4 Lógica difusa

4.1 Introducción

La lógica difusa o *fuzzy logic* (también se le llama lógica borrosa, es una rama de la inteligencia artificial que permite lidiar con situaciones y datos inciertos en la toma de decisiones y el razonamiento. A diferencia de la lógica tradicional, que trabaja con valores binarios (verdadero o falso, si o no), la lógica difusa aborda la naturaleza borrosa y subjetiva de muchos problemas del mundo real al permitir la representación de grados de verdad entre 0 y 1.
Por ejemplo, en la vida real usamos cualificadores del tipo "muy", "poco", "mucho", etc. Esos son grados intermedios entre la verdad absoluta (1) y la falsedad completa (0). Esa cuantificación intermedia de variables cualitativas es lo que aborda la logica difusa.

Se trata de una herramienta matemática fundamental para dar una respuesta a la limitación de la lógica clásica, pudiendo manejar la incertidumbre y la vaguedad inherentes a numerosos contextos. Existen multitud de aplicaciones en una amplia variedad de campos, desde sistemas de control automático (como automóviles y procesos industriales, donde las condiciones pueden ser variables y difíciles de modelar con precisión), diagnóstico médico (diagnóstico de enfermedades, donde los síntomas pueden ser vagos y difíciles de interpretar de manera estrictamente binaria) y sistemas de recomendación (para adaptar recomendaciones a preferencias cambiantes y difusas de los usuarios), entre otros.
Un ejemplo sería el control del termoestato de un aire acondicionado. En la lógica tradicional, le diríamos que cuando la temperatura llegue a 27 grados se apague. Sin embargo, cuando llegue a esa temperatura aproximada, el aire se estaría apagando y encendiendo continuamente con variaciones pequeñas de temperatura, siendo muy ineficiente energéticamente. Esto se podría resolver mediante lógica difusa.

4.1 Introducción 57
4.2 Conjuntos difusos 58
4.2.1 Tipos de funciones de pertenencia 60
4.2.2 Etiquetado lingüistico y modificaciones a las funciones de pertenencia 63
4.3 Operaciones con conjuntos difusos 65
4.3.1 Unión 65
4.3.2 Intersección 66
4.3.3 Complementario 67
4.3.4 Otras diferencias entre conjuntos clásicos y difusos 68
4.4 Relaciones y razonamiento difuso 69
4.4.1 Relaciones difusas 69
4.4.2 Razonamiento difuso . . 70
4.5 Implicación e inferencia difusa 70
4.5.1 Implicación de Mamdani 70
4.6 Defuzzificación 72
4.7 Modelado difuso 74
4.7.1 Modelo de Mamdani . . . 74
4.7.2 Modelo TSK 77

Uno de los conceptos fundamentales en la lógica difusa es el de "conjunto difuso". A diferencia de los conjuntos convencionales, en los que un elemento pertenece o no pertenece claramente a un conjunto, los conjuntos difusos permiten que los elementos tengan grados de pertenencia parciales, lo que refleja de manera más precisa la incertidumbre presente en el mundo real. Además, esta lógica hace uso de funciones de pertenencia y operadores difusos para representar y manipular información imprecisa. En este capítulo se veran las diferentes representaciones de estos conjuntos.

Finalmente, se explican técnicas de inferencia difusa para tratar la incertidumbre y la vaguedad en situaciones del mundo real, y técnicas de defuzzificación, para convertir los resultados borrosos o difusos en valores concretos y no ambiguos (valores numéricos en un rango específico).

4.2 Conjuntos difusos

Definiciones

Conjunto difuso: conjunto que sirve para realizar una evaluación cualitativa de alguna cantidad física.

Para tratar de explicar la lógica difusa, vamos a usar un ejemplo concreto: determinar si una persona está delgada.

Según la teoría de la lógica clásica, se establecería el conjunto "personas obesas" a aquel conjunto al que pertenece toda persona con un peso mayor que un umbral (80 kg, por ejemplo). Toda persona con un peso igual o menor quedaría fuera de ese conjunto. De esta forma, una persona que pese 81 kg pertenecería al conjunto de personas obesas, y otra de 79 kg no. No tiene mucho sentido decir que una persona es obesa, y otra que pesa 2 kg menos no lo es. El enfoque de la lógica difusa viene a corregir esto, considerando que el conjunto "personas obesas" es un conjunto que no tiene una delimitación (frontera) clara para pertenecer o no a él. Para tratar el problema, se definen previamente dos conceptos: grado de pertenencia y función de pertenencia.

Definiciones

Grado de pertenencia: es un número que indica en qué medida un elemento dado pertenece a un conjunto difuso. Este número es un valor entre 0 y 1, donde 0 representa una no pertenencia (ausencia total) y 1 representa una pertenencia completa (presencia total). Los grados de pertenencia intermedios representan niveles de pertenencia parciales.
Función de pertenencia: es una curva o función matemática que describe cómo varía el grado de pertenencia a lo largo del rango de valores posibles de una variable. Es decir, es una representación gráfica de cómo un elemento se relaciona con un conjunto difuso en función de su valor en una variable. Pueden tomar diversas formas, como triangular, trapezoidal, gaussiana, sigmoide, etc.

Notación: Un conjunto difuso A en el dominio χ (discreto o continuo) se define como sigue:
$A = \{(x, \mu_A(x)) | x \in \chi)\}$, donde
$\mu_A(x)$ es la función de pertenencia para el conjunto difuso A tal que $\mu_A : \chi \longrightarrow [0, 1]$. Es decir, asigna a cada elemento $x \in \chi$ un valor entre 0 y 1, que sería el grado de pertenencia de x al conjunto A.

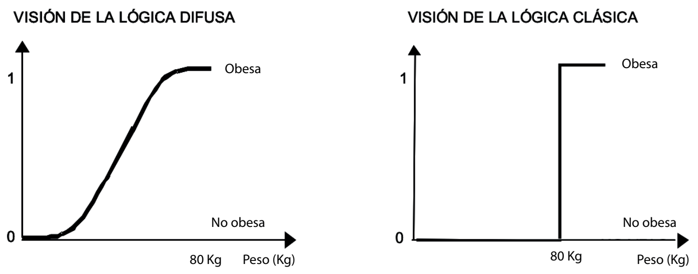

Figure 4.1: Visión de la lógica. Representación gráfica de la lógica clásica y difusa en el ejemplo.

Por lo tanto, siguiendo el ejemplo anterior, una persona con un peso de 79 kg podría pertenecer al conjunto difuso "personas obesas" con un grado de pertenencia de 0.8, una persona que pese 81 kg tendrá un grado de 0.85 y una que pese 60 kg tendrá un grado de pertenencia de 0.05. En definitiva, la persona de 79 kg diremos que es una persona obesa en gran medida.

