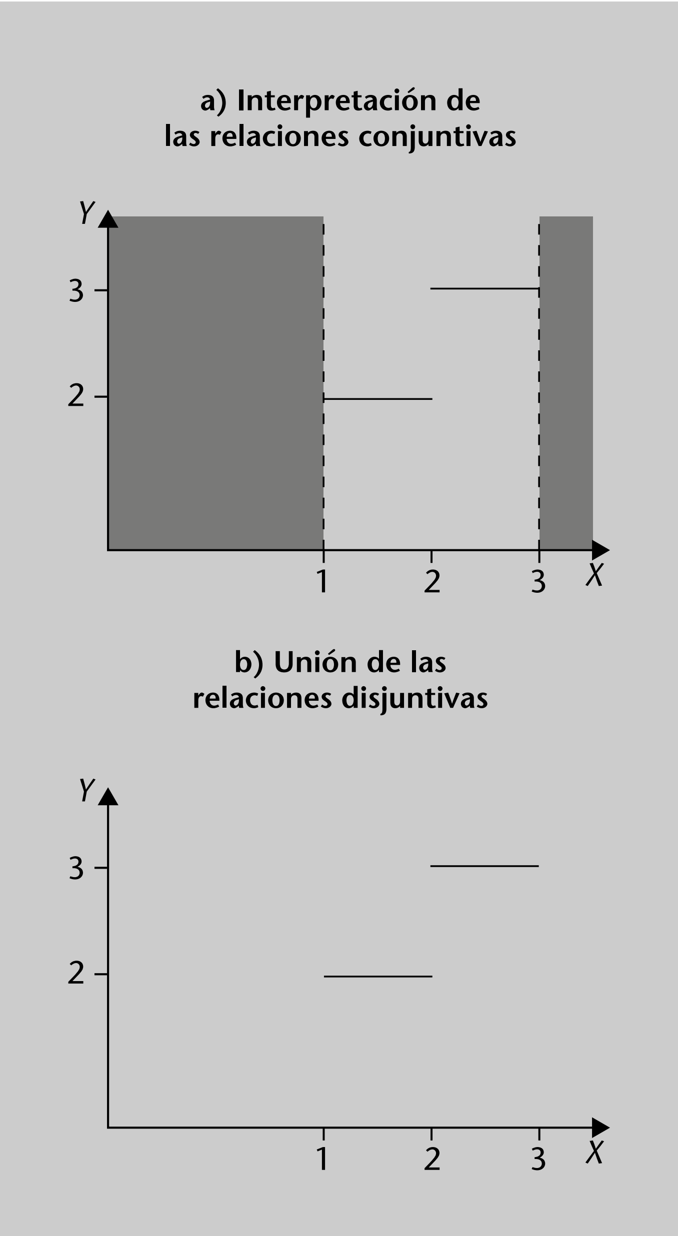
Figura 18




Parece razonable pensar que si en un sistema tenemos las dos reglas, lo que querremos es que cuando la $x$ esté en el intervalo [1, 2] la conclusión sea $y$ = 2 y que cuando esté en el intervalo [2, 3] la conclusión sea $y$ = 3. Si efectuamos la intersección de las relaciones conjuntivas –intersección de las figuras 18 y 20(a) que representan las relaciones conjuntivas– obtendremos la relación conjuntiva de la figura 19(a). Si, en cambio, efectuamos la unión de las disyuntivas –unión de las figuras 19 y 20(b) que representan las relaciones disyuntivas– obtendremos la relación disyuntiva de la figura 19(b). En ambos casos, tenemos que para x en el intervalo [1, 3] la figura es la misma. Podéis ver que la intersección de las relaciones disyuntivas da la relación vacía y que la unión de las relaciones conjuntivas nos dará la relación que siempre es 1.

Aunque en la descripción y en el ejemplo nos hemos centrado en las reglas binarias, lo mismo sirve para el caso difuso. Cuando los antecedentes y los consecuentes se encuentran definidos mediante conjuntos difusos, lo que tendremos será que la implicación (que aparece en la interpretación de las reglas conjuntivas) ha de ser una función de implicación difusa. Estas funciones se construyen a partir de propiedades de manera parecida a como hemos hecho con las negaciones, *t*-normas y *t*-conormas. En el caso de la interpretación disyuntiva no es necesaria ninguna función nueva porque al modelizar la regla con una conjunción se utiliza una *t*-norma.

Figura 19




Una vez hemos interpretado las reglas, debemos ver cómo combinar la información correspondiente a un estado concreto con las relaciones que hemos construido. Esto nos dará la conclusión de todo el sistema. Cuando sólo se cuenta con una relación, la combinación de las dos informaciones se realiza mediante la composición de la relación con la información que poseemos. En el caso nítido, esto resulta sencillo. Por ejemplo, si tenemos que el estado del sistema corresponde a $x$ =1,5, sólo hemos de mirar qué elementos se encuentran relacionados con esta $x$. En nuestro caso, sólo hay uno ($y$ = 2) pero en general podría haber más. La conclusión es el conjunto de todos los puntos del dominio de la variable $Y$ relacionados. En el caso difuso, la conclusión será un conjunto difuso sobre el dominio de la variable $Y$.

Sin entrar en detalle en este proceso, podemos escribir que si hay una regla que combina dos variables $X$ e $Y$ que se interpreta, como en el ejemplo, mediante una relación difusa $\mu_R(x, y)$, la conclusión de esta regla cuando el estado corresponde al conjunto difuso $\mu_A$ es:

$$\mu_B(Y) = \sup_{x \in X} \min(\mu_R(x, y), \mu_A(x))$$

Notad que la información del estado actual debe ser un conjunto sobre el dominio de las $X$ y que la conclusión es un conjunto difuso sobre el dominio de las $Y$. En el caso concreto en el que la información es un valor concreto $x_o$ del dominio de las $X$ en lugar de ser un conjunto difuso $\mu_A$, lo que tenemos es: $\mu_B(y) = \min(\mu_R(x, y), \mu_A(x)) = \mu_R(x_o, y)$ porque $\mu_A(x_o) = 1$ y $\mu_A(x) = 0$ para cualquier otra x diferente ($x \neq x_o$).

La descripción de la inferencia de Mamdani que hemos proporcionado en el subapartado 2.3 corresponde a un sistema con reglas disyuntivas donde la *t*-norma para combinar el antecedente y el consecuente de las reglas es el mínimo. Observad que en este caso podemos decir que:

$$\mu_B(y) = \mu_R(x, y) = T(\mu_{\text{antecedente}}(x_o), \mu_{\text{consecuente}}(y)),$$

que cuando $T$ = mínimo es lo mismo que cortar el conjunto $\mu_{\text{consecuente}}(y)$ según el valor de $\mu_{\text{antecedente}}(x_o)$.

