Tema 3: Lógica Difusa

Universidad Pontificia de Salamanca

Manuel Martín-Merino

Contenido

Introducción (I)

- Los expertos humanos controlan procesos muy complejos utilizando reglas heurísticas y sin necesidad de recurrir a complejos modelos matemáticos.

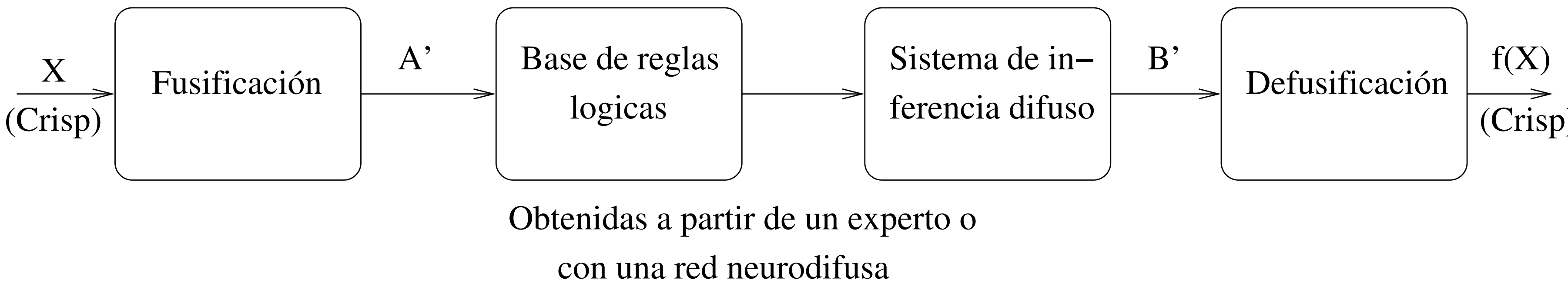
- Un experto controlaría la presión sobre el freno de un tren utilizando reglas del siguiente tipo:

  Si la VELOCIDAD es *baja* y la DISTANCIA es *intermedia* la PRESIÓN sobre el freno es *baja*

- Las reglas empleadas por el experto utilizan variables difusas como VELOCIDAD de valores difusos {baja, media, alta }. Por eso no pueden ser modeladas por la lógica clásica.

Introducción (II)

- Sin embargo las variables de entrada salida del controlador deben ser numéricas (crisp). Por tanto, el proceso de diseño de un sistema difuso consta de las siguientes fases:



- El sistema resultante se puede interpretar en forma de reglas comprensibles por el experto humano.

- Es posible incorporar conocimiento heurístico a priori en el sistema.

Conjuntos difusos (I)

- En la lógica clásica una proposición o es verdadera o es falsa. De igual forma, en un conjunto clásico, un elemento o pertenece a un conjunto o no pertenece.

- Los elementos de un conjunto difuso pueden tener diferentes grados de pertenencia a un conjunto.
  Sea X una colección de objetos. Se define el conjunto difuso A en X como:

$$A = \{(x, \mu_A(x)) | x \in X\} \tag{1}$$

  $\mu_A(x)$ indica el grado de pertenencia de x al conjunto A y es no negativa y acotada.

Conjuntos difusos (II)

- Las funciones de pertenencia no son probabilidades.

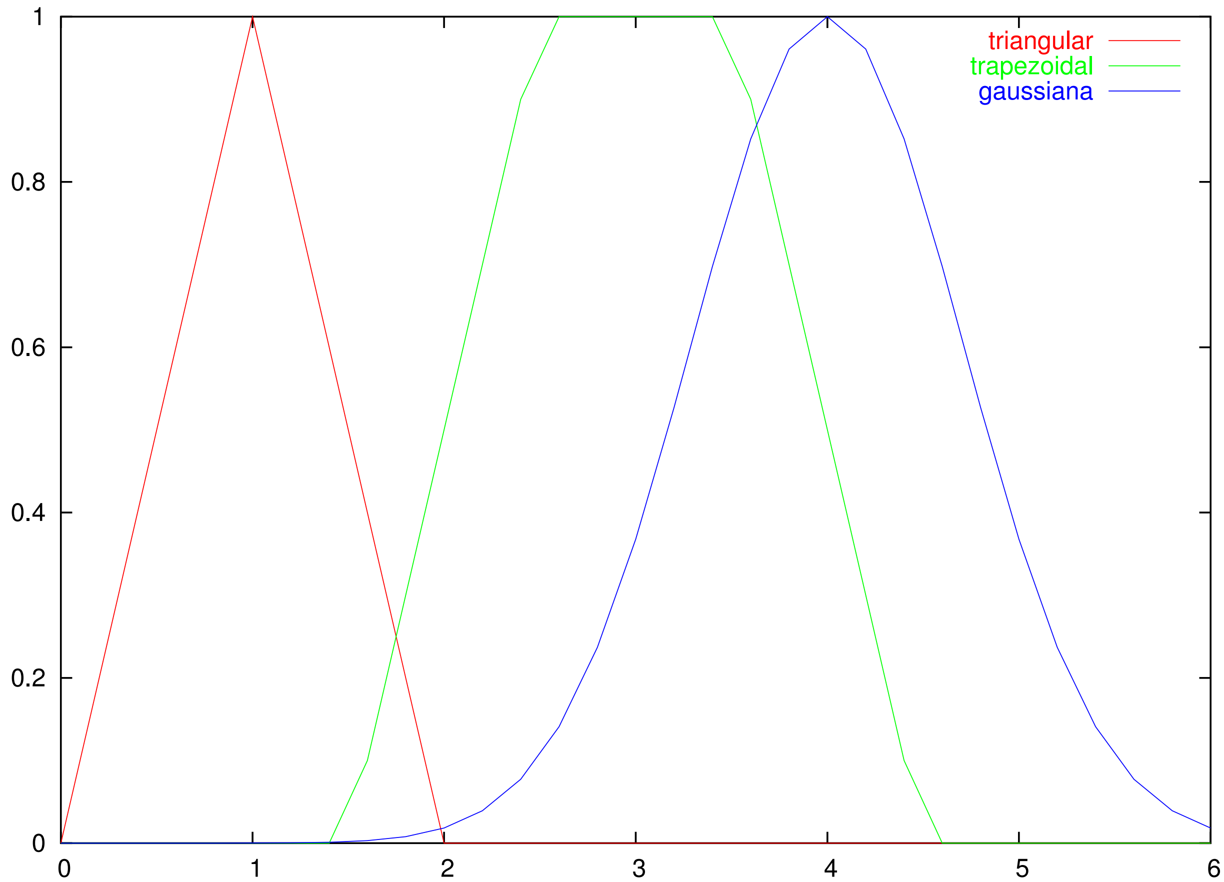
  Ejemplo:

  $p(x \in A) = 0{,}25$ indica que si repetimos un experimento 100 veces en 25 ocasiones pertenece a A.

  $\mu_A(x) = 0{,}25$ indica que x pertenece a la frontera imprecisa del conjunto con grado $0{,}25$.

Conjuntos difusos (III)

Funciones de pertenencia usuales



La función de pertenencia más utilizada es la triangular por ser más sencillo trabajar con ella.

Conjuntos difusos (IV)

Matemáticamente se definen como:

- Triangular:

$$\mu_A(x) = \begin{cases} \frac{x-0}{1-0} & \text{si} \quad 0 \leq x \leq 1 \\ \frac{2-x}{2-1} & \text{si} \quad 1 \leq x \leq 2 \end{cases} \tag{2}$$

- Trapezoidal:

$$\mu_A(x) = \begin{cases} \frac{x-1{,}5}{2{,}5-1{,}5} & \text{si} \quad 1{,}5 \leq x \leq 2{,}5 \\ 1 & \text{si} \quad 2{,}5 \leq x \leq 3{,}5 \\ \frac{4{,}5-x}{4{,}5-3{,}5} & \text{si} \quad 3{,}5 \leq x \leq 4{,}5 \end{cases} \tag{3}$$

- Gaussiana:

$$\mu_A(x) = \exp(-(x-4)^2/1) \tag{4}$$

Operaciones con conjuntos difusos (I)

Una vez que sabemos modelar conjuntos difusos es necesario definir las conectivas que permitirán combinarlos para formar expresiones más complejas. A diferencia de la lógica clásica las definiciones no son únicas.

Definición (Intersección difusa, t-normas): Sea $t: A \times B \rightarrow [0,1]$ una función definida como $t(\mu_A(x), \mu_B(x)) = \mu_{A \wedge B}(x)$. Para que t sea una $t-$norma debe cumplir las siguientes propiedades:

- $t(0,0) = 0$; $t(a,1) = a$ (condición de frontera).
- $t(a,b) = t(b,a)$ (commutativa).
- Si $a \leq a'$ y $b \leq b'$ entonces $t(a,b) \leq t(a',b')$ (no decreciente).
- $t[t(a,b), c] = t[a, t(b,c)]$ (asociativa)

El producto y el mínimo cumplen las condiciones anteriores.

Operaciones con conjuntos difusos (II)

Definición (Unión difusa, s-conormas): Sea $s: A \times B \to [0,1]$ una función definida como $s(\mu_A(x), \mu_B(x)) = \mu_{A \wedge B}(x)$. Para que s sea una $s-$norma debe cumplir las siguientes propiedades:

- $s(1,1) = 1$; $s(a,0) = a$ (condición de frontera).
- $s(a,b) = s(b,a)$ (commutativa).
- Si $a \leq a'$ y $b \leq b'$ entonces $s(a,b) \leq s(a',b')$ (no decreciente).
- $s[s(a,b),c] = s[a,s(b,c)]$ (asociativa)

El máximo y la suma algebraica cumplen las condiciones anteriores.

Operaciones con conjuntos difusos (III)

Los operadores difusos más utilizados se resumen en la siguiente tabla.