Algunos conceptos de la teoría clásica sirven lara los conjuntos difusos, pero hay otros conceptos nuevos que conforman la teoría de conjuntos difusos. A continuación se definen los más importantes:

Definiciones

Variable lingüística: concepto que se va a calificar de forma difusa. En el ejemplo es el "peso", como podrían ser otros ejemplos de "joven" o "bajo".
Universo de discurso: rango de valores que pueden tomar los elementos que poseen la propiedad expresada por la variable lingüística. En el ejemplo, sería un rango de valores que comprende todos los pesos posibles de una persona (por ejemplo, de 0 a 350 kg)
Valor lingüístico: las diferentes clasificaciones que se pueden efectuar sobre la variable lingüística. En el ejemplo, podría ser "delgada", "normal" y "obesa", o en otros ejemplos "joven", "adulto", "anciano".
Conjunto nítido: un conjunto es nítido si su función de pertenencia toma valores en 0,1, y difuso si toma valores en [0,1].
Alfa-corte: es el conjunto de elementos que pertenecen al conjunto difuso dado, con grado mayor o igual que alfa ($A_\alpha = \{x \in \chi / \mu_A(x) \geq \alpha\}$). Si fuese mayor que alpha, se denominaría *alfa corte estricto*.
Soporte: es el conjunto de todos los puntos $x \in \chi$ tales que su función de pertenencia es mayor que 0. En otras palabras, es el conjunto nítido de elementos que tienen grado de pertenencia estrictamente mayor que 0 (alfa-corte estricto de nivel 0).
Núcleo: es el conjunto nítido de elementos que tienen grado de pertenencia 1 (alfa-corte de nivel 1).
Altura: es el valor más grande de la función de pertenencia.

Conjunto normal o normalizado: es aquel conjunto que su núcleo es no vacío, es decir, se puede encontrar un punto $x \in \chi$ tal que $\mu_A(x) = 1$ encontrar un punto (su altura es 1).
Punto de cruce: es el elemento x de χ para el cual $\mu_A(x) = 0.5$.
Conjunto difuso singleton: aquel conjunto cuyo soporte es un solo punto que cumple la condición $\mu_A(x) = 1$. Es decir, el soporte coincide con el núcleo y tienen un único punto.
Convexidad: un conjunto es convexo si no hay discontinuidades.

Un ejemplo gráfico de un problema genérico se puede ver en la siguiente imagen:

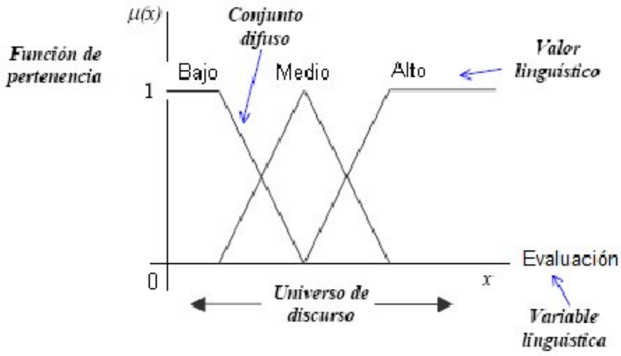

Figure 4.2: Conceptos específicos de la lógica difusa. Representación gráfica de un problema genérico.

4.2.1 Tipos de funciones de pertenencia

En la lógica difusa, los conjuntos difusos sirven para modelar diferentes problemas o comportamientos. Para ello, se usan funciones de pertenencia, ya sea de forma discreta o continua. Por facilidad de computación y representación, hay una serie de funciones típicas que se usan más frecuentemente, como las que se describen a continuación:

Función Gamma.

Su uso es para representar conjuntos como el del ejemplo: "persona obesa". Sigue la siguiente distribución:

$$\mu(x) = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{m-a} & \text{si } a < x < m \\ 1 & \text{si } x \geq m \end{cases}$$

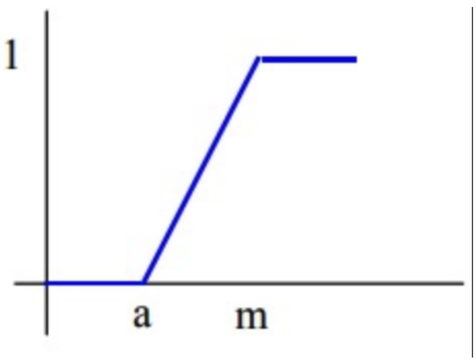

Figure 4.3: Función de pertenencia Gamma.

Función L.

Es la inversa de la función Gamma. Su uso es para representar conjuntos inversos a la gamma, como el del ejemplo: "persona delgada". Sigue la siguiente distribución:

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

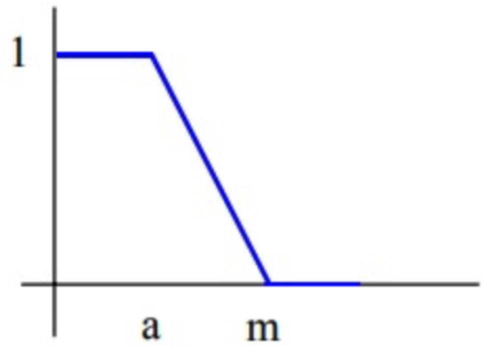

Figure 4.4: Función de pertenencia L.

Función Triangular.

Se usa cuando se desea hacer una transición entre lo verdadero y falso de forma suave, como podría ser en el ejemplo del termoestado: si se desea que cuando llegue a 27 se apague, para que no sea brusco y no este continuamente encendiéndose y apagándose, se puede establecer esta función en un rango pequeño de 26-28 grados. Sigue la siguiente distribución:

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

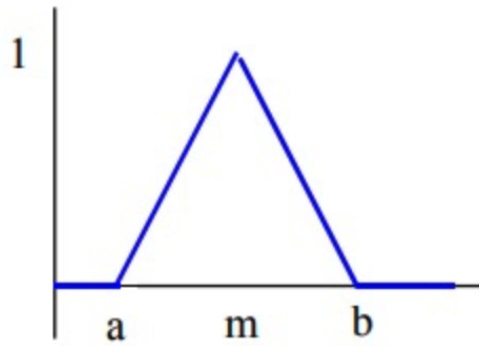

Figure 4.5: Función de pertenencia Triangular.

Función Trapezoidal.

Es la más usada de todas. Es parecida a la triangular, pero se usa cuando se desea que la zona transición sea más prolongada, como podría ser cuando se desea establecer una "velocidad normal en ciudad", que en lugar de ser un pico como ejemplo del termoestado, se trata de un rango de velocidades como Verdadero (30-50 km/h, por ejemplo). Sigue la siguiente distribución:

$$\mu(x) = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } a < x \leq b \\ 1 & \text{si } b < x \leq c \\ \frac{d-x}{b-c} & \text{si } c < x \leq d \\ 0 & \text{si } x > d \end{cases}$$

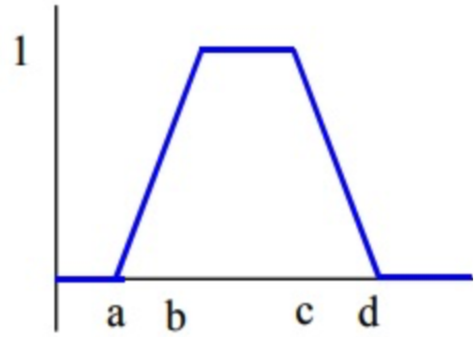

Figure 4.6: Función de pertenencia Trapezoidal.

Función Sigmoidea.