La composición que se ha dado aquí presupone que la información sobre el estado actual viene dada por un conjunto difuso y que la base de conocimiento del sistema es una relación. Cuando en lugar de tener una única regla disponemos de un conjunto de reglas, podemos operar primero las reglas con el fin de construir una relación que resuma la información de todas las reglas (realizando la unión o la intersección de las relaciones según corresponda) y aplicar después el proceso descrito. Sin embargo, existe una alternativa. Podemos componer de manera independiente cada regla y después combinar las conclusiones (de manera parecida a como hemos descrito en el sistema de Mamdani). Si tomamos la opción de combinar las conclusiones, hemos de tener en cuenta que cuando las reglas son disyuntivas deberemos efectuar la unión de las conclusiones (como hemos realizado en el caso del sistema de Mamdani), pero que cuando las reglas son conjuntivas lo que hay que hacer es la intersección de los resultados.

Cuando consideramos el hecho de tener reglas conjuntivas y disyuntivas, y la posibilidad de realizar primero una relación que resuma todas las reglas o bien combinar las conclusiones que da cada regla individualmente, tenemos cuatro maneras diferentes de operar:

**1)** Reglas disyuntivas, componiendo la información del estado con cada una de las reglas y después haciendo la unión de las conclusiones.

**2)** Reglas disyuntivas, primero efectuando la unión de las reglas y después componiendo la información correspondiente al estado con la relación que resulta de ellas.

**3)** Reglas conjuntivas, componiendo la información del estado con cada una de las reglas y después realizando la intersección de las conclusiones.

**4)** Reglas conjuntivas, primero haciendo la intersección de las reglas y después componiendo la información correspondiente al estado con la relación que resulta de ellas.

El sistema de Mamdani corresponde al primer tipo de inferencia. Debemos subrayar que el primero y el segundo tipo de inferencia dan exactamente el mismo resultado, los otros dan resultados diferentes. De hecho, se puede demostrar que el conjunto resultante del tercer tipo de inferencia es siempre un subconjunto del resultado del segundo tipo y que el resultado del cuarto tipo es siempre un subconjunto del resultado del tercero.

## 2.5. Aspectos prácticos de estos sistemas

En la descripción que hemos comentado de un sistema difuso, se ha visto que existen muchos parámetros que se deben fijar. Una vez se han definido las reglas, hemos de dar una función de pertenencia para cada término, después deberemos definir las *t*-normas, *t*-conormas y el procedimiento de nitidificación que utilizaremos (además de la negación en algunos casos). Además, hemos visto también que no existe una única manera de realizar la inferencia, sino que hay más de una. A todo esto se debe añadir el problema de la discretización de las variables. Generalmente, no implementaremos un sistema difuso considerando funciones de pertenencia continuas y esto provoca que se deban considerar discretizaciones del dominio (cuanto más fina sea la discretización, mayor será el coste computacional del sistema).

Para completar el sistema es necesario definir todos estos elementos. Su elección no resulta sencilla y depende de muchos factores. Para facilitar este proceso, existen procedimientos basados en algoritmos genéticos y en redes neuronales para encontrar valores óptimos para un sistema concreto. En general, sin embargo, habrá que encontrar manualmente valores que hagan que el sistema se porte correctamente.

**Ved también**

Los algoritmos genéticos y las redes neuronales se verán con detalle en el módulo 5, "Introducción al aprendizaje computacional".

A continuación presentamos algunas indicaciones para construir estos sistemas (aunque no siempre son buenas):

**1)** El producto da resultados más cuidados que el mínimo. Esto es así por la forma de la función mínimo($x$, $y$) y de $x * y$.

**2)** La utilización de funciones paramétricas (como la familia de Yager) permite modificar fácilmente las funciones para adaptarlas a nuestros requerimientos. Además, también podemos aplicar técnicas de aprendizaje.

**3)** Cuando se construyen funciones de pertenencia a partir de observaciones, se pueden utilizar las llamadas *B-splines*. El uso de las *B-splines* resulta muy adecuado, especialmente cuando no hay errores en las observaciones. Cuando los hay, su uso no es tan recomendable. Aparte de esto, a menudo se utilizan funciones de pertenencia *triangulares* (construidas por trozos lineales, como las que hemos utilizado en los ejemplos de este apartado).

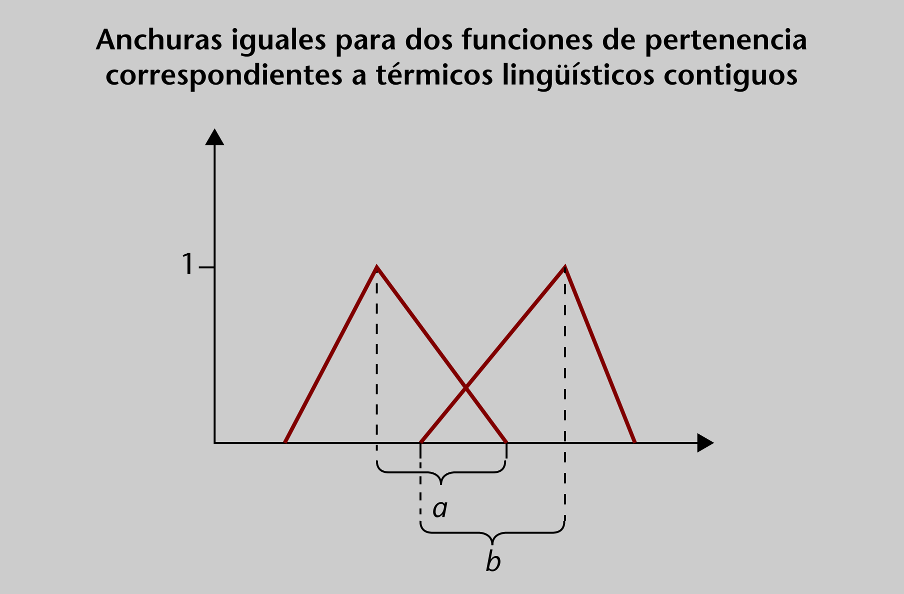
**4)** Dos funciones de pertenencia correspondientes a términos lingüísticos contiguos deberían tener intersección no nula. Si la intersección es nula, habrá discontinuidad en la conclusión.

**5)** Es bueno que las anchuras $a$ y $b$ de dos funciones de pertenencia correspondientes a términos lingüísticos contiguos sean la misma (podéis ver la figura 20). Es adecuado que el valor de pertenencia máximo de la intersección de estas funciones sea 0,5. Esto da buenos comportamientos en los sistemas de control.