Operación	Def. función pertenencia
Intersección (and)	$\mu_{A \wedge B}(x) = \min(\mu_A(x), \mu_B(x))$ $\mu_{A \wedge B}(x) = \mu_A(x)\mu_B(x)$
Unión (or)	$\mu_{A \cup B}(x) = \max\{\mu_A(x), \mu_B(x)\}$ $\mu_{A \cup B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x)$
Complementación (not)	$\mu_{\neg A}(x) = 1 - \mu_A(x)$

Operaciones con conjuntos difusos (IV)

Finalmente se suele definir la función de pertenencia para problemas multivariantes como el producto tensorial de funciones univariantes:

$$\mu_{A_1} \otimes, \ldots, \otimes \mu_{A_p} \tag{5}$$

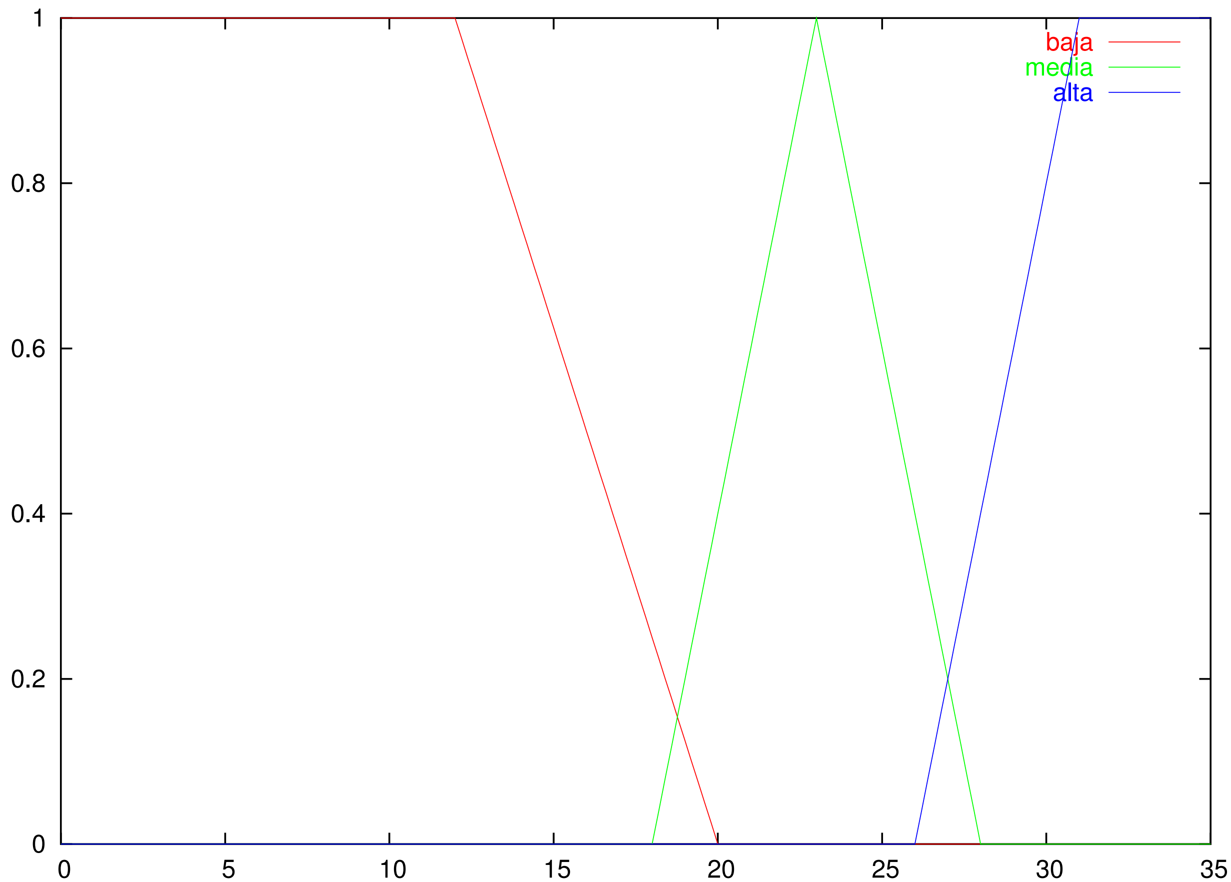
Así, la función de pertenencia sobre un dato $\boldsymbol{x}$ vale:

$$\mu_A(\boldsymbol{x}) = \mu_{A_1}(x_1) \ldots \mu_{A_p}(x_p) \tag{6}$$

Operaciones con conjuntos difusos (V)

Ejemplo:

Sea la variable temperatura de valores lingüísticos $\{$Baja, Media, Alta $\}$ y cuyas funciones de pertenencia se muestran en la figura. Obtener las ecuaciones que modelan dichos conjuntos difusos. Obtener las funciones de pertenencia de los conjuntos difusos *Baja y Media, Media o Alta* y no Media y Alta



Relaciones difusas (I)

Definición: Una relación clásica en $U \times V$ es un subconjunto $Q(U \times V) \subset U \times U$ contenido el el espacio producto cartesiano.