Las anteriores son lineales a trozos. Si se quiere modelar un problema no lineal a trozos, un par de buenas opciones son la sigmoidea, su opuesta o la función PI, que sería la suma de ambas (subida y bajada, similar a la triangular y trapezoidal. Como ejemplo, la sigmoidea se usa para transiciones suaves del falso al verdadero, como podria ser la "satisfacción" de un cliente (como va cambiando conforme a una bajada del precio, por ejemplo). Sigue la siguiente distribución:

$$\mu(x) = \begin{cases} 0 & \text{si } x \leq a \\ 2(\frac{x-a}{c-a})^2 & \text{si } a < x \leq \frac{a+c}{2} \\ 1 - 2(\frac{x-a}{c-a})^2 & \text{si } \leq \frac{a+c}{2} < x < c \\ 1 & \text{si } x \geq c \end{cases}$$

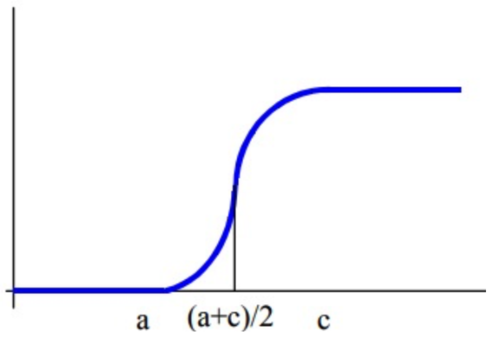

Figure 4.7: Función de pertenencia Sigmoidea.

Finalmente, como se ha dicho al principio, un conjunto A puede expresarse de forma discreta (4.1) o continua (4.2), mediante sumatorio o integrales del ratio de la función de pertenencia y el elemento, de la siguiente forma:

$$A = \sum_{U} \frac{\mu_A(X_i)}{X_i} \tag{4.1}$$

$$A = \int_{U} \frac{\mu_A(X_i)}{X_i} \tag{4.2}$$

Por lo tanto, muy a *grosso modo*, se puede simplificar el uso de un tipo u otro de la siguiente forma: las funciones Gamma y L se usan para calificar valores lingüísticos extremos, como "recién nacido" o "anciano", respectivamente. Las funciones triangular y trapezoidal se usan para describir valores intermedios (como "adolescente", "joven", "adulto").

4.2.2 Etiquetado lingüistico y modificaciones a las funciones de pertenencia

Los conjuntos difusos tienen asociados normalmente etiquetas lingüisticas, que normalmente en lenguaje natural tienen asociados adverbios. Para componer la función de pertenencia con estos etiquetados, se suelen usar operaciones aritméticas muy simples. Por ejemplo:

- Grado superlativo: potencia al cuadrado. Ejemplo: "Persona MUY obesa" modificaría la función de pertenencia de la forma $\mu_{MUYobesa}(x) = (\mu_{obesa(x)})^2$, o lo que es lo mismo, $\mu_{MUY}(x) = (\mu_(x))^2$.
- Grado no especificado: raiz cúbica. Ejemplo: "Persona ALGO obesa" modificaría la función de pertenencia de la forma $\mu_{ALGOobesa}(x) = \sqrt[3]{(\mu_{obesa(x)}}$.
- Grado de inexactitud aproximada: raiz cuadrada. Ejemplo: "Persona MAS O MENOS obesa" modificaría la función de pertenencia de la forma $\mu_{MASOMENOSobesa}(x) = \sqrt{(\mu_{obesa(x)}}$.
- Grado altamente superlativo: potencia al cubo. Ejemplo: "Persona EXTREMADAMENTE obesa" modificaría la función de pertenencia de la forma $\mu_{EXTREMADAMENTEobesa}(x) = (\mu_{obesa(x)})^3$.
- Negación: inversa de la función. Ejemplo: "Persona NO obesa" modificaría la función de pertenencia de la forma $\mu_{NOobesa}(x) = 1 - \mu_{obesa(x)}$.

Este etiquetado puede ser personalizado para cada sistema difuso. Por ejemplo, hay modificaciones que se aplican frecuentemente a los conjuntos difusos, como los siguientes:

Normalización de un conjunto difuso. Convierte un conjunto difuso no normalizado a normalizado, dividiendo por la altura del conjunto.

Concentración de un conjunto difuso. Cuando se desea que la función de pertenencia tome valores más pequeños de los actuales, focalizándose en los valores mas grandes, se compone con una función de pertenencia tal que (μ^p, siendo $p > 1$)

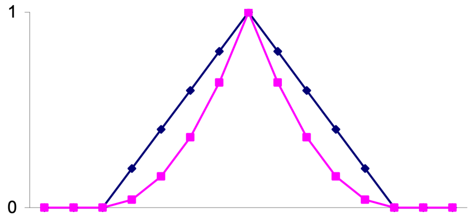

Figure 4.8: Concentración en un conjunto difuso. En azul, la función de activación original, y en rosa la función modificada por concentración.

Dilatación de un conjunto difuso. Si se desea un efecto contrario a la concentración (tomar valores más grandes para focalizarse en los pequeños), se compone una función de pertenencia tal que (μ^p, siendo $0 < p < 1$).

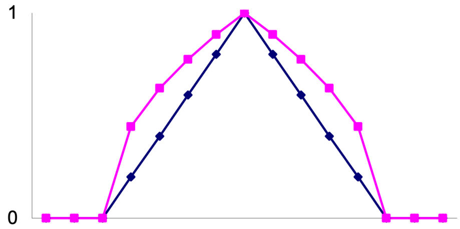

Figure 4.9: Dilatación en un conjunto difuso. En azul, la función de activación original, y en rosa la función modificada por dilatación.

Intensificación del contraste de un conjunto difuso. Cuando se desea un efecto híbrido (dilatar una parte y concentrar otra), se sitúa un umbral (por ejemplo 0.5), y se disminuyen los valores menores al umbral y se aumentan los mayores que el umbral. La composición de la función quedaría de la siguiente forma:

$$\mu_{mod}(x) = \begin{cases} 2^{p-1}\mu(x)^p & \text{si } \mu(x) \leq 0.5 \\ 1 - 2^{p-1}(1-\mu(x))^p & \text{en otro caso} \end{cases}$$

siendo $p > 1$ (cuanto mayor es p, mayor es la intensificación.

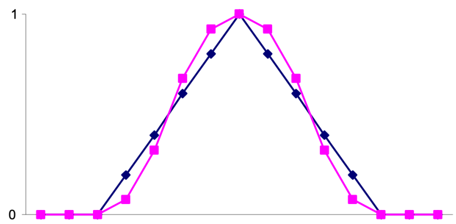

Figure 4.10: Intensificación del contraste en un conjunto difuso. En azul, la función de activación original, y en rosa la función modificada por intensificación.

Difuminación de un conjunto difuso. Cuando se desea un efecto contrario a la intensificación, es decir, se disminuyen los valores mayores al umbral y se aumentan los menores que el umbral. La composición de la función quedaría de la siguiente forma:

$$\mu_{mod}(x) = \begin{cases} \sqrt{\frac{\mu(x)}{2}} & \text{si } \mu(x) \leq 0.5 \\ 1 - \sqrt{\frac{1-\mu(x)}{2}} & \text{en otro caso} \end{cases}$$

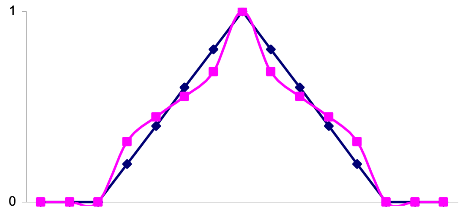

Figure 4.11: Difuminación en un conjunto difuso. En azul, la función de activación original, y en rosa la función modificada por intensificación.

4.3 Operaciones con conjuntos difusos

De forma análoga a la teoría clásica de conjuntos, los conjuntos difusos tienen las operaciones de unión, intersección o complementario, entre otras. A continuación se realiza un repaso de las mismas y sus particularidades en este tipo de conjuntos.