**6)** En relación con el sistema de inferencia, resulta más fácil de implementar un sistema que primero aplica las reglas y después combina las conclusiones (como el método de Mamdani), que uno que primero combina las reglas y después aplica la regla resultante.

**7)** Los sistemas con reglas disyuntivas son más usados que los de reglas conjuntivas.

Figura 20




### 2.5.1. Sistemas complejos

La aplicación típica de un sistema difuso basado en reglas se basa en un conjunto de reglas como el que se ha visto hasta ahora: un conjunto de reglas con conjuntos difusos en el antecedente y en el consecuente donde todas las reglas se aplican a la vez para obtener la conclusión. Este tipo de sistemas plantea dificultades cuando se quieren modelizar o controlar sistemas complejos.

En primer lugar, encontramos que un sistema complejo requiere un gran número de variables de entrada porque hay muchos elementos para tener en cuenta antes de ofrecer una conclusión. Por otra parte, los cambios imprevistos en el entorno pueden motivar que las reglas que se habían definido dejen de ser válidas.

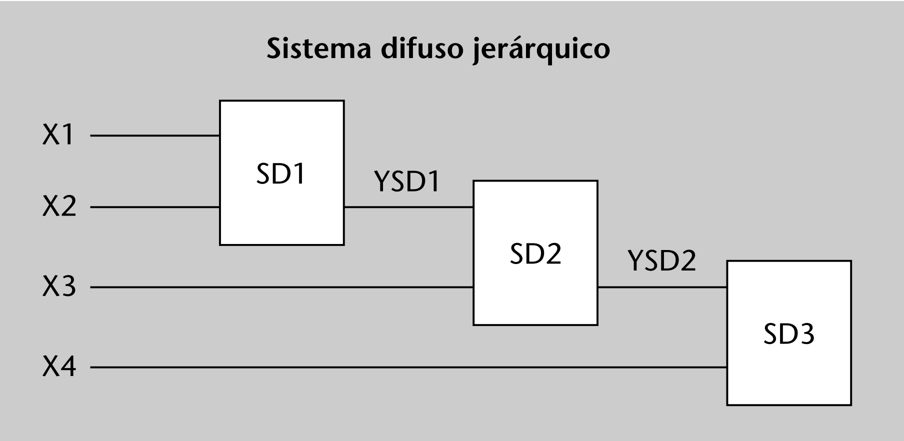
**Ejemplo**

Recordad que tenemos dos variables, tres términos lingüísticos por variable, con un total de nueve reglas.

Se debe subrayar que la dificultad de construir un sistema con muchas variables se debe a la cantidad de reglas que es necesario definir. Si tenemos un sistema con $n$ variables de entrada y cada variable tiene $m$ términos, esto dará que es necesario definir $m^n$ reglas para cubrir todo el dominio. Cuando $n$ es grande, el número de reglas que hemos de definir resulta excesivo.

Para tratar estos dominios complejos, tenemos los sistemas jerárquicos (o sistemas multietapa). Éstos se componen de una jerarquía de sistemas difusos (subsistemas) en los que las conclusiones de un nivel son utilizadas como entradas en el nivel siguiente. Las variables del sistema se van introduciendo en niveles sucesivos. En la figura 21, se da un esquema de un sistema difuso jerárquico con cuatro variables. Se puede ver que en el primer subsistema difuso (SD1) se utilizan las variables X1 y X2. La conclusión de este subsistema YSD1 junto con la variable X3 son las entradas del subsistema SD2. La salida de este sistema (YSD2) y la variable X4 son las entradas del susbsistema SD3. La salida de este sistema es la salida de todo el sistema.

Figura 21




**Ejemplo**

Mientras que si no consideramos un sistema jerárquico donde las variables tienen cada una tres términos necesitamos $3^4 = 81$ reglas, en el caso jerárquico (suponiendo que las variables YSD1 y YSD2 también tienen tres términos) tendremos nueve reglas para cada subsistema y, por lo tanto, un total de $3 * 9 = 27$ reglas.

Con esta construcción, el número de reglas totales necesarias es menor que en el otro caso.

De todos modos, el uso de estos sistemas plantea algunas dificultades. Enumeramos algunas a continuación. La primera es que la capacidad de modelización de un sistema jerárquico no es la misma que la de un sistema en el que todas las reglas se aplican a la vez. La elección de una jerarquía adecuada resulta, por lo tanto, muy importante. Además, el hecho de descomponer un sistema en subsistemas requiere considerar los parámetros de cada sistema (*t*-norma, *t*-conorma, inferencia de cada subsistema) y además también se deberá considerar si es necesario nitidificar o no la conclusión de cada subsistema.

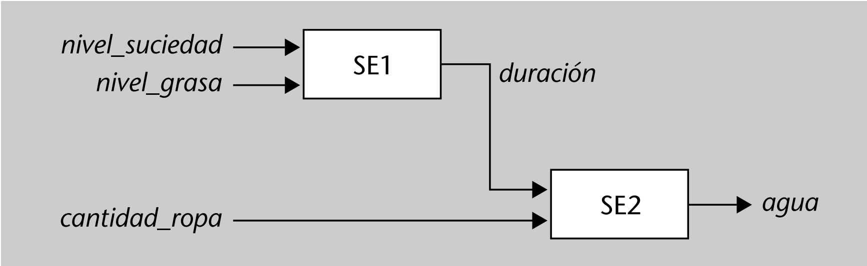
Para resolver el problema del entorno cambiante, una de las opciones es introducir nuevas variables en el sistema que permitan controlar los cambios en el entorno. Otra opción es construir un sistema jerárquico en el que un módulo controle el comportamiento de los otros y modifiquen sus reglas cuando las condiciones cambian.

**Ejemplo de sistema jerárquico difuso**

En la sección 2.2 se ha analizado un sistema de control de una lavadora que tenía como objetivo determinar la duración de un programa de lavado. Imaginemos que queremos diseñar un sistema algo más complejo que determine la cantidad de agua que necesita la lavadora. Para tomar esta decisión, el sistema tendrá en cuenta la duración del programa de lavado que hayamos determinado y la cantidad de ropa que se ha metido en la lavadora.

Con estas premisas, podemos diseñar un sistema de control difuso con dos bloques de reglas, tres variables de entrada, una variable de salida y una variable intermedia.

Queremos conocer la cantidad de agua que hay que emplear en la lavadora para el valor *cantidad_ropa* = 7, y los valores ya conocidos de *nivel_suciedad* = 70 y *nivel_grasa* = 50.




En este sistema, ya conocemos mucha información. Tenemos definidas las variables lingüísticas *nivel_suciedad*, *nivel_grasa* y *duración*. Nos faltará definir las variables *cantidad_ropa* y *agua*.

