Apuntes de Inteligencia artificial (II)

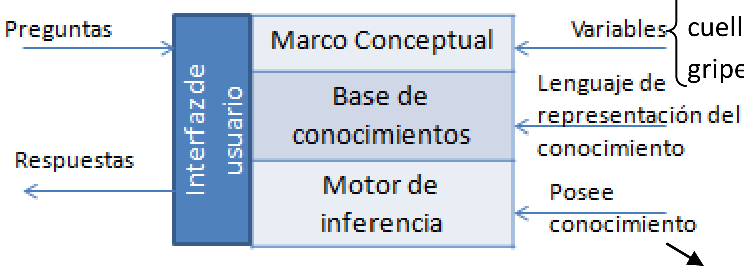
Tema 2 – Razonamiento Aproximado

Profesor: Daniel Manrique

Índice

1. Introducción (Volver)

En cualquier sistema basado en el conocimiento tenemos:



Variables: Por ejemplo: temperatura ∊ [30-50] ºC, Rigidez de cuello ∊ {Sí, No} boolean, diagnóstico ∊ {meningitis, gripe}.

Lenguaje de representación del conocimiento: Lenguaje forma, matemático que especifica nuestro conocimiento acerca de un dominio. Por ejemplo: Reglas, Jerarquías de marcos,…

Permite inferir conocimiento a partir de la base de conocimientos y el marco conceptual.

Se diseñan en base a la base de conocimientos.

Un sistema de este tipo se supone creado para solucionar problemas difíciles, por tanto, no interesa que sea capaz de dar respuestas ópticas sino válidas.

Sistemas con incertidumbre

Hasta ahora hemos visto bases de conocimientos y marcos conceptuales nítidos. Es decir, basado en reglas de probabilidad uno, con seguridad en su respuesta. Por ejemplo:

Si Tº > 37, rigidez _cuello=Sí
Entonces meningitis

Sin embargo, en la vida real esto no funciona de esta manera sino que siempre existe cierto grado de incertidumbre, donde se dan respuestas con un cierto grado de probabilidad. Si se utilizan las leyes clásicas de la probabilidad es un sistema sencillo. Sin embargo, la probabilidad tiene el inconveniente de no diferenciar entre la ignorancia total acerca de un hecho y la certeza parcial.

diagnóstico médico. Este se basa en medidas de certeza. Una medida de certeza (en inglés, *Certain Factor*, CF) que toma valores entre -1 y 1:

$$CF \in [-1, 1]$$

Donde -1 indica que es seguro que la evidencia no apoya la hipótesis y 1 indica que es seguro que la evidencia sí apoya la hipótesis. Por supuesto, el valor 0 representa la ignorancia. No podemos inferir hipótesis alguna con la información disponible.

Lógica borrosa

Como decíamos, hasta ahora hemos visto lógica *crisp*, nítida. Una lógica donde los hechos están claramente representados:

Tº > 37 → Fiebre

La lógica borrosa, *fuzzy logic*, es aquella en la que las fronteras entre los valores de una variable son imprecisas, difíciles de discernir de forma discreta:

Si altura = alto ∧ peso = elevado
Entonces persona = robusta

Donde los valores **alto** y **elevado** son difusos, poco definidos. De esta forma añadiremos un conjunto borroso a los valores de una variable donde los valores originales pertenecen o no a él con un cierto grado de posibilidad de 0 a 1, donde 0 es que NO pertenece seguro y 1 que SÍ pertenece seguro. También añadiremos valores cualitativos (que representan a este conjunto borroso). Por ejemplo:

altura = {bajo, alto, muy alto} Tº ambiental = {muy frío, frío, templado, claro, mucho claro}

Ejemplos

Lanzamiento de un dado

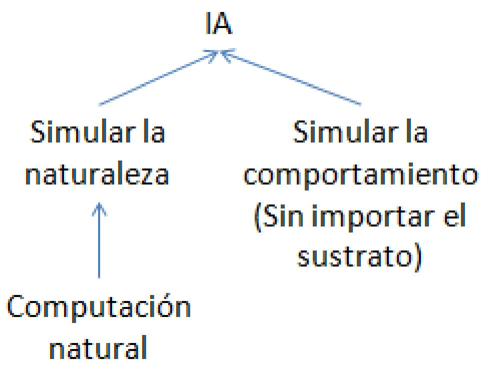
Toma valores del conjunto, {1, 2, 3, 4, 5, 6}. Son valores precisos y sabes que uno de ellos sale seguro. Sin embargo, hay incertidumbre, no sabemos el valor exacto.