Definición: Una relación difusa en $U \times V$ es un conjunto difuso definido como:

$$Q = \{((x_1, x_2), \mu_Q(x1, x2)) | (x_1, x_2) \in U \times V\} \tag{7}$$

Ejemplos:

- “x aproximadamente igual a y”
  $\mu_Q(x, y) = e^{-(x-y)^2}$
- “x es mucho mayor que y”
  $\mu_Q(x, y) = \frac{1}{1+e^{-(x-y)}}$

Relaciones difusas (II)

Composición de relaciones difusas

Sean $P(U,V)$ y $Q(V,W)$ dos relaciones crisp que comparten V. $P \circ Q$ es una relación en $U \times W$ tal que $(x,z) \in P \circ Q$ si y sólo si existe al menos un $y \in V$ tal que $(x,y) \in P$ y $(y,z) \in Q$. Esta definición se puede extender al caso difuso.

Lema 1 (Composición relaciones difusas)**.** $P \circ Q$ *es la composición de las relaciones difusas* $P(U,V)$ *y* $Q(V,W)$ *si y sólo si*

$$\mu_{P\circ Q}(x,z) = \max_{y \in V} t[\mu_P(x,y), \mu_Q(y,z)] \tag{8}$$

Proof. Se deja como ejercicio. $\square$

Reglas difusas (I)

En la lógica clásica las reglas *IF* p *THEN* q se expresan como: $p \to q$. Esta expresión es equivalente a:

$$\bar{p} \cup q \tag{9}$$

$$(p \wedge q) \cup \bar{p} \tag{10}$$

Sea una regla de la forma *IF* x es A *THEN* y es B donde A y B son relaciones difusas en $U = U_1 \times \ldots \times U_n$, $V = U_1 \times \ldots \times U_n$ respectivamente.
La implicación difusa es una relación difusa definida en el espacio producto cartesiano $U \times V$.

Reglas difusas (II)

- Implicación *Dienes-Rescher*: Sustituyendo en la ecuación (9) la conectiva or por el $\max$ y la negación por su definición obtenemos:

$$\mu_{A\to B}(x,y) = \max(1-\mu_A(x), \mu_B(y)) \tag{11}$$

- Implicación de *Zadeh*: Sustituyendo la conectiva or por el $\max$ y la and por el $\min$ obtenemos la siguiente definición:

$$\mu_{A\to B}(x,y) = \max[\min(\mu_A(x), \mu_B(y)), 1-\mu_A(x)] \tag{12}$$

Las reglas anteriores están definidas globalmente. Sin embargo las implicaciones utilizadas por el experto tienen grado de activación elevado sólo cuando lo tienen p y q.

Reglas difusas (III)

Por tanto, muchos expertos interpretan la implicación como: $p \to q = p \wedge q$.

- Implicación de Mamdani: Sustituyendo la conectiva and por el máximo o el producto obtenemos:

$$\mu_{A\to B}(x,y) = \text{mín}(\mu_A(x), \mu_B(y)) \tag{13}$$

o

$$\mu_{A\to B}(x,y) = \mu_A(x), \mu_B(y) \tag{14}$$

Esta es la implicación más utilizada en sistemas difusos de control.

Reglas difusas (IV)

Ejemplo: Sea x_1 la velocidad de un coche, x_2 la aceleración e y la fuerza aplicada al acelerador. Sea la siguiente regla difusa:

If x_1 es baja y x_2 es pequeña *THEN* y es alta

donde las funciones de pertenencia de los conjuntos difusos vienen dadas por las siguientes ecuaciones:

$$\mu_{lento}(x_1) = \begin{cases} 1 & \text{si} \quad x_1 \leq 35 \\ \frac{55-x_1}{20} & \text{si} \quad 35 < x_1 \leq 55 \\ 0 & \text{si} \quad x_1 > 55 \end{cases} \qquad \mu_{peq.}(x_2) = \begin{cases} \frac{10-x_2}{10} & \text{si} \quad 0 \leq x_2 \leq 10 \\ 0 & \text{si} \quad x_2 > 10 \end{cases}$$

$$\mu_{grande}(y) = \begin{cases} 0 & \text{si} \quad y \leq 1 \\ y-1 & \text{si} \quad 1 \leq y \leq 2 \\ 1 & \text{si} \quad y > 2 \end{cases}$$

Considérese que los dominios de x_1, x_2 e y son $U_1 = [0, 100]$, $U_2 = [0, 30]$, and $V = [0, 3]$, respectivamente.

Regla Modus Ponens (I)

Premisa 1: x es A'

Premisa 2: *IF* x es A *THEN* y es B

Conclusión: y es B'.

Cuanto más cerca esté A' de A más cerca estará B' de B.

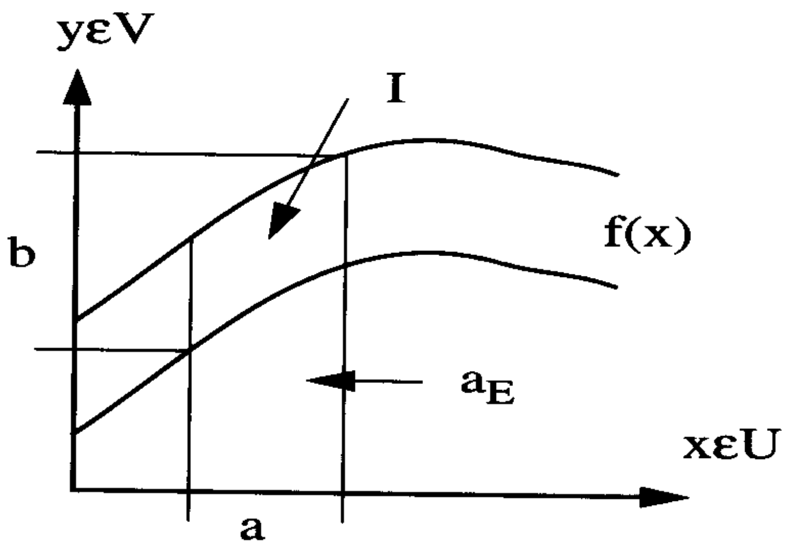


Fig. 1: Inferencia de b a partir del intervalo a y la función $f(x)$.