Como se ha dicho anteriormente, hay que decidir cómo deben trabajar los dos subsistemas en continuo. La forma habitual es que sólo se calcula el valor nítido de salida final del sistema. Para la salida de SE1, como veremos, no nos hará falta el valor nítido y meteremos directamente los valores de las activaciones de los términos de la variable duración en el nuevo bloque SE2. Por ese motivo, consideramos la variable *duración* como variable intermedia del sistema.

Definimos las variables que nos faltan (ved figura 22).

Variable *cantidad_ropa*:

<cantidad_ropa, $L_{cantidad\_ropa}$ = {poca, normal, mucha}, $U_{cantidad\_ropa}$ = [0, 10], $S_{cantidad\_ropa}$>

y ahora definimos las funciones de pertenencia de los tres términos:

$$\mu_{\text{poca}}(x)=\begin{cases} 1, & \text{si } 0\le x<2{,}5 \\ -0{,}4x+2, & \text{si } 2{,}5\le x<5 \\ 0, & \text{si } 5\le x\le 10 \end{cases}$$

$$\mu_{\text{normal}}(x)=\begin{cases} 0, & \text{si } 0\le x<2{,}5 \\ 0{,}4x-1, & \text{si } 2{,}5\le x<5 \\ -0{,}4x+3, & \text{si } 5\le x<7{,}5 \\ 0, & \text{si } 7{,}5\le x\le 10 \end{cases}$$

$$\mu_{\text{mucha}}(x)=\begin{cases} 0, & \text{si } 0\le x<5 \\ 0{,}4x-2, & \text{si } 5\le x<7{,}5 \\ 1, & \text{si } 7{,}5\le x<10 \end{cases}$$

Variable *agua*:

<agua, $L_{agua}$ = {poca, media, bastante, mucha}, $U_{agua}$ = [0, 50], $S_{agua}$>

y ahora definimos las funciones de pertenencia de los cuatro términos:

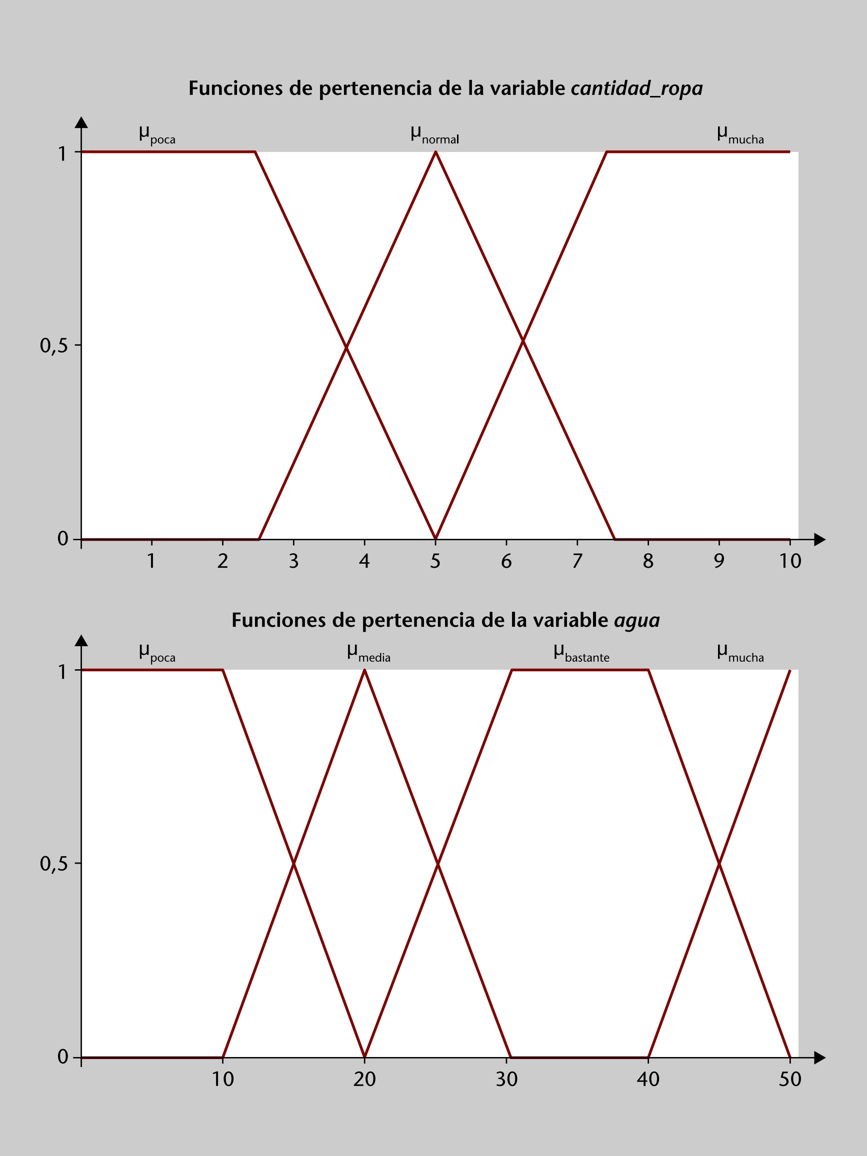
$$\mu_{\text{poca}}(x)=\begin{cases} 1, & \text{si } 0\le x<10 \\ -0{,}1x+2, & \text{si } 10\le x<20 \\ 0, & \text{si } 20\le x\le 50 \end{cases}$$

$$\mu_{\text{media}}(x)=\begin{cases} 0, & \text{si } 0\le x<10 \\ 0{,}1x-1, & \text{si } 10\le x<20 \\ -0{,}1x+3, & \text{si } 20\le x<30 \\ 0, & \text{si } 30\le x\le 50 \end{cases}$$

$$\mu_{\text{bastante}}(x)=\begin{cases} 0, & \text{si } 0\le x<20 \\ 0{,}1x-2, & \text{si } 20\le x<30 \\ 1, & \text{si } 30\le x<40 \\ -0{,}1x+5, & \text{si } 40\le x<50 \end{cases}$$

$$\mu_{\text{mucha}}(x)=\begin{cases} 0, & \text{si } 0\le x<40 \\ 0{,}1x-4, & \text{si } 40\le x\le 50 \end{cases}$$

Figura 22




El conjunto de reglas definido en SE2 es el siguiente:

| duración\cantidad_ropa | poca | normal | mucha |
|---|---|---|---|
| **corta** | poca | media | media |
| **media** | poca | media | bastante |
| **larga** | media | bastante | mucha |
| **extra_larga** | bastante | mucha | mucha |

Para el valor de entrada de *cantidad_ropa* = 7, se activan los términos lingüísticos *normal* con un grado 0,2 y *mucha* con un nivel 0,8.