4.3.1 Unión

La teoría clásica define la unión como la pertenencia de un elemento a un conjunto unión si pertenece, al menos, a uno de los conjuntos. En su análogo a lógica difusa, se refiere a su función de pertenencia. La ecuación de la unión u de dos conjuntoas A y B quedaría de la siguiente forma (4.3.1):

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

donde u: [0,1] x [0,1] $\longrightarrow [0,1]$(4.3)

Hay una serie de propiedades que se deben cumplir, partiendo de $\forall \alpha, \beta, \gamma \in [0,1]$:

- Concordante con el caso nítido: $u(0,1) = u(1,1) = u(1,0) = 1$; mientras $u(0,0) = 0$
- Propiedad commutativa: $u(\alpha,\beta) = u(\beta,\alpha)$
- Propiedad asociativa: $u(\alpha, u(\beta,\gamma)) = u(u(\alpha,\beta),\gamma)$
- Identidad: $u(\alpha, 0) = \alpha$
- Monotonía: si $\alpha \leq \alpha'$ y $\beta \leq \beta'$, entonces $u(\alpha,\beta) \leq u(\alpha',\beta')$

También es deseable que cumpla las Leyes de De Morgan para calcular el grado de la unión en función de los grados de la intersección y el complementario. Si a las propiedades de obligado cumplimiento se le suman que siguen estas leyes, se dice que las funciones de pertenencia se les llama conormas triangulares.

En estos casos, la unión puede ser representada mediante las t-conorma. Hay muchas variedades, como las drásticas de la suma o el producto. Sin embargo, las más sencillas y usadas son:

T-conorma máxima: toma el valor máximo de pertenencia en cada punto del dominio de los conjuntos difusos que se están uniendo. Dicho de otra forma, si un elemento tiene un grado de pertenencia mayor en uno de los conjuntos que se están uniendo, su grado de pertenencia en la unión será el mayor de los dos. Esto significa que la T-conorma máxima tiende a dar más peso a las regiones donde al menos uno de los conjuntos tiene un alto grado de pertenencia. Su notación es: $u_{max}(\alpha, \beta) = \max(\alpha, \beta)$

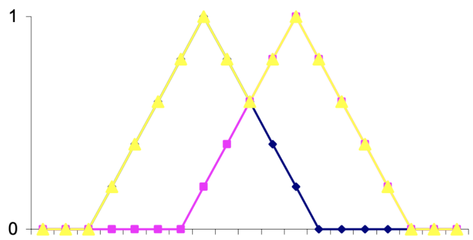

Figure 4.12: t-norma del máximo. En azul y rosa las funciones de pertenencia de los dos conjuntos, y en amarillo la unión de los dos conjuntos difusos.

T-conorma suma: se toma el valor mínimo de pertenencia en cada punto del dominio de los conjuntos difusos que se están uniendo. Es decir, si tienes dos conjuntos difusos y deseas calcular su unión, asignará a cada elemento en el conjunto resultante el valor mínimo de pertenencia entre los dos conjuntos originales en ese punto. Su notación es: $u * (\alpha, \beta) = \alpha + \beta - \alpha * \beta$

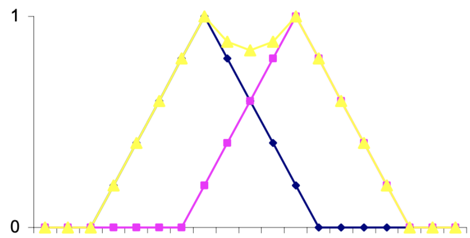

Figure 4.13: t-norma del producto. En azul y rosa las funciones de pertenencia de los dos conjuntos, y en amarillo la intersección de los dos conjuntos difusos.

4.3.2 Intersección

La teoría clásica define la intersección como la pertenencia de un elemento a un conjunto intersección si pertenece a ambos. Su análogo difuso busca determinar el grado de pertenencia a dicho conjunto intersección partiendo de los dos grados de pertenencia de los conjuntos originales. La notación será la siguiente, donde los dos conjuntos a valorar se definen como A y B:

$$\mu_{A \cap B}(x) = \text{inter}(\mu_A(x), \mu_B(x)),$$

donde inter: [0,1] x [0,1] $\longrightarrow [0, 1]$(4.4)

Similar a la unión, hay una serie de propiedades que se deben cumplir, partiendo de $\forall \alpha, \beta, \gamma \in [0, 1]$:

- Concordante con el caso nítido: $\text{inter}(0, 1) = \text{inter}(0, 0) = \text{inter}(1, 0) = 0$; mientras $\text{inter}(1, 1) = 1$
- Propiedad commutativa: $\text{inter}(\alpha, \beta) = \text{inter}(\beta, \alpha)$
- Propiedad asociativa: $\text{inter}(\alpha, \text{inter}(\beta, \gamma)) = \text{inter}(\text{inter}(\alpha, \beta), \gamma)$
- Identidad: $\text{inter}(\alpha, 1) = \alpha$
- Monotonía: si $\alpha \leq \alpha'$ y $\beta \leq \beta'$, entonces $\text{inter}(\alpha, \beta) \leq \text{inter}(\alpha', \beta')$

Cuando en un problema se usan funciones de pertenencia triangulares, la intersección puede ser representada mediante las t-norma. Se usan para calcular la intersección de dos conjuntos difusos, es decir, representa la parte en común o la superposición entre los dos conjuntos originales.

Hay muchas variedades, como las drásticas de la suma o el producto. Sin embargo, las más sencillas y usadas son:

T-norma mínima: se usa para calcular la intersección de conjuntos difusos tomando el valor mínimo de pertenencia en cada punto del dominio. Es decir, si un elemento tiene un grado de pertenencia mayor en un conjunto que en otro, su grado de pertenencia en la intersección será el menor de los dos. Su notación es: inter$_{min}(\alpha, \beta) = \min(\alpha, \beta)$

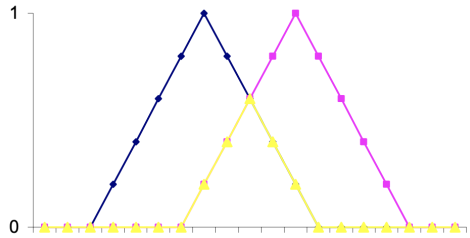

Figure 4.14: t-norma del mínimo. En azul y rosa las funciones de pertenencia de los dos conjuntos, y en amarillo la intersección de los dos conjuntos difusos.

T-norma producto: se usa para calcular la intersección de conjuntos difusos multiplicando los grados de pertenencia en cada punto del dominio. Esto tiende a dar más peso a las regiones donde ambos conjuntos tienen altos grados de pertenencia. Su notación es: inter $* (\alpha, \beta) = \alpha * \beta$

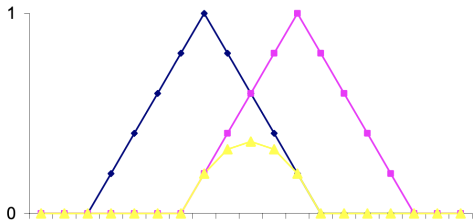

Figure 4.15: t-norma del producto. En azul y rosa las funciones de pertenencia de los dos conjuntos, y en amarillo la intersección de los dos conjuntos difusos.