Precisión e incertidumbre.

Temperatura ambiental

Medida en un momento dado, Tº = 15,3 ºC. Es una medida real, exacta. Sin embargo, ¿esta temperatura es fría o templada? Es imprecisa, no se sabe con claridad.

Imprecisión y certidumbre.



Nosotros nos encontramos en el punto más a la derecha. Intentamos simular el comportamiento de un sistema inteligente.

2. Razonamiento con imprecisión: Lógica borrosa (Volver)

Índice

1. Fundamentos de lógica borrosa.
2. Razonamiento en lógica borrosa.
3. Controladores difusos.

Fundamentos de lógica borrosa (Volver)

Índice

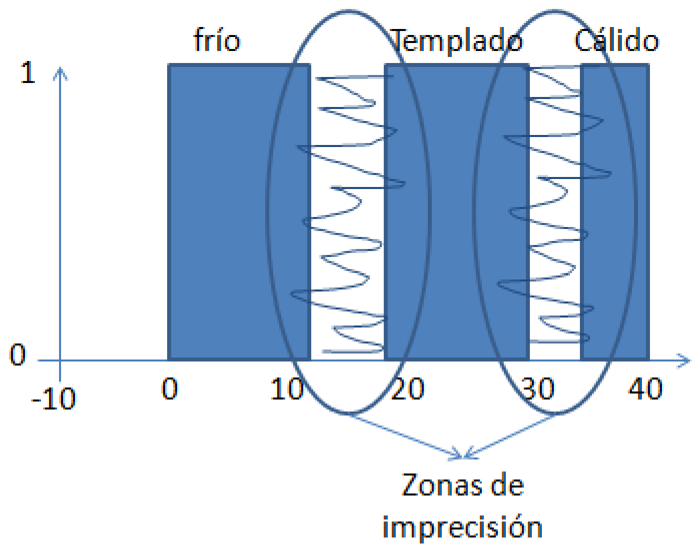
1. Borrosidad, vaguedad, imprecisión.
2. Conjuntos borrosos.
3. Operaciones con conjuntos borrosos.

En un sistema basado en el conocimiento que maneje imprecisión, tenemos un marco conceptual sobre una base de conocimiento sobre un motor de inferencia. El marco conceptual en un sistema impreciso posee las variables, su dominio y el significado de los mismos. La diferencia es que las variables tendrán un dominio de valores numéricos (Temperatura=20, 25, 19º) y otro dominio de conceptos imprecisos cualitativos, llamados etiquetas lingüísticas que poseen unas funciones de pertenencia y funciones de distribución de posibilidad (Temperatura = muy frio, frio, calor).

Estas funciones asignan el grado de pertenencia de cada valor numérico a las etiquetas lingüísticas, esta función toma valores entre 0 y 1, pero esto no es probabilidad sino posibilidad. Las etiquetas describen un conjunto borroso.

Borrosidad, vaguedad, imprecisión (Volver)

Representa aquellos conceptos cuyas fronteras no están bien definidas.



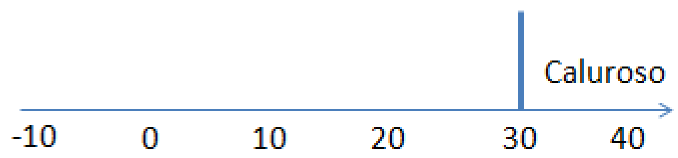
En este pequeño ejemplo suponemos que el valor de etiqueta de una variable temperatura es T = {frío, templado, cálido} y que sus valores numéricos son T = [-10..40]ºC

Como vemos, en el gráfico de posibilidad de dichos valores existe una zona intermedia a cada valor de etiqueta que es la zona de imprecisión donde no sabríamos decir a qué etiqueta pertenece cada número. Veamos esto con más profundidad.