Para ver cómo entran los valores de *duración* y *cantidad_ropa* en SE2, utilizaremos los grados de activación de los consecuentes de cada variable. No utilizaremos el valor nítido de SE1 obtenido anteriormente. Los valores de *duración* que ya habímos calculado son el valor *media* con valor (0,6), el valor *larga* (0,4) y el valor *extra_larga* (0,375).

Aprovechamos que tenemos dos variables de entrada para el bloque SE2 y mostramos las activaciones de las reglas en forma tabular. Entre paréntesis se muestran las activaciones. Aplicamos la *t*-norma del mínimo para obtener los consecuentes.

| duración\cantidad_ropa | poca | normal (0,2) | mucha (0,8) |
|---|---|---|---|
| corta | poca | media | media |
| media (0,6) | poca | media (0,2) | bastante (0,6) |
| larga (0,4) | media | bastante (0,2) | mucha (0,4) |
| extra_larga (0,375) | bastante | mucha (0,2) | mucha (0,375) |

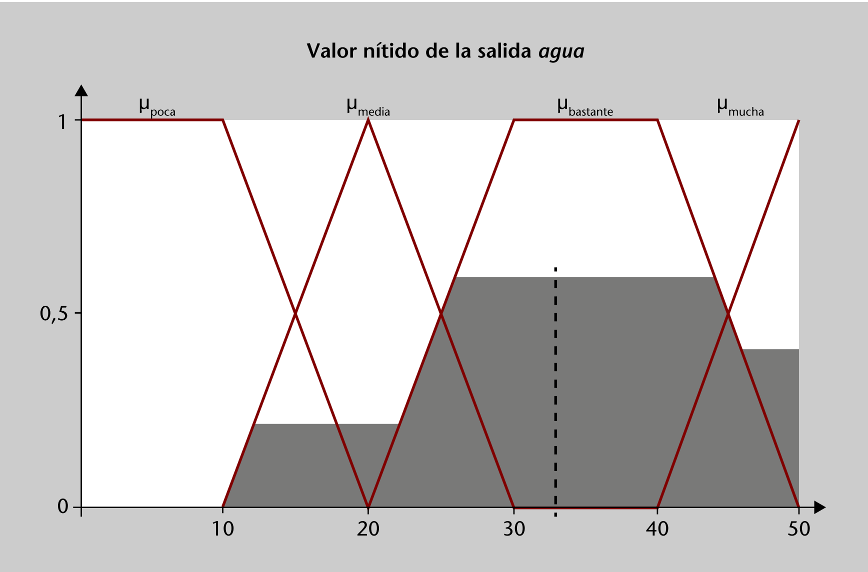
Así, pues, ya sólo nos queda obtener los grados de activación finales aplicando la *t*-conorma del máximo. Obtenemos que el término de la variable agua *media* queda activo con un grado 0,2, el valor *bastante* con un grado max(0,2; 0,6) = 0,6, y el valor *mucha* con un grado max(0,2; 0,4; 0,375) = 0,4.

La función de pertenencia resultante es (ved figura 23):

$$\mu_{\text{agua}}(x) = \begin{cases} 0, & \text{si } 0 \le x < 10 \\ 0,1x - 1, & \text{si } 10 \le x \le 12 \\ 0,2, & \text{si } 12 \le x \le 22 \\ 0,1x - 2, & \text{si } 22 \le x \le 26 \\ 0,6, & \text{si } 26 \le x \le 44 \\ -0,1x + 5, & \text{si } 44 \le x \le 46 \\ 0,4, & \text{si } 46 \le x \le 50 \end{cases}$$

La cantidad de agua que gastará la lavadora es de (el valor nítido) **33,42** litros.

Figura 23

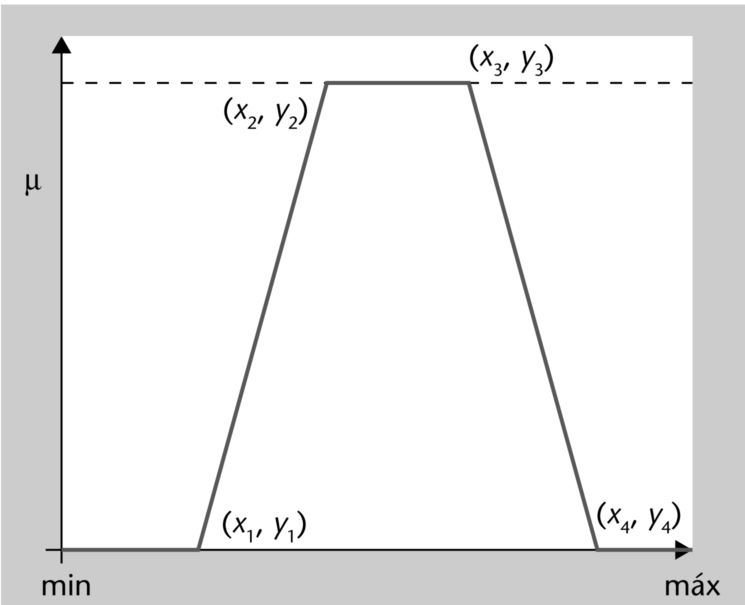


## Ejercicios de autoevaluación

**1.** Considerad un sistema experto difuso con *tres* variables de entrada (VarA, VarB, VarC) y *una* de salida (*Out*). A continuación se detallan los puntos de corte y los términos lingüísticos asociados a estas cuatro variables.

| Variable | Tipo | Rango | Término lingüístico: puntos (*x*, *y*) |
|---|---|---|---|
| **VarA** | Entrada | Mín: 0<br>Máx: 15 | VL: (0, 1) (2, 1) (3, 0) (15, 0)<br>L: (0, 0) (3, 1) (4, 1) (7, 0) (15, 0)<br>M: (0, 0) (5, 0) (7, 1) (9, 1) (10, 0) (15, 0)<br>H: (0, 0) (9, 0) (10, 1) (12, 1) (15, 0)<br>VH: (0, 0) (12, 0) (13, 1) (15, 1) |
| **VarB** | Entrada | Mín: 0<br>Máx: 10 | L: (0, 1) (1, 1) (4, 0) (10, 0)<br>M: (0, 0) (1, 0) (5, 1) (7, 0) (10, 0)<br>H: (0, 0) (6, 0) (9, 1) (10, 1) |
| **VarC** | Entrada | Mín: -5<br>Máx: 10 | VL: (-5, 1) (-4, 1) (-2, 0) (10, 0)<br>L: (-5, 0) (-4, 0) (-2, 1) (0, 1) (4, 0) (10, 0)<br>M: (-5, 0) (0, 0) (2, 1) (4, 0) (10, 0)<br>H: (-5, 0) (2, 0) (5, 1) (10, 1) |
| ***Out*** | Salida | Mín: 0<br>Máx: 10 | L: (0, 1) (3, 1) (6, 0) (10, 0)<br>M: (0, 0) (4, 0) (5, 1) (6, 0) (10, 0)<br>H: (0, 0) (3, 0) (7, 1) (10, 1) |

La secuencia de puntos de los términos se lee de la forma siguiente:




Tomad un sistema Mamdani con *t*-norma mínimo y *t*-conorma máximo.