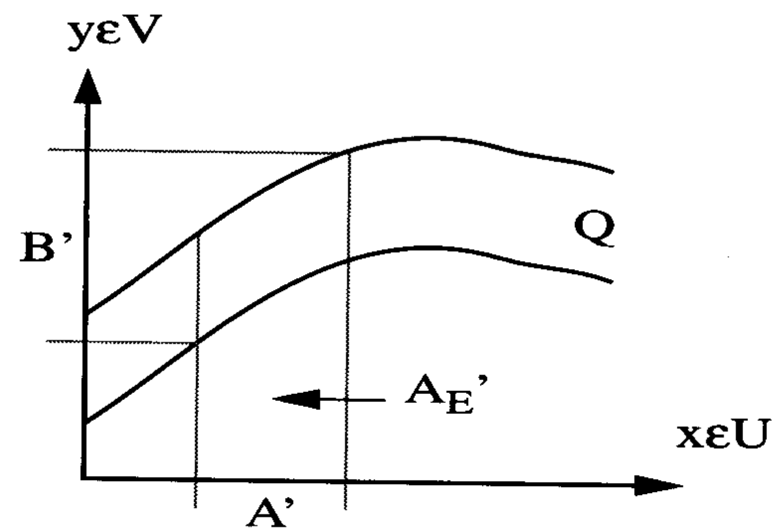


Fig. 2: Inferencia de B' a partir del conj. difuso A' y la relación difusa Q.

La función de pertenencia $\mu_{B'}$ se puede obtener utilizando la regla de composición que motivamos a continuación.

Regla Modus Ponens (II)

Pasos en la obtención de $\mu_{B'}$ vía la relación difusa $\mu_{A \to B}$.

- Extensión cilíndrica del conjunto difuso A'.

$$\mu_{A'_E}(x, y) = \mu_{A'}(x)$$

- Intersección de las relaciones difusas A'_E y $Q \equiv A \to B$:

$$\mu_{A'_E \wedge Q} = t[\mu_{A'_E}(x, y), \mu_Q(x, y)] = t[\mu_{A'}(x), \mu_Q(x, y)]$$

- Realizar la proyección de $A'_E \wedge Q$ sobre V

$$\mu_{B'}(y) = sup_{x \in U} t[\mu_{A'}(x), \mu_Q(x, y)]$$

Sistemas de inferencia difusos (I)

Basados en la agregación de reglas

Sea $R^{(l)} : A_1^l \times \ldots \times A_n^l \rightarrow B^l$ la l-ésima regla difusa definida en $U \times V$. Las M reglas difusas se pueden combinar en una relación difusa Q_M definida como:

$$Q_M = \bigcup_{l=1}^{M} R^{(l)} \tag{15}$$

Esta es la llamada combinación de Mamdani. Utilizando como operador de unión $\dot{+}$, la relación difusa se expresa como:

$$\mu_{Q_M}(x, y) = \mu_{R^{(1)}}(x, y) \dot{+} \ldots \dot{+} \mu_{R^{(M)}}(x, y) \tag{16}$$

Las reglas se pueden combinar utilizando también el operador de intersección:

$$Q_G = \bigcap_{l=1}^{M} R^{(l)} \tag{17}$$

Esta se conoce como combinacón de Gödel, pero su significado es menos intuitivo.

Agregadas las reglas difusas se puede obtener el consecuente B' para una base de M reglas utilizando la regla Modus Ponens:

$$\mu_{B'}(y) = \sup_{x \in U} t[\mu_{A'}(x), \mu_{Q_M}(x,y)] \tag{18}$$

Sistemas de inferencia difusos (II)

Algoritmo:

- Para las M reglas difusas obtener las funciones de pertenencia de los antecedentes de las reglas $\mu_{A_1^l \times \ldots A_n^l}(x_1, \ldots, x_n)$ para $l = 1, 2, \ldots, M$.
- Obtener la función de petenencia $\mu_{R^{(l)}}$ para cada regla $A^l \rightarrow B^l$.
- Determinar $\mu_{Q_M}(x, y)$ o $\mu_{Q_G}(x, y)$ utilizando la combinación de Mamdani o Gödel.
- Para un conjunto de entrada A' obtener el consecuente utilizando la regla Modus Ponens (18).