Conceptos

La lógica borrosa permite representar conjuntos con **fronteras no precisas**. La afirmación "x pertenece a A" (Donde x es un valor del dominio de la variable y A una etiqueta lingüística del sistema) no es cierta o falsa, sino que es medible mediante una **posibilidad** en [0,1]. Este sistema permite manejar la vaguedad o imprecisión. Por ejemplo:

- "Día caluroso": la frontera entre caluroso y templado no es exacta. Podemos decir:



- Pero, entonces, ¿se puede suponer que si hay 30 ºC el día es caluroso, pero si hay 29 ºC entonces ya no lo es? No existe una respuesta precisa.

Imprecisión se refiere a vaguedad, fronteras mal definidas. Podemos definir la **Incertidumbre** con un ejemplo: si lanzamos un dado hay seis posibilidades (precisas), pero no se desconoce qué saldrá. No hay imprecisión pero si hay incertidumbre.

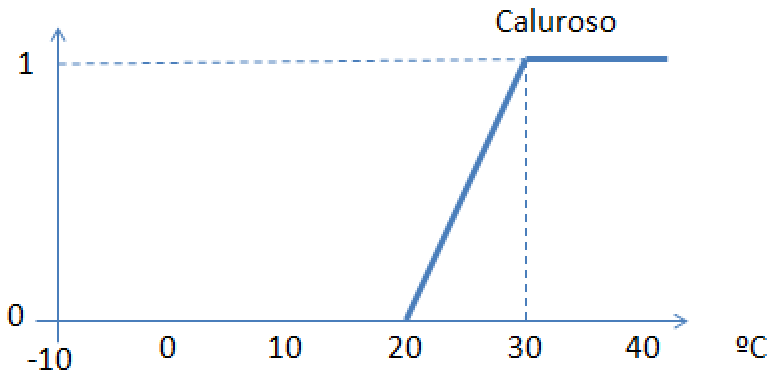
Aplicaciones de la lógica borrosa:

Conjuntos borrosos (Volver)

Un conjunto borroso es aquél en donde la pertenencia de los elementos se define mediante una **función de pertenencia** o **función de distribución de posibilidad**. Definida como:

$$\mu_A: X \to [0,1]$$

El conjunto borroso A tiene una función de pertenencia $\mu_A(x)$ que asigna a cada valor $x \in X$ un número entre 0 (no pertenece) y 1 (sí pertenece). Los valores intermedios de esta función representan pertenencia parcial. Por ejemplo: $\mu_{caluroso}$: Tª → [0,1]:

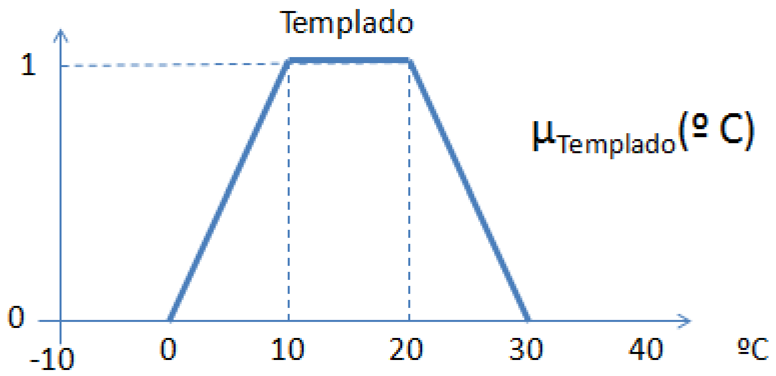


Distribuciones típicas

Existen tres distribuciones típicas, dependiendo de lo bien definido que esté el conjunto de valores numéricos y las etiquetas asociadas a esa distribución.

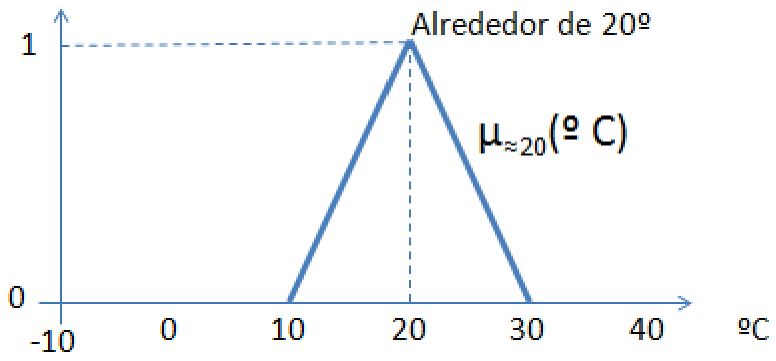
Trapezoidales

Los laterales no están bien definidos pero hay una parte del dominio bien definido.