**a)** Representad gráficamente las variables lingüísticas de entrada y de salida. Calculad todas las funciones de pertenencia.

**b)** Detallando todo el proceso de inferencia que se realiza, representad la salida del sistema y el valor nítido, teniendo en cuenta estos valores de entrada:

(VarA, VarB, VarC) = (6, 2, 3)

El conjunto de reglas de este sistema experto difuso se detalla a continuación.

R1: SI (VarA es L) o (VarB es L) o (VarC es L) ENTONCES (Out es L)
R2: SI (VarA es M) o (VarB es L) o (VarC es M) ENTONCES (Out es L)
R3: SI (VarA es M) o (VarB es L) o (VarC es L) ENTONCES (Out es L)

R4: SI (VarA es VL) o (VarB es L) o (VarC es VL) ENTONCES (Out es L)
R5: SI (VarA es M) o (VarB es M) o (VarC es L) ENTONCES (Out es M)
R6: SI (VarA es M) o (VarB es M) o (VarC es H) ENTONCES (Out es H)
R7: SI (VarA es H) o (VarB es H) o (VarC es H) ENTONCES (Out es H)

# Solucionario

## Ejercicios de autoevaluación

**1. a)** Las funciones de pertenencia son las siguientes:

**Variable: VarA**

**Funciones de pertenencia ($\mu_{\text{término}}$)**

$$\mu_{VL}(x) = \begin{cases} 1 \text{ si } 0 < x \leq 3 \\ -x + 3 \text{ si } 3 < x \leq 4 \\ 0 \text{ si } x > 4 \end{cases}$$

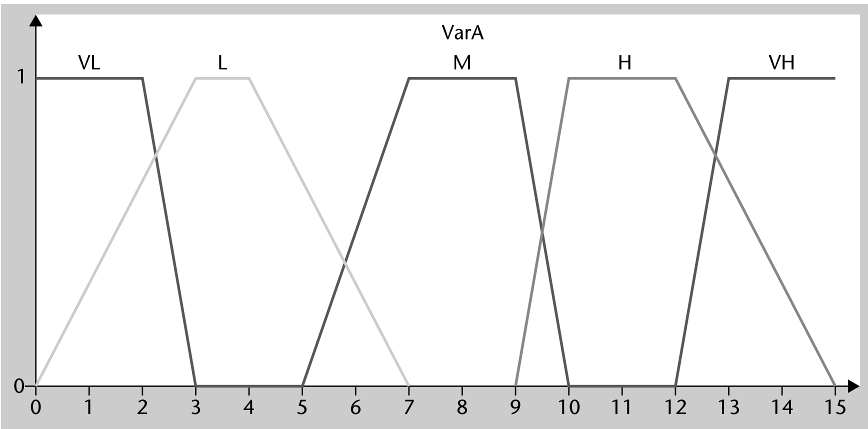
$$\mu_{L}(x) = \begin{cases} 0{,}33x \text{ si } 0 < x \leq 3 \\ 1 \text{ si } 3 < x < 4 \\ -0{,}33x + 2{,}33 \text{ si } 4 < x \leq 7 \\ 0 \text{ si } x > 7 \end{cases}$$

$$\mu_{M}(x) = \begin{cases} 0 \text{ si } x < 5 \\ 0{,}5x - 2{,}5 \text{ si } 5 < x \leq 7 \\ 1 \text{ si } 7 < x \leq 9 \\ -x + 10 \text{ si } 9 < x \leq 10 \\ 0 \text{ si } x > 10 \end{cases}$$

$$\mu_{H}(x) = \begin{cases} 0 \text{ si } x < 9 \\ x - 9 \text{ si } 9 < x \leq 10 \\ 1 \text{ si } 10 < x \leq 12 \\ -0{,}33x + 5 \text{ si } 12 < x \leq 15 \\ 0 \text{ si } x > 10 \end{cases}$$

$$\mu_{VH}(x) = \begin{cases} 0 \text{ si } x < 12 \\ x - 12 \text{ si } 12 < x \leq 13 \\ 1 \text{ si } 13 < x < 15 \end{cases}$$

Representación gráfica de la variable VarA




**Variable: VarB**

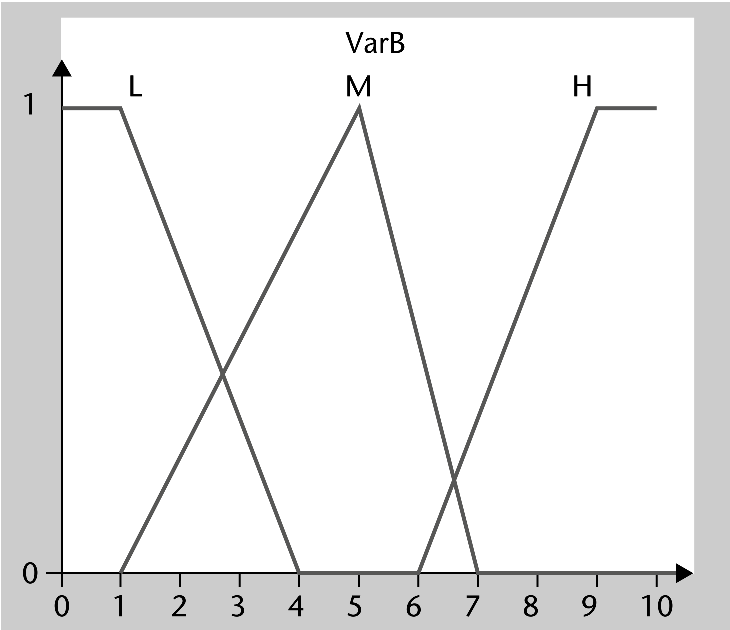
**Funciones de pertenencia ($\mu_{\text{término}}$)**

$$\mu_L(x) = \begin{cases} 1 \text{ si } 0 < x \leq 1 \\ -0{,}33x + 1{,}33 \text{ si } 1 < x \leq 4 \\ 0 \text{ si } x > 4 \end{cases}$$

$$\mu_M(x) = \begin{cases} 0 \text{ si } x < 1 \\ 0{,}25x - 0{,}25 \text{ si } 5 < x < 7 \\ -0{,}5x + 3{,}5 \text{ si } 5 < x \leq 7 \\ 0 \text{ si } x > 7 \end{cases}$$

$$\mu_H(x) = \begin{cases} 0 \text{ si } x < 6 \\ 0{,}33x - 2 \text{ si } 6 < x < 9 \\ 1 \text{ si } x > 9 \end{cases}$$