Sistemas de inferencia difusos (III)

Agregando los consecuentes de las reglas

Para cada regla individual se obtiene el consecuente de A', B'_l utilizando la regla modus ponens:

$$\mu_{B'_l}(y) = \sup_{x \in U} t[\mu_{A'}(x), \mu_{R^{(l)}}(x, y)] \quad \forall l = 1, 2, \ldots, M \tag{19}$$

Los conjuntos difusos de cada regla B'_l se agregan ahora mediante su unión:

$$\mu'_B(y) = \mu_{B'_1}(y) \dot{+} \ldots \dot{+} \mu_{B'_M}(y) \tag{20}$$

o utilizando el operador de intersección:

$$\mu'_B(y) = \mu_{B'_1}(y) * \ldots * \mu_{B'_M}(y) \tag{21}$$

Sistemas de inferencia difusos (IV)

- Sistema de inferencia producto: Se considera: Agregación consecuentes de reglas con operador unión, implicación de Mamdani producto, operador t-norma producto, y $\max$ para las conectivas de unión. Así obtenemos:

$$\mu_{B'}(y) = \max_{l=1}^{M}[\sup_{x \in U}(\mu_{A'}(x) \prod_{i=1}^{n} \mu_{A_i^l}(x_i)\mu_{B^l}(y))] \tag{22}$$

- Sistema de inferencia mínimo: Se considera: Agregación consecuentes de reglas con operador de unión, implicación de Mamdani $\min$, operador t-norma $\min$ y $\max$ para los operadores de unión. Así obtenemos:

$$\mu_{B'}(y) = \max_{l=1}^{M}[\sup_{x\in U}(\min(\mu_{A'}(x), \mu_{A_1^l}(x_1), \ldots, \mu_{A_n^l}(x_n), \mu_{B^l}(y))] \tag{23}$$

Los sistemas anteriores son los más utilizados en sistemas de control difuso. Son computacionalmente eficientes e intuitivos.

Fusificadores (I)

El fusificador es un mapa que transforma $x^* \in U$ en un conjunto difuso A' en U.

- $\mu_{A'}$ debe tomar valores grandes en x^*.
- A' debe ayudar a suprimir el ruido en x^*.
- Debería ayudar a simplificar las operaciones en el sistema de inferencia difuso.

Fusificadores (II)

- Fusificador singleton:

$$\mu_{A'}(x) = \begin{cases} 1 & \text{si} \quad x = x^* \\ 0 & \text{en otro caso} \end{cases} \tag{24}$$

- Fusificador gaussiano:

$$\mu_{A'}(x) = e^{-(\frac{x_1 - x_1^*}{\sigma 1})^2} * \ldots * e^{-(\frac{x_n - x_n^*}{\sigma 1})^2} \tag{25}$$

- Fusificador triangular:

$$\mu_{A'}(x) = \begin{cases} (1 - \frac{|x_1 - x_1^*|}{b_1}) * \ldots * (1 - \frac{|x_n - x_n^*|}{b_n}) & \text{si} \quad |x_i - x_i^*| \leq b_i \\ 0 & \text{en otro caso} \end{cases} \tag{26}$$

Defusificadores (I)

El defusificador es un mapa que transforma un conjunto difuso B' en V en un valor crisp $y^* \in V$. Este valor debe ser el que mejor representa al conjunto difuso B'.

Se suelen considerar los siguientes criterios:

- Plausibilidad: y^* debería representar intuitivamente el conjunto difuso B'
- Eficiencia computacional: Esta condición depende de la aplicación a considerar.
- Continuidad: Un pequeño cambio en B' debe producir un cambio pequeño en y^*.

Defusificadores (II)

- Defusificador centro de gravedad:

$$y^* = \frac{\int_V y\mu_{B'}(y)dy}{\int_V \mu_{B'}(y)dy} \quad (27)$$

  Si las funciones de pertenencia están normalizadas coincide con $E(y)$.
  El principal inconveniente es su alto coste computacional.

- Defusificador promedio de los centros: Puesto que B' suele ser la unión de M conjuntos difusos, el defusificador anterior se puede aproximar por:

$$y^* = \frac{\sum_{l=1}^{M} \bar{y}^l w_l}{\sum_{l=1}^{M} w_l} \quad (28)$$

donde $\bar{y}_l$ es el centro del l-ésimo conjunto difuso y w_l el valor de la función de pertenencia en dicho punto.
Es el defusificador más utilizado en control.

- Defusificador máximo: Elige y^* como el valor para el que $\mu_{B'}(y)$ es máximo. Si definimos

$$h(B') = \{y \in V | \mu_{B'}(y) = \sup_{y \in V} \mu_{B'(y)}\} \tag{29}$$

si el conjunto tiene más de un punto se puede elegir y^* aleatoriamente, como el promedio, el ínfimo o el supremo.