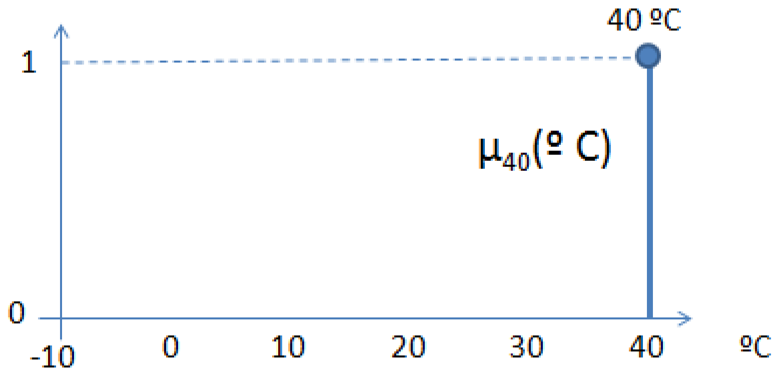
Triangulares

Solo existe un punto preciso y los laterales a ese punto son imprecisos. Este tipo de distribución se puede utilizar para representar valores difusos que parecen nítidos. Por

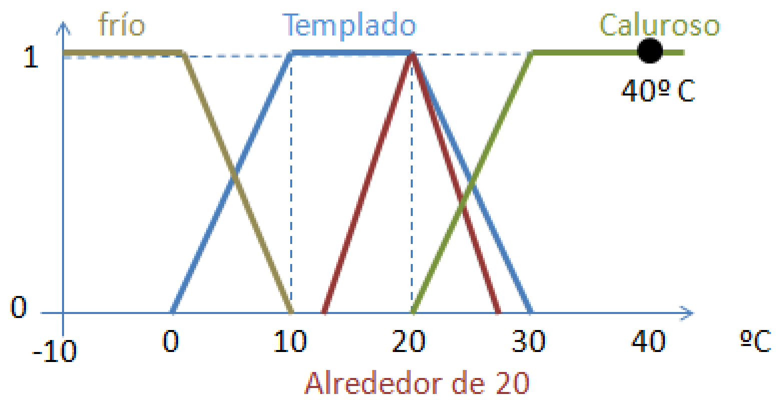


Nítidas

Solo existe un punto preciso, sin valores imprecisos a su alrededor. No hay posibilidades parciales.



Ejemplo



$$\mu_{Frio}(ºC),\ \mu_{Templado}(ºC),\ \mu_{\approx 20}, \mu_{caluroso}(ºC), \mu_{40ºC}$$

Hay que tener cuidado con los intervalos de los dominios, ya que si un extremo está abierto

Resumiendo

Variables
- Dominio numérico: **T ϵ [-10, 40]**
- Dominio Etiquetas cualitativas **T = {frio, templado, caluroso}**

Conjunto borroso $\xleftarrow{\text{define}}$ Distribución de posibilidad **μ_{frio}(T), T → [0,1]**

Operaciones con conjuntos borrosos (Volver)

A partir de la cuantificación de la posibilidad (Cuantificación de la permanencia) de que sea cierta "x es p" expresado como $\mu_p(x)$:

- x es p y q: $\mu_{p \wedge q}(x)$. El valor x pertenece a los dos conjuntos difusos. Una conjunción. Para calcularla necesitamos realizar una operación que veremos más adelante que resultará en una nueva distribución de posibilidad: $\mu_{(p \wedge q)}(x)$.
- x es p ó q: $\mu_{p \vee q}(x)$. Se eliminan los valores de p y q que sean conjuntos. Por tanto es la disyunción entre los conjuntos p y q.
- x no es p: $\mu_{\sim p}$ (x). Negación. Si se niega un conjunto quiere decir que el valor es todo lo demás.
- Si x es p, entonces x es q: $\mu_{p \to q}(x)$. Implicación.