4.3.3 Complementario

Dado un conjunto A, su complementario $\overline{A}$ (en la literatura hay diferentes formas de representar la notación del complementario) se forma con todos aquellos elementos del universo que no pertenecen al conjunto A. En conjuntos difusos, la representación de si un elemento pertenece o no a un conjunto A viene dado por su función de pertenencia, por lo que la representación del complementario se denota como sigue (4.5):

$$\mu_{\overline{A}}(x) = c(\mu_A(x)) \tag{4.5}$$

A nivel de propiedades, la función complementaria $c(\alpha)$ tiene que cumplir que en el rango de 0 a 1, se cumple que $c : [0, 1] \rightarrow [1, 0]$, siendo una función concordante con el caso nítido ($C(1) = 0$ y $C(0) = 1$), forzosamente decreciente ($\forall \alpha, \beta \in [0,1]\ \alpha > \beta \longrightarrow c(\alpha) < c(\beta)$), y que cumple la anulación de la doble negación ($c(c(\alpha)) = \alpha$).

Resumiendo, la función del complementario se puede representar como $c(\alpha) = 1 - \alpha$ y gráficamente como:

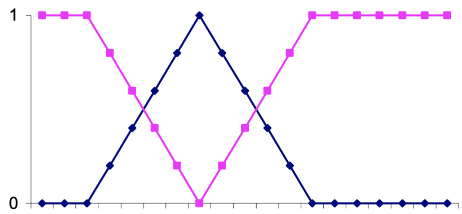

Figure 4.16: Complementario de la función de pertenencia triangular del conjunto difuso. En azul la función de pertenencia original y en rosa su complementario.

Hay otras variantes del complementario, que cumplen todas las propiedades mencionadas, que sirven para suavizar la función complementaria clásica (por ejemplo, complementario de Yager o Sugeno).

4.3.4 Otras diferencias entre conjuntos clásicos y difusos

Hay otros ciertos aspectos que diferencian a las operaciones y propiedades de los conjuntos clásicos de los difusos. Por ejemplo, las condiciones requeridas en la unión e intersección no garantizan que se cumplan propiedades básicas como la idempotencia ($A \cap A = A$) y la distributividad ($A \cap (B \cup C)$) de la intersección, o la idempotencia ($A \cup A = A$) y la distributividad de la unión ($A \cup (B \cap C)$).
Para verificar éstas, se usan la t-normal del mínimo y su t-conorma del máximo correspondiente.

También ha cierta diferencia en la definición del conjunto vacío y el conjunto universal. En concepto clásico del conjunto vacío dice que es aquel conjunto que no contiene ningún elemento. Su correspondiente definición en teoría de conjuntos difusos sería:

$$\forall x \in \chi, \mu_{\emptyset}(x) = 0 \tag{4.6}$$

Por tanto, el conjunto universal se definirá como:

$$\forall x \in \chi, \mu_{\chi}(x) = 1 \tag{4.7}$$

No obstante, si se parte de estas definiciones no se consiguen verificar en la teoría de conjuntos difusos algunos teoremas famosos de la teoría de conjuntos clásica, como pueden ser $A \cap \overline{A} = \emptyset$ o $A \cup \overline{A} = \chi$. A esta contradicción se le llama principio de contradicción y principio del tercio excluso, respectivamente. Por ejemplo, una persona, por lógica difusa, puede pertenecer y no pertenecer al conjunto obeso, con diferentes grados. Por lo tanto, la intersección del conjunto y su complementario no sería vacío.

Una posible solución para que cumplan estos teoremas sería la definición de una t-norma y una t-conorma acotada (también llamadas de Luckasiewicz) que los satisfagan, pero no cumplirían las propiedades de idempotencia y distributividad de la intersección y unión.

4.4 Relaciones y razonamiento difuso

Una vez vista la teoría de conjuntos difusos, y como objetivo final de modelar el razonamiento difuso, hay que prestar atención a las relaciones que construyen el razonamiento difuso.

4.4.1 Relaciones difusas

Las relaciones difusas son una extensión de las relaciones binarias clásicas en las que los elementos de un conjunto pueden estar relacionados de una manera más flexible y gradual en lugar de estar simplemente relacionados o no relacionados. Como ya se ha visto, se pasa de de tener valores binarios (verdadero o falso) a que las relaciones difusas permitan la representación de grados de pertenencia o grado de relación entre elementos.

Existen diferentes relaciones difusas, destacando las más comunes:

Relaciones difusas

Igualdad difusa: a diferencia de afirmar que dos elementos son iguales o no lo son, la igualdad difusa permite representar cuánto se asemejan dos elementos en una escala de 0 a 1. Por ejemplo, si estamos comparando dos elementos difusos, podemos decir que tienen un grado de igualdad de 0.8, lo que indicaría una alta similitud.
Mayor o Menor que (difuso): a diferencia de decir que un número es mayor o menor que otro, las relaciones difusas pueden representar cuánto uno es mayor o menor que el otro en una escala de 0 a 1. Esto es útil en situaciones donde se necesita expresar relaciones de orden difusas.
Conjunto relacional difuso: no solo trata de establecer relaciones directas entre elementos, sino que se utilizan conjuntos difusos para describir las relaciones. Por ejemplo, se puede tener un conjunto difuso que representa la relación "es un buen compañero de trabajo de". Cada elemento en el conjunto tendría un grado de pertenencia que indica cuánto es un buen compañero de otros elementos.
Relaciones temporales difusas: describen cómo evolucionan las relaciones a lo largo del tiempo en lugar de ser fijas. Por ejemplo, se podría tener una relación difusa que representa "es cercano a" en un contexto de amistad, y esta relación podría cambiar con el tiempo.
Relaciones espaciales difusas: describen relaciones en el espacio. Por ejemplo, en un mapa, podríamos tener una relación difusa que representa "proximidad" entre ubicaciones en lugar de una relación binaria de "está cerca" o "no está cerca".

4.4.2 Razonamiento difuso

El razonamiento difuso es el proceso de realizar inferencias (se ve en el siguiente apartado) partiendo de hechos, relaciones imprecisas (vagas), evidencias difusas y su precisión actualizada mediante creencias.

Lo primero es definir dos tipos de proposiciones:

Proposiciones difusas

Simple: es aquella que da un valor a una variable difusa, asociando a un conjunto difuso su valor lingüistico y su correspondiente función de pertenencia. Por ejemplo, "la altura de Lebron es alta"
Compuesta: es aquella que agrupa dos o más proposiciones difusas simples, o incluso podrían ser modificadas por su agrupación. Para ello, se usan los operadores lógicos (se ven a continuación). Un ejemplo compuesto sería "La persona es alta" Y "La persona trabaja muy cerca".

Los **operadores lógicos difusos** sirven para agrupar (conectivos) o para modificar (negación) proposiciones. Definiéndolos de una forma más formal, partiendo de dos proposiciones difusas p y q, dos conjuntos difusos A y B que intervienen en ellas con funciones de pertenencia μ_A y μ_B, los operadores tienen la siguiente notación:

- NO (¬p): $\mu_A(x) = 1 - \mu_A(x)$
- Y ($p \wedge q$): $\mu_{A \wedge B}(x, y) = \min(\mu_A(x), \mu_B(y))$
- O ($p \vee q$): $\mu_{A \vee B}(x, y) = \max(\mu_A(x), \mu_B(y))$

4.5 Implicación e inferencia difusa

Con toda la información anterior, el siguiente paso para el modelado de un problema que se desea resolver con razonamiento basado en reglas es la definición de unas reglas difusas del tipo SI-ENTONCES (SI "Persona es muy alta" ENTONCES "juega mucho al baloncesto").
Antes de profundizar en las reglas y la inferencia difusa, hay que definir un concepto previo necesario llamado "implicación".