Sistemas difusos como mapas no lineales (I)

Lema 2. *Sea B^l un conjunto difuso con centro $\bar{y}^l$ y $\mu_{B^l}(\bar{y}^l) = 1$. Consideramos el Sistema de Inferencia Producto (22), Fusificador singleton (24), y defusificador Promedio de los centros (28). La función aproximadora resultante tiene la siguiente forma:*

$$f(x) = \frac{\sum_{l=1}^{M} \bar{y}^l (\prod_{i=1}^{n} \mu_{A_i^l}(x_i))}{\sum_{l=1}^{M} (\prod_{i=1}^{n} \mu_{A_i^l}(x_i))} \tag{30}$$

Proof. Se deja como ejercicio. $\square$

En particular si consideramos como funciones de pertenencia

$$\mu_{A_i^l}(x_i) = a_i^l \exp(-\frac{x_i - \bar{x}_i^l}{\sigma_i})^2, \qquad \mu_{B'}(y) = \exp(-(y - \bar{y}^l)^2) \tag{31}$$

Sistemas difusos como mapas no lineales (II)

La función aproximadora implementada por el sistema difuso se expresa como:

$$f(x) = \frac{\sum_{l=1}^{M} \bar{y}^l [\prod_{i=1}^{n} a_i^l \exp(-\frac{x_i - \bar{x}_i^l}{\sigma_i})^2]}{\sum_{l=1}^{M} [\prod_{i=1}^{n} a_i^l \exp(-\frac{x_i - \bar{x}_i^l}{\sigma_i})^2]} \tag{32}$$

Ejercicio: Obtener la expresión para la función f cuando el sistema de inferencia es el mínimo (23).

Sistemas difusos como mapas no lineales (III)

Lema 3. *Consideramos ahora el Sistema de Inferencia Producto (22), Fusificador Gaussiano (25), y defusificador Promedio de los centros (28) y funciones de pertenencia Gaussianas con $a_i^l = 1$. La función aproximadora resultante tiene la siguiente forma:*

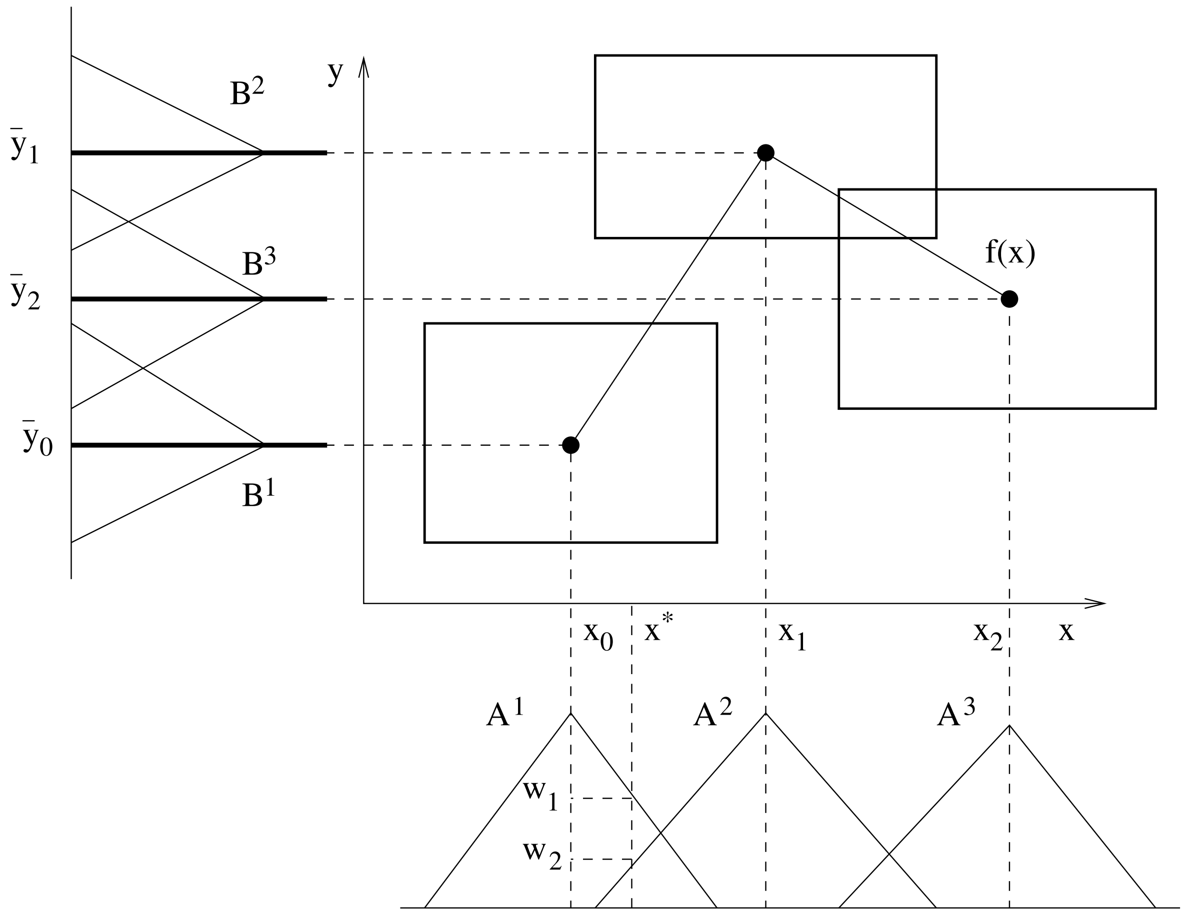
$$f(x) = \frac{\sum_{l=1}^{M} \bar{y}^l \prod_{i=1}^{n} \exp(-\frac{(x_i - \bar{x}_i^l)^2}{a_i^2 + (\sigma_i^l)^2})}{\sum_{l=1}^{M} (\prod_{i=1}^{n} \exp(-\frac{(x_i - \bar{x}_i^l)^2}{a_i^2 + (\sigma_i^l)^2})} \tag{33}$$