Representación gráfica de la variable VarB




**Variable: VarC**

**Funciones de pertenencia ($\mu_{\text{término}}$)**

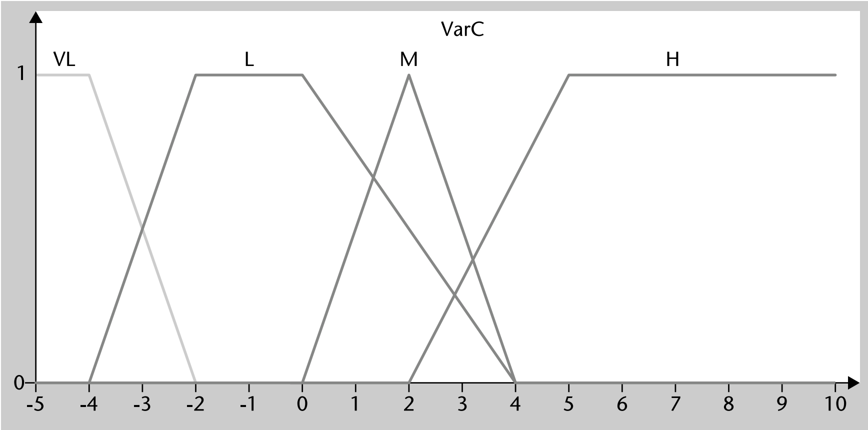
$$\mu_{VL}(x) = \begin{cases} 1 \text{ si } -5 < x \leq -4 \\ -0{,}5x - 1 \text{ si } -4 < x \leq -2 \\ 0 \text{ si } x > -2 \end{cases}$$

$$\mu_L(x) = \begin{cases} 0 \text{ si } x < -4 \\ 0{,}5x + 2 \text{ si } -4 < x \leq -2 \\ 1 \text{ si } -2 < x \leq 0 \\ -0{,}25x + 1 \text{ si } 0 < x \leq 4 \\ 0 \text{ si } x > 4 \end{cases}$$

$$\mu_M(x) = \begin{cases} 0 \text{ si } x < 0 \\ 0{,}5x \text{ si } 0 < x \leq 2 \\ -0{,}5x + 2 \text{ si } 2 < x \leq 4 \\ 0 \text{ si } x > 4 \end{cases}$$

$$\mu_H(x) = \begin{cases} 0 \text{ si } x \leq 2 \\ 0{,}33x - 0{,}66 \text{ si } 2 < x \leq 5 \\ 1 \text{ si } x > 5 \end{cases}$$

Representación gráfica de la variable VarC




**Variable: *Out***

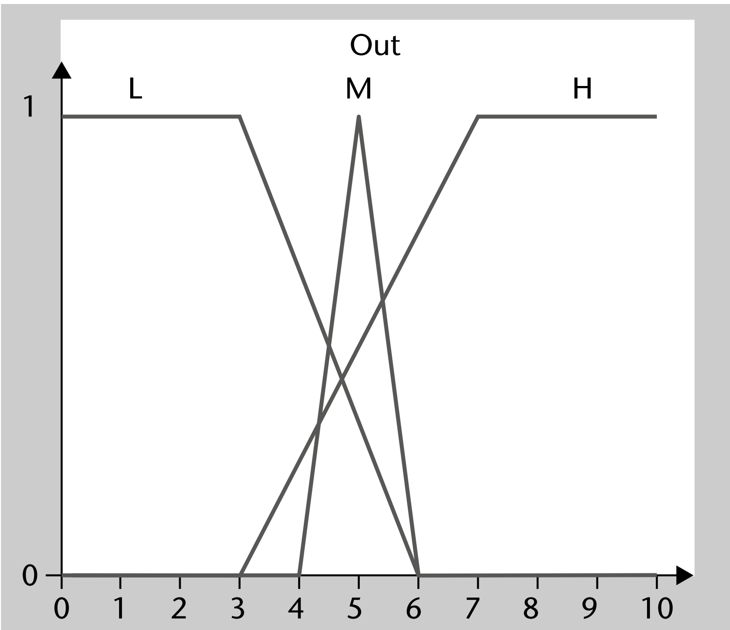
**Funciones de pertenencia ($\mu_{\text{término}}$)**

$$\mu_L(x) = \begin{cases} 1 \text{ si } x < 3 \\ -0{,}33x + 2 \text{ si } 3 < x \leq 6 \\ 0 \text{ si } x > 6 \end{cases}$$

$$\mu_M(x) = \begin{cases} 0 \text{ si } x < 4 \\ x - 4 \text{ si } 4 < x \leq 5 \\ -x + 6 \text{ si } 5 < x \leq 6 \\ 0 \text{ si } x > 6 \end{cases}$$

$$\mu_H(x) = \begin{cases} 0 \text{ si } x < 3 \\ 0{,}25x - 0{,}75 \text{ si } 3 < x \leq 7 \\ 1 \text{ si } x > 7 \end{cases}$$