Generalizando a n dimensiones:

$$\mu_{p \wedge q}(x_1, x_2, \ldots, x_n),\ \mu_{p \vee q}(x_1, x_2, \ldots, x_n), \text{ etc.}$$

Teoría de conjuntos

Matemáticas

Frío y templado son conjuntos difusos. Con funciones de posibilidad del tipo $\mu_p(x)$ donde x ϵ p. En ella se pueden establecer las operaciones:

- Intersección: ∩
- Unión: ∪
- Negación: ¬
- Implicación: No podemos hacer directamente p → q pero podemos decir:

$$p \to q = \neg p \vee q \to \neg p \cup q$$

Lógica

Frío y templado son etiquetas con su distribución de posibilidad. En ella se pueden establecer

Existe una relación muy estrecha entre ambas visiones de la teoría de conjuntos y, hasta cierto punto, podríamos decir que son equivalentes, aunque hay ciertos matices diferentes. No son equivalentes pero dan los mismos resultados.

Extensión cilíndrica

Es una técnica matemática que nos permite dadas dos funciones de posibilidad de dos variables diferentes crear otra función de posibilidad de ambas variables. Para componer dos funciones, deben estar **referidas a las mismas variables**. No es posible componer x es p e y es q:

$$\mu_p(x) \wedge \mu_q(y)$$

Para ello se calcula la extensión cilíndrica de $\mu_p(x)$ con y y la extensión cilíndrica de $\mu_q(y)$ con x para obtener $\mu_p(x,y)$ y $\mu_q(x,y)$. Entonces es posible:

$$\mu_{p \wedge q}(x,y)$$

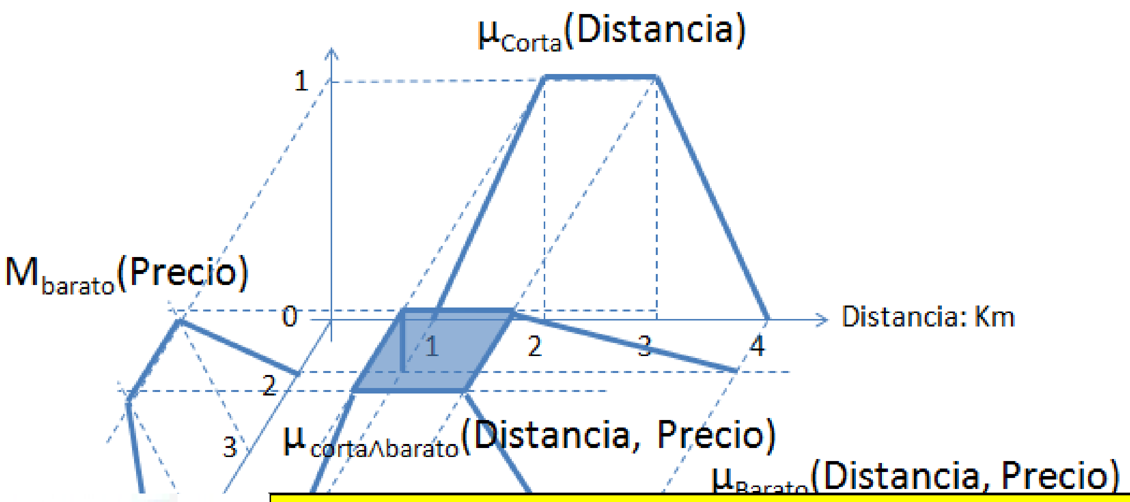
La extensión cilíndrica de $\mu(x_1, x_2,\ldots, x_n)$ con y es $\mu(x_1, x_2, \ldots, x_n,y)$ tal que cumple:

$$\forall\ y\ \mu(x_1, x_2, \ldots, x_n, y) = \mu(x_1, x_2, \ldots, x_n)$$

Básicamente lo que quiere decir es que añadas la nueva variable, una nueva dimensión, pero que no afecte al resto de variables. Vamos, que no le hagas ni caso ya que no importa el valor de la nueva variable.