4.5.1 Implicación de Mamdani

Definiciones

Implicación: usado en lógica y razonamiento para expresar la relación entre dos proposiciones o afirmaciones, generalmente en la forma "SI... ENTONCES...". Establece una conexión lógica o causal entre dos proposiciones, donde la primera, llamada "antecedente", es una condición o premisa que, si se cumple, lleva a la segunda, llamada "consecuente", como resultado". La notación tradicional es: $p \longrightarrow q$

Entonces, una implicación en lógica difusa tiene que asignar una función de pertenencia a una agrupación del antecedente (p) y el consecuente (q) que permita razonar con reglas del tipo SI-ENTONCES.

La función de pertenencia se define como:

$$\begin{aligned} \mu_{p\to q} :& U \times V \to [0,1] \\ & (x,y) \to \mu_{p\to q}(x,y) \end{aligned} \tag{4.8}$$

El asunto central de la relación de implicación reside en decidir qué se desea representar, puesto que es la base para el razonamiento basado en reglas. Hay dos tipos diferentes:

1. Implicación lógica: hacer coincidir el significado de la implicación con el de la lógica clásica. Un ejemplo sería la equivalencia $p \to q \equiv \neg p \vee q$. Como se explicó en el apartado de los operadores lógicos difusos, se reformula la función de pertenencia como $\mu_{p\to q}(x,y) = \max(1 - \mu_A(x), \mu_B(y))$.
2. Implicación causa-efecto: dar un significado de relación causa-efecto como se usan en los Sistemas basados en conocimiento (SBC).

La implicación lógica difusa es la mejor a nivel teórico, pero falla en la formulación en muchos SBC con relaciones causa-efecto no consistentes con la lógica. Por ello, se deben definir el segundo tipo. La forma más sencilla y común es mediante la implicación de Mamdani:

Definiciones

Implicación de Mamdani: el grado de verdad de la regla $p \to q$ es igual a la proposicion A y B. Es decir, una condición solo es verdad cuando el antecente y el consecuente son ciertos. Su notación es la siguiente: $p \to q \equiv p \wedge q \Rightarrow \mu_{p\to q}(x,y) = \min(\mu_A(x), \mu_B(y))$

Por tanto, la implicación de Mamdani se basa en dos pasos:
Evaluación de reglas: cada regla en el sistema de inferencia difusa se evalúa en función de las entradas y se activa con un grado de activación difuso que refleja la adecuación de esa regla para la situación actual.
Combinación de reglas: una vez que todas las reglas se han evaluado y activado, los resultados se combinan para obtener una salida difusa. La implicación de Mamdani utiliza operadores de lógica difusa (como la T-conorma máxima para la unión y la T-norma mínima para la intersección) para combinar los grados de activación de las reglas y producir la salida.

Por último, aclarar que en implicación difusa, las reglas de la lógica clásica varían un poco. Las reglas difusas pueden tomar diferentes formas (con antecedentes difusos, sin ellos, etc.). En este caso, vamos a explicar brevemente dos de ellos:

- Con antecedentes difusos: cuando en el antecedente hay un hecho difuso, del cual se pretende saber su función de pertenencia.
- Con antecedentes nítidos: cuando en el antecedente hay un hecho nítido y se desea saber la función de pertenencia del hecho difuso de la conclusión.

Por ejemplo, la regla de Modus Ponens se generaliza. En lógica clásica se formula de la siguiente forma:

- Proposición 1: "x es A"
- Proposición 2: "SI x es A ENTONCES y es B"
- Consecuencia: "y es B"

siendo generalizada en difusa:

- Premisa 1: "x es A*"
- Premisa 2: "SI x es A ENTONCES y es B"
- Consecuencia: "y es B*"

donde el conjunto difuso A* no tiene por qué ser el mismo que el conjunto difuso A del antecendente de la premisa 2, ni el conjunto difuso B* tampoco tiene que ser el mismo que el que sale en el consecuente de la premisa 2.

En conclusión, en la lógica clásica la regla solo funciona si la premisa 1 coincide exactamente con el antecedente de la premisa 2, y en la difusa funciona si hay un grado de similitud diferente a cero entre ambas.

4.6 Defuzzificación

Al igual que en muchos sistemas tenemos una fase de codificación y otra de decodificación, en lógica borrosa existe algo similar.
La asociación de un cojunto difuso con sus valores de pertenencia, que ya se han explicado a lo largo de este capítulo, se le llama **fuzzificación** (el equivalente al *enconder*). Con los valores fuzzificados ya se puede trabajar con las reglas lingüisticas para lograr una salida. Ésta salida puede ser de dos tipos: difusa o **defuzzificada** (el equivalente al *decoder*). En otras palabras, la salida será un valor difuso (llamado *fuzzy*) o discreto (también llamado *crisp*).

Un ejemplo de defuzzificación sería la altura de una persona, que quedaría de la siguiente forma:

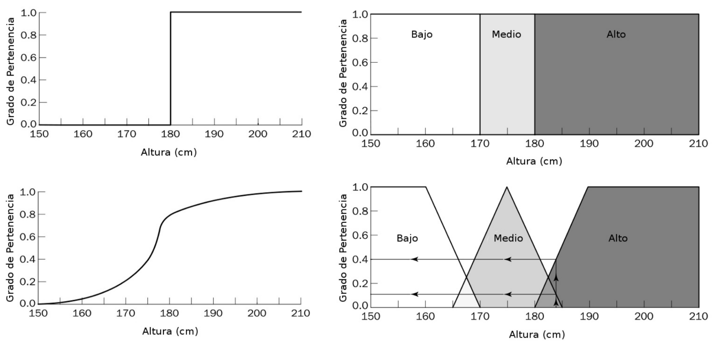

Figure 4.17: Defuzzificación. Ejemplo de equivalencia entre valores fuzzy (abajo) y crisp (arriba), y sus funciones de pertenencia.

En el ejemplo de la figura, una persona de 181 cm es catalogado como "alto" en valores discretos. Sin embargo, el conjunto difuso permite expresar que esa persona tiene un grado de pertenencia del 82% al conjunto de los altos ($\mu_{Alto}(180) = 0.82$).

Existen decenas de métodos para defuzzificar, siendo el más común el del centroide. A continuación se formulan los más usados:

Centroide o centro de área.
Se calcula el valor de salida X_C donde una línea vertical divide el conjunto en dos áreas con igual masa (ese punto lo marca el centro de gravedad) con la fórmula siguiente:

$$X_C = \frac{\sum_{x=a}^{b} \mu_A(x)x}{\sum_{x=a}^{b} \mu_A(x)} \tag{4.9}$$

donde a y b son los límites del dominio del conjunto difuso A.
Cabe recordar que esta formula es para los casos en los que la función de pertenencia es discreta. Para el caso continuo, habría que integrar en lugar de sumar.

Este método es muy "pesado" computacionalmente. Por ello, se usan otros métodos con menor carga.