Representación gráfica *Out*



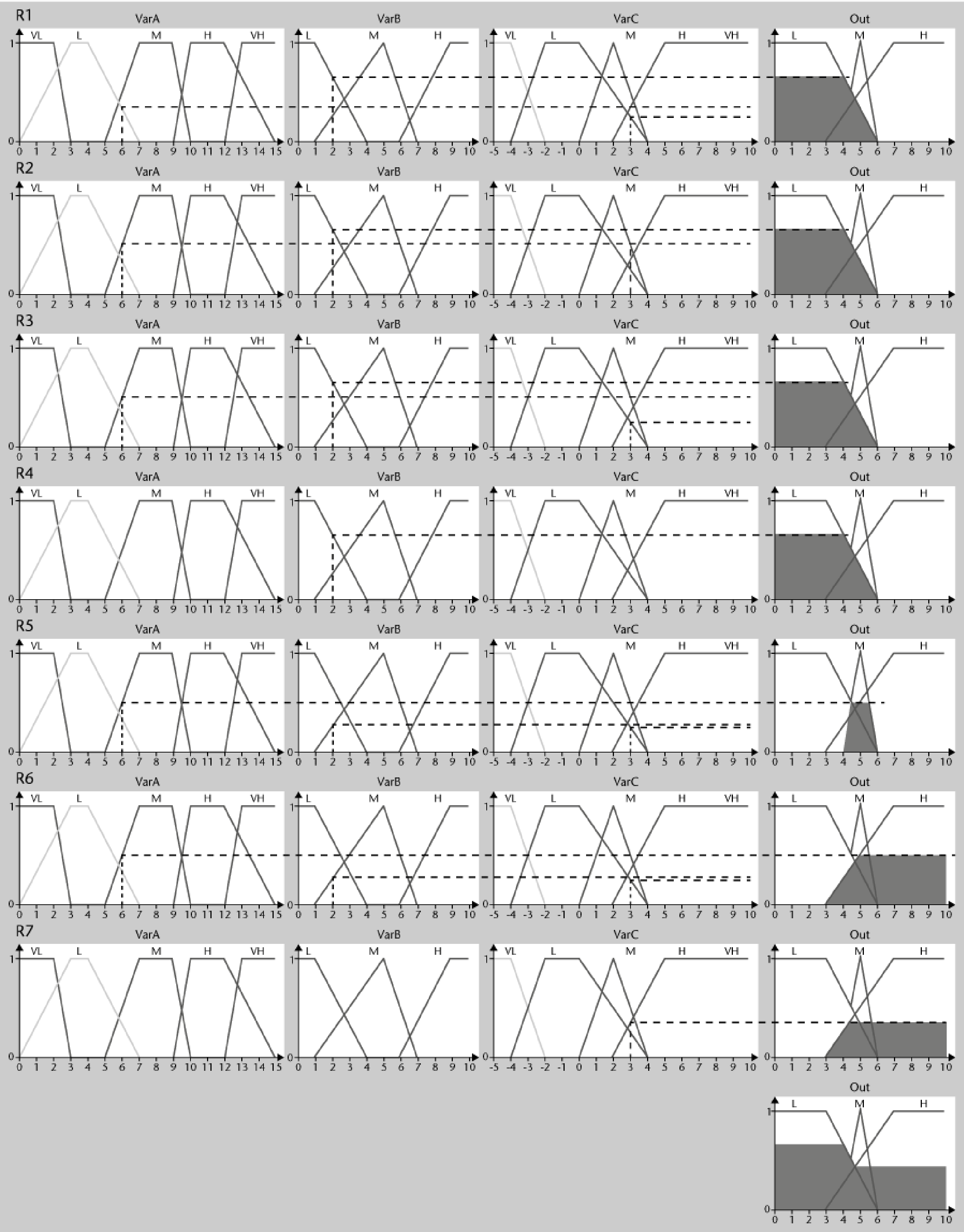

**b)** Cálculo del valor nítido.

Cálculo del valor nítido con las mismas entradas (6,2 y 3), cambiando las reglas de funcionamiento por las reglas del sistema difuso 2.

El sistema de reglas contiene OR en lugar de AND. Por lo tanto, los consecuentes se obtienen aplicando las *t*-conormas en lugar de las *t*-normas.

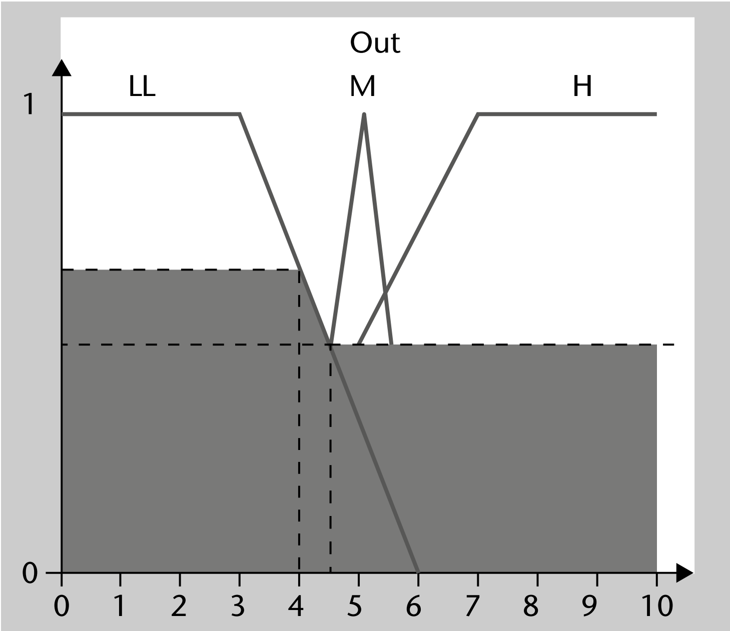
R1: Se ACTIVA. ***Out* es L truncada en 0,666**
R2: Se ACTIVA. *Out* es L truncada en 0,666
R3: Se ACTIVA. *Out* es L truncada en 0,666
R4: Se ACTIVA. *Out* es L truncada en 0,666
R5: Se ACTIVA. ***Out* es M truncada en 0,5**
R6: Se ACTIVA. ***Out* es H truncada en 0,5**
R7: Se ACTIVA. *Out* es H truncada en 0,333

Si mostramos gráficamente las activaciones y la aplicación de las *t*-conormas, tenemos las transiciones siguientes.




Finalmente, en este caso, la representación de la variable *Out*, su función de pertenencia y el valor nítido resultante son:

Representación gráfica de la variable *Out*




**Función de pertenencia ($\mu_{\text{término}}$)**

$$\mu_L(x) = \begin{cases} 0{,}66 \text{ si } 0 < x \leq 4 \\ -0{,}33x + 2 \text{ si } 4 < x < 4{,}5 \\ 0{,}5 \text{ si } 4{,}5 < x \leq 10 \end{cases}$$

**Defuzzificación** = 4.638.

# Bibliografía

## Bibliografía básica

**Klir, G.; Yuan, B.** (1995). *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice-Hall.

**Neapolitan, R.** (1990). *Probabilistic Reasoning in Expert Systems: Theory and Algorithms*. John-Wiley & Sons Inc.

## Bibliografía complementaria

**Dubois, D.; Prade, H.** (1988). *Possibility Theory. An aproach to Computerized Processing of Uncertainty*. Nova York: Plenum Press.

**Kahraman, C. (ed).** (2006). *Fuzzy Applications in Industrial Engineering*. Springer.

**López de Mántaras, R.** (1990). *Aproximate Reasoning Models. Ellis Horwood series in Artificial Intelligence*. Ellis Horwood Limited.

**Pearl, J.** (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers, Inc.

**Trillas, E.; Alsina, C.; Terricabras, J. M.** (1995). *Introducción a la lógica borrosa*. Ariel.

## Software

**jFuzzyLogic**. http://jfuzzylogic.sourceforce.net

**Wiris CalcMe**. https://calcme.com/

**fuzzyTECH**. https://www.fuzzytech.com/