Proof. Se deja como ejercicio. □

Teorema 1 (Teorema de aproximación universal)**.** *Sea U un entorno compacto en $\mathbb{R}^n$. Para cualquier función continua $g(x)$ en U y un $\epsilon > 0$, existe un sistema difuso $f(x)$ en la forma (33) tal que $\sup_{x \in U} |f(x) - g(x)| < \epsilon$.*
Es decir, el sistema difuso anterior es un aproximador universal.

Sistemas difusos como mapas no lineales (IV)

Los sistemas difusos realizan una interpolación localmente lineal entre los centros $\bar{y}_l$.



Puesto que w_l en (28) implementa un entorno local puede ser sustituida por una funcion base como la RBF $w_l = K_l(x, c_l, \sigma_l)$. Con esta notación el sistema difuso se expresa como:

$$f(x) = \sum_{l=1}^{M} g_l(x)\bar{y}_l, \tag{34}$$

donde $g_l(x) = K_l(x, c_l, \sigma_l) / \sum_j K_j$ son las funciones base normalizadas.

Los métodos de optimización de las redes RBF pueden ser utilizadas para calcular los parámetros de los sistemas difusos.

Diseño de Sistemas difusos con datos de entrada-salida

Diseño por gradiente descendente (I)

Consideramos un sistema difuso con sistema de inferencia producto, fusificador singleton, defusificador centro de gravedad y funciones de pertenencia gaussianas.
f(x) viene dado por (32), donde M es el número de reglas y $\bar{y}^l$, $\bar{x}_i^l$ y σ_i^l son los parámetros a calcular.

El sistema difuso se puede considerar una red neuronal de tres capas que realiza las siguientes operaciones:

$$z^l = \prod_{i=1}^{n} \exp(-(\frac{x_i - \bar{x}_i^l}{\sigma_i^l})^2) \quad (35)$$

$$b = \sum_{l=1}^{M} z^l \qquad a = \sum_{l=1}^{M} \bar{y}^l z^l \quad (36)$$

$$f(x) = \frac{a}{b} \quad (37)$$

Diseño por gradiente descendente (II)

Definimos el error para el patrón p como:

$$e^p = \frac{1}{2}[f(x_0^p) - y_0^p]^2 \tag{38}$$

Aplicando la regla del gradiente descendente obtenemos reglas de actualización para $\bar{y}^l$, $\bar{x}_i$, σ_i^l:

$$\bar{y}^l(t+1) = \bar{y}^l(t) - \alpha \frac{\partial e}{\partial \bar{y}^l} \tag{39}$$

Utilizando la regla de la cadena:

$$\frac{\partial e}{\partial \bar{y}^l} = (f - y)\frac{\partial f}{\partial a}\frac{\partial a}{\partial \bar{y}^l} = (f - y)\frac{1}{b}z^l \tag{40}$$

Sustituyendo la regla de actualización queda como:

$$\bar{y}^l(t+1) = \bar{y}^l(t) - \alpha(f-y)\frac{1}{b}z^l \tag{41}$$

Diseño por gradiente descendente (III)

- Para determinar $\bar{x}_i^l$ se aplica gradiente descendente:

$$\bar{x}_i^l(t+1) = \bar{x}_i^l(t) - \alpha \frac{\partial e}{\partial \bar{x}_i^l} \tag{42}$$

Como en el caso anterior e depende de $\bar{x}_i^l$ solo a través de z^l. Aplicando la regla de la cadena:

$$\frac{\partial e}{\partial \bar{x}_i^l} = (f - y)\frac{\partial f}{\partial z^l}\frac{\partial z^l}{\partial \bar{x}_i^l} = (f - y)\frac{\bar{y}^l - f}{b} z^l \frac{2(x_{0i}^t - \bar{x}_i^l)}{\sigma_i^{l2}} \tag{43}$$

Quedando la regla de actualización para el parámetro $\bar{x}_i^l$ como:

$$\bar{x}_i^l(t+1) = \bar{x}_i^l(t) - \alpha(f - y)\frac{\bar{y}^l - f}{b} z^l \frac{2(x_{0i}^t - \bar{x}_i^l)}{\sigma_i^{l2}} \tag{44}$$

Diseño por gradiente descendente (IV)

- Por último la actualización del parámetro σ_i^l por gradiente descencente se obtiene como:

$$\sigma_i^l(t+1) = \sigma_i^l(t) - \alpha \frac{\partial e}{\partial \sigma_i^l} \tag{45}$$

de donde haciendo las derivadas obtenemos:

$$\sigma_i^l(t+1) = \sigma_i^l(t) - \alpha \frac{f-y}{b}(\bar{y}^l(t) - f) z^l \frac{2(x_{0i}^t - \bar{x}_i^l(t))^2}{\sigma_i^{l3}(t)} \tag{46}$$