Ejemplo

Se desea representar que la tarifa de un taxi es barata cuando la carrera es corta. Se tiene **μ_{corta}(distancia)** y **μ_{barato}(precio)**. Se desea calcular **$\mu_{corta \wedge barato}$(distancia, precio).**



t-Norma, t-Conorma y negación

Se manejan las siguientes funciones:

Funciones T: t-normas

Toma dos distribuciones conocidas de posibilidad y devuelve una nueva distribución, que no era conocida, con la conjunción de las distribuciones de la entrada. Conjunción (lógica), intersección (conjuntos).

$$T(x,y).\ \mu_{p \wedge q}(x) = T(\mu_p(x), \mu_q(x))$$

Es una función que devuelve un valor [0,1]. Representa la conjunción o intersección. Si $\forall x, y, z \in [0,1]$, las propiedades de la t-norma son:

1. $T(x,1) = x$ (elemento neutro)
2. $x \leq y \rightarrow T(x,z) \leq T(y,z)$ (monotonía)
3. $T(x,y) = T(y,x)$ (conmutativa)
4. $T(x,T(y,z)) = T(T(x,y),z)$ (asociativa)

Ejemplos:

- $T(x,y) = \text{mín}(x,y)$ (mínimo) Posiblemente la más utilizada
- n? $T(x,y) = P(x,y) = x \cdot y$ (producto)
- n? $T(x,y) = W(x,y) = \text{máx}(0,x+y-1)$ (Lukasiewicz)
- n? $T(x,y) = Z(x,y) = x$ si $y=1$; y si $x=1$; 0 resto (drástica)

Funciones S: t-conormas

Toma dos distribuciones conocidas de posibilidad y devuelve una nueva distribución, que no era conocida, con la disyunción de las distribuciones de la entrada. Disyunción (lógica), unión (conjuntos).

$$S(x,y).\ \mu_{p \vee q}(x) = S(\mu_p(x), \mu_q(x))$$

Es una función binaria que devuelve un valor [0,1]. Representa la disyunción. Si $\forall x, y, z \in [0,1]$, las propiedades de la t-conorma son:

1. $S(x,0) = x$ (elemento neutro)
2. $x \leq y \rightarrow S(x,z) \leq S(y,z)$ (monotonía)
3. $S(x,y) = S(y,x)$ (conmutativa)
4. $S(x,T(y,z)) = S(T(x,y),z)$ (asociativa)

Ejemplos:

Negación

Toma una distribución conocida de posibilidad y devuelve una nueva distribución, que no era conocida, con la negada de la distribución de la entrada.

$$N(x).\ \mu_{\sim p}(x) = N(\mu_p(x)).$$

Es una función unaria que devuelve un valor [0,1]. Para que sea intuitiva debe cumplir $\forall x \in [0,1]$. Sus propiedades son:

1. $N(0) = 1; N(1) = 0$ (condiciones frontera)
2. $x \leq y \rightarrow N(x) \geq N(y)$ (inversión de monotonía)

Ejemplos:

- $N(x) = (1-x)$
- $N(x) = (1-x^2)^{1/2}$

Dualidad

Se dice que una t-norma y una t-conorma son duales con respecto a una negación sii:

- $N(T(x,y)) = S(N(x), N(y))$
- $N(S(x,y)) = T(N(x), N(y))$

Son duales con respecto a $N(x) = 1-x$:

- $< \text{mín}(x,y), \text{máx}(x,y) >$
- $< P(x,y), P'(x,y) >$
- $< W(x,y), W'(x,y) >$
- $< Z(x,y), Z'(x,y) >$

$T(x, y)$ y $S(x, y)$ deben ser duales con respecto a $N(x)$.

Ejercicio

Demostrar que $P(x,y)$ y $P'(x, y)$ son duales con respecto a la negación $N(x)=1-x$.