Criterio de máximo (MC).
Se calcula el valor de salida X_{max} mediante el valor máximo que alcanza la función de pertenencia.

$$X_{max} = \arg\max_x[\mu_A(x)] \tag{4.10}$$

Máximo central (o media del máximo, MOM).
Se calcula el punto C tomando los valores en los que las funciones de pertenencia alcanzan el máximo, calculando la media con la fórmula siguiente:

$$C = \sum_{i=1}^{n} \frac{z_i}{n} \tag{4.11}$$

donde z_i es el punto en el que la función de pertenencia logra el máximo valor, y n es el número de veces que se alcanza el valor máximo.

Máximo menor (SOM).
Se calcula el valor de salida X_{SOM} tomando el valor más pequeño de todos aquellos que logran alcanzar el valor más alto de la función de pertenencia.

$$X_{SOM} = \min(\arg\max[\mu_A(x)], \arg\max[\mu_B(x)], \ldots]) \tag{4.12}$$

Máximo mayor (LOM).
Similar al anterior, pero en este caso, el valor de salida X_{LOM} se calcula tomando el valor mayor.

$$X_{LOM} = \max(\arg\max[\mu_A(x)], \arg\max[\mu_B(x)], \ldots]) \tag{4.13}$$

Bisectriz.
Se calcula el valor de salida X_{bis} como el valor que separa el área bajo la curva en dos sub-áreas iguales.

$$X_{bis} = \frac{X_{min} + X_{max}}{2} \tag{4.14}$$

donde X_{min} y X_{max} son los valores más bajo y alto que da la función de pertenencia (sin contar el cero).

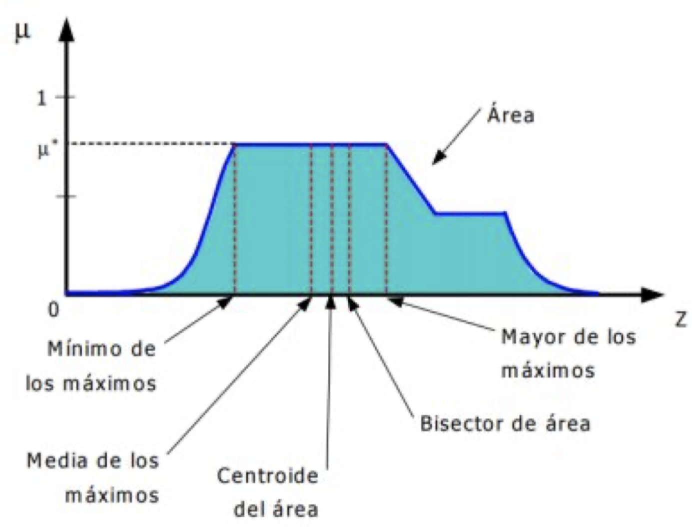

Figure 4.18: Esquema con los diferentes método de defuzzificación.

4.7 Modelado difuso

El último paso para lograr un sistema de razonamiento que se base en técnicas de lógica difusa es modelar un sistema de inferencia difuso.

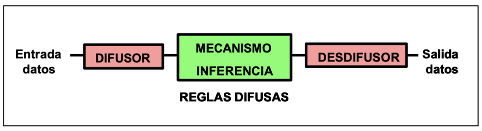

Figure 4.19: Esquema general de un sistema basado en lógica difusa.

Existen diferentes mecanismos de inferencia, como puede ser el del mínimo de Mamdani el del producto de Larsen, el TSK, etc. En este apartado se van a ver dos tipos: con defuzzificador y sin él.

4.7.1 Modelo de Mamdani

Es el modelo más conocido, y se basa en 3 fases principales: un difusor y fuzzyficador, un mecanismo de inferencia (se puede dividir en dos fases diferenciadas, evaluación de reglas y agregación de sus salidas) y un desdifusor o defuzzificador. A continuación se describe con más detalle, poniendo un ejemplo sencillo para una mayor comprensión:

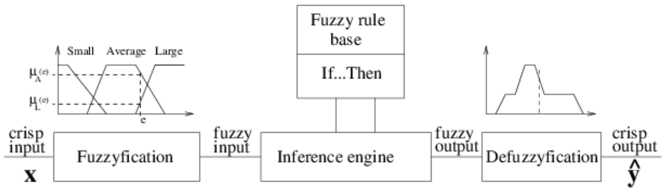

Figure 4.20: Esquema del modelo Mamdani.

Se desea evaluar el riesgo de colesterol de una persona con sus datos de altura y peso. Se dispone de tres reglas que van a hacer uso de las siguientes variables lingüisticas:

- Altura (x). El conjunto se define sobre el dominio $x_{bajo}, x_{mediano}, x_{alto}$.
- Peso (y). El conjunto se define sobre el dominio $y_{delgado}, y_{normal}, y_{obeso}$.
- Riesgo (z). El conjunto se define sobre el dominio $z_{pequeo}, z_{medio}, z_{elevado}$.

Y las tres reglas serían las siguientes:

1. SI x es x_{alto} O y es $y_{delgado}$ ENTONCES z es z_{pequeo}
2. SI x es $x_{mediano}$ O y es y_{normal} ENTONCES z es z_{medio}
3. SI x es x_{bajo} O y es y_{obeso} ENTONCES z es $z_{elevado}$

Por tanto, se parte de unos valores discretos (crisp), y unas reglas, que será el punto de partida del modelo, donde la primera etapa de inferencia es la fuzzificación.

Sistema de Fuzzificación. En este paso, las entradas del sistema definidas en valores lingüísticos (variables difusas) se transforman en valores difusos. Esto significa que las entradas, que pueden ser medidas o datos precisos, se transforman en conjuntos difusos que representan la incertidumbre o la vaguedad asociada con esas entradas. Para ello, se usan las funciones de pertenencia almacenadas en la estructura de conocimiento. En este caso se ponen los siguientes valores de pertenencia: x (160 y 180 cm) e y (60 y 100 kg). De esta forma se fuzzifican todas las entradas con las funciones de pertenencia μ_x e μ_y. Por ejemplo, una persona con 161 cm de altura y un peso de 67 kg pertenecerá a los conjuntos difusos $x_{mediano}$ con un grado de pertenencia de 0.2 y al x_{bajo} con un 0.5, y a los conjuntos difusos $y_{delgado}$ con un grado de 0.1 y al y_{normal} con un 0.7.

Evaluación y combinación de reglas. Se utilizan reglas de inferencia difusa que establecen relaciones entre las entradas anteriores y las salidas. Esas entradas se aplican a los antecedentes de las reglas difusas y obtener el consecuente.
Si una regla tiene más de un antecedente, se utilizan los operadores para obtener un solo valor. En ese caso, para la evaluación se usan la t-norma del mínimo (Y) o la t-conorma máxima (O).

Finalmente, el resultado del antecedente se aplica al consecuente, haciendo uso del recorte o el escalado según el valor obtenido en el antecedente. El recorte acota el valor del consecuente con el valor de verdad del antecedente, y el escalado multiplica todos los valores por el valor de verdad del antecedente para dar un valor más exacto y presertvando la forma original del conjunto difuso.

Teniendo todas las reglas evaluadas y los conjuntos difusos de salida modificados, se combinan todas las reglas para tener un resultado único. Esto se hace normalmente con operadores de lógica difusa, como la T-conorma máxima (para la unión) y la T-norma mínima (para la intersección).

Sistema de Defuzzificación. Como el resultado, normalmente, hay que representarlo mediante un valor discreto, se toma como entrada en este sistema el conunto difuso obtenido de la etapa anterior y se "decodifica". Para ello, hay que usar un método de defuzzificación de los mencionados anteriormente.