Solución 1

Debemos llegar a $N(T(x, y)) = S(N(x), N(y))$. Con las operaciones elegidas será:

$N(P(x,y)) = P'(N(x), N(y)) \Rightarrow 1-x\cdot y = P'(1-x, 1-y) \Rightarrow 1-x\cdot y = 1-x\cdot 1-y-[(1-x)\cdot(1-y)] \Rightarrow$

$\Rightarrow 1-x\cdot y = 1-x\cdot 1-y-[1-y-x+x\cdot y] \Rightarrow 1-x\cdot y = 1-x+1-y-1+y+x-x\cdot y \therefore 1-x\cdot y = 1-x\cdot y$ **CQD**

Solución 2

- T(x,y), S(x,y) y N(x)
- Conjuntos:
    - Intersección, unión y complementario
- Lógica:
    - Conjunción, disyunción y negación

En lógica borrosa se define la implicación a través de la función binaria J (x,y) que devuelve un valor [0,1].

$$\mu_{p\rightarrow q}(x) = J(\mu_p(x), \mu_q(x))$$

De la lógica clásica, se tiene que:

- A→B = ¬AVB
- J(x,y) = S (N(x),y).

Tomando N(x) = 1-x, se tiene:

- Con S(x,y) = máx (x,y) **J(x,y) = máx (1-x,y)**
- Con S(x,y) = P'(x,y) = x+y-x·y **J(x,y) = 1-x+x·y**
- S(x,y) = W'(x,y) = mín (1,x+y) **J(x,y) = mín (1,1-x+y)**
- S(x,y) = Z'(x,y) = x si y=0; y si x=0; 1 resto **J(x,y) = y si x=1; 1-x si y=0; 1 resto**

En la lógica clásica se cumple:

AVB = AV (¬A^B)

Puesto que:

A→B = ¬AVB

Se tiene que:

A→B = ¬A V (A^B)

Por tanto:

J(x,y) = S (N(x), T(x,y))

Utilizando la negación N(x) = 1-x:

- Con <min,max> **J(x,y) = máx (1-x,mín(x,y))**
- Con <P , P'> **J(x,y) = 1-x + x2y**

Razonamiento en lógica borrosa (Volver)

A partir de aquí seremos capaces de preguntar al sistema otorgándole datos y recibir una respuesta basándose en sistemas de inferencia. En estas transparencias veremos dos tipos de inferencia.

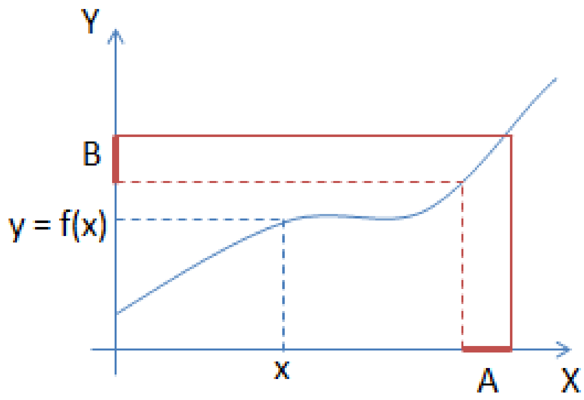
Regla composicional de inferencia. Que es un modus ponens borroso.

Inferencia en conjuntos

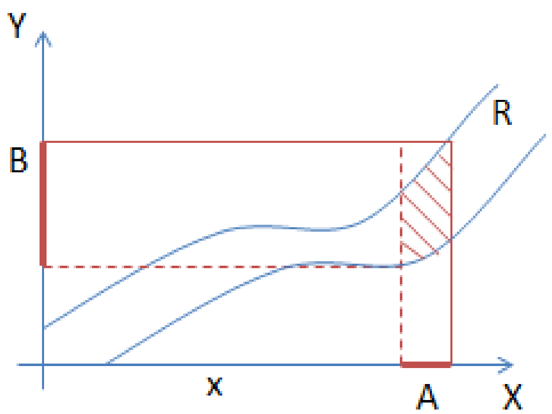
La inferencia en conjuntos clásicos se puede ver de la siguiente forma:

Dados X e Y dos conjuntos, la relación entre ambos viene dada por la función f(X) donde:

- Dado un valor X=x, se puede inferir y=f(x).
- Dado un subconjunto A⊂X, se puede inferir el subconjunto B⊂Y: B={y∊Y / y=f(x), x ∊ A}



Dada una relación R cualquiera entre los conjuntos X e Y. Dado el subconjunto A⊂X, se puede inferir el subconjunto B⊂Y, B={y∈Y / <x,y>∈R, x ∈ A}