Un esquema general del ejemplo dado dos valores de entrada se puede ver en el esquema siguiente:

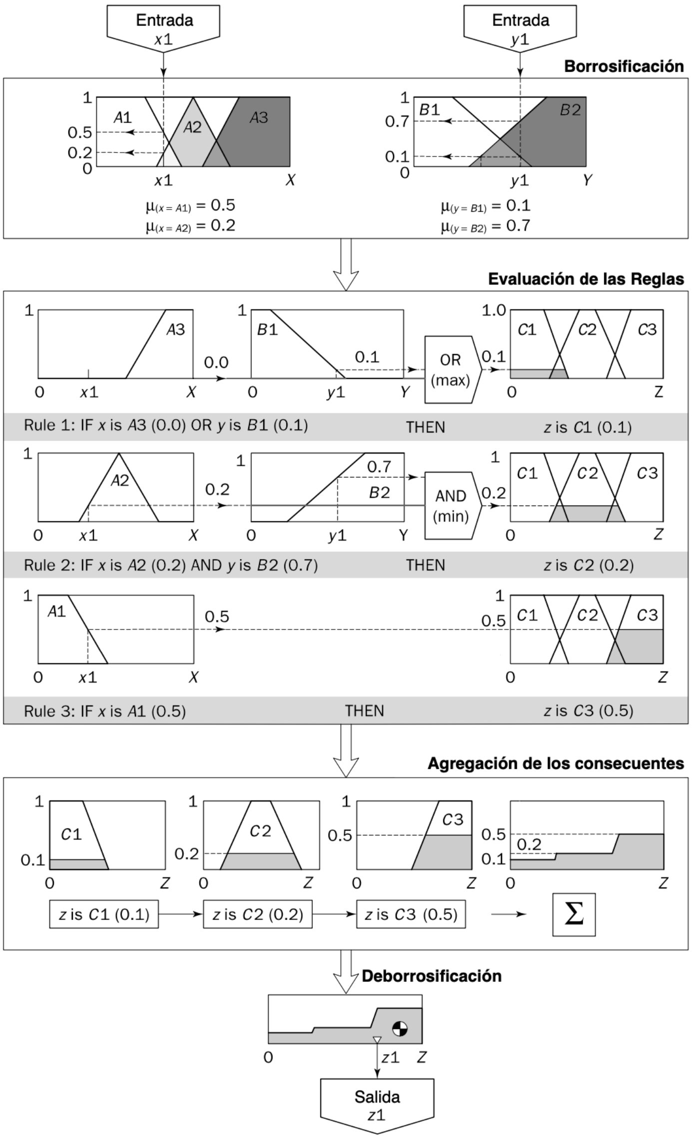

Figure 4.21: Esquema del ejemplo mediante modelo Mamdani. Nota: en lugar de x, y y z se usan A, B y C.

4.7.2 Modelo TSK

El modelo de inferencia de Mamdani hace uso de algún método para la defuzzificación. Por regla general, este método no es muy eficiente desde el punto de vista computacional. Por ello, es posible reducir este tiempo mediante una función matemática en el consecuente.

Uno de los modelos que siguen este enfoque es el modelo TSK, tiene el formato general de regla de inferencia siguiente:

p: SI x es A Y y es B ENTONCES z es $f(x, y)$

De esta forma es mucho más eficiente, pero no es tan natural su representación. Por ello, se usa una aproximación basada en un *singleton* (un solo valor), que consiste en dar un valor 1 en un punto concreto y 0 en el resto del universo. De esta forma, la regla quedaría de la siguiente forma:

p: SI x es A Y y es B ENTONCES z es k

siendo k un valor constante (singleton).

Con esta aproximación, el ejemplo anterior quedaría de la siguiente forma:

Por último, insistir que hay muchos más modelos de inferencia, adecuados cada uno a un problema concreto. Eso sí, tienen una serie de objetivos comunes (aproximadores generales, transparencia e interpretabilidad), y para ello cumplen con unas características concretas: **distinguibilidad** (que los valores lingïsticos sean lo más significativos y diferenciables posible), **normalidad** (que para cada función de pertenencia exista al menos un elemento del universo con pertenencia máxima), **moderación** (número acotado de entidadaes a manejar) y **cubrimiento** (que cubra todo el espacio de la variable representada).

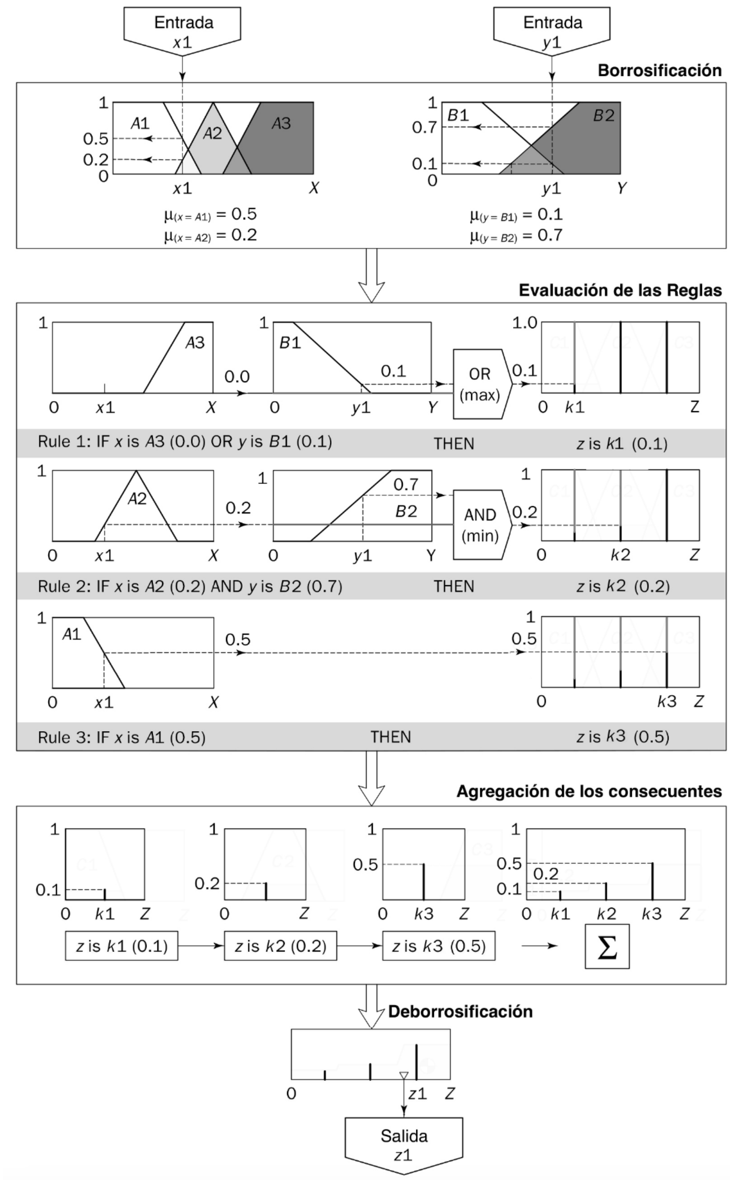

Figure 4.22: Esquema del ejemplo mediante modelo TSK. Nota: en lugar de x, y y z se usan A, B y C.