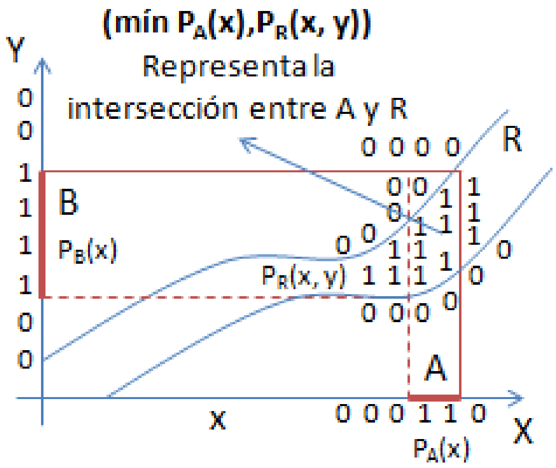


Función característica

Se define la función característica de un conjunto A, $P_A(x)$, a una función de pertenencia definida como:

- $P_A(x)$=1, si x ∊ A

$$P_B(y) = \operatorname*{Sup}_{x \epsilon X}\{mín\{P_A(x), P_R(x,y)\}\}$$



Regla composicional de inferencia

Partiendo de la expresión anterior:

$$P_B(y) = \operatorname*{Sup}_{x \epsilon X}\{mín\{P_A(x), P_R(x,y)\}\}$$

Generalizando para conjuntos difusos:

$$\mu_B(y) = \operatorname*{Sup}_{x \epsilon X}\{mín\{\mu_A(x), \mu_R(x,y)\}\}$$

La intersección puede emplear una T-norma cualquiera:

$$\mu_B(y) = \operatorname*{Sup}_{x \epsilon X}\{T\{\mu_A(x), \mu_R(x,y)\}\}$$

Si R puede ser una relación de implicación, entonces, la **Regla Composicional de Inferencia** (RCI) es:

$$\mu_{B'}(y) = \operatorname*{Sup}_{x \epsilon X}\{T\{\mu_A(x), \mu_{A\to B}(x,y)\}\}$$

Modus Ponens generalizado

Con la RCI, se puede hacer inferencia con modus ponens:

Regla: Si X es A, entonces Y es B
Hecho: X es A.
Conclusión: Y es B

Inferencia borrosa con RCI

Tenemos los datos:

- X es A: $\mu_A(x)$
- Y es B: $\mu_B(y)$
- X es A': $\mu_{A'}(x)$

Nuestro objetivo es calcular $\mu_{B'}(y)$. Debemos hacer lo siguiente:

1. Calcular **$\mu_{A\rightarrow B}(x,y)$** a partir de $\mu_A(x)$ y $\mu_B(y)$
    a. Es necesario realizar las **extensiones cilíndricas**: $\mu_A(x)$ con y, $\mu_B(y)$ con x.
    b. Se emplea para la implicación cualquier función J(x,y).
        i. Zadeh: **J(x,y)=máx {1-x, mín(x,y)}**.
2. Calcular **T($\mu_{A'}(x)$, $\mu_{A\rightarrow B}(x,y)$)**.
    a. Se realiza la **extensión cilíndrica** de $\mu_{A'}(x)$ con y.
    b. Se emplea una T-norma. **T(x,y) = mín (x,y)**.
3. **Calcular el supremo** en x de T($\mu_{A'}(x)$, $\mu_{A\rightarrow B}(x,y)$), que está en función de x e y.
    a. Se obtiene la distribución de posibilidad sobre la dimensióny: **$\mu_{B'}(y)$**, es decir, las posibilidades de y es B'.

De forma más resumida podemos decir que los pasos para convertir la fórmula a lógica borrosa son:

1) Extensión cilíndrica de $\mu_A(x)$ con y para obtener $\mu_A(x,y)$
2) Extensión cilíndrica de μ_B (x) con x: $\mu_B(x,y)$
3) J(μ_A (x,y), μ_B (x,y)) = mu_A->B(x,y)
4) Extensión cilíndrica de $\mu_{A'}(x)$ con y: $\mu_{A'}(x,y)$
5) T($\mu_{A'}(x,y)$, $\mu_{A\text{->}B}(x,y)$) = $\mu_{A\wedge(A\text{->}B)}(x,y)$
6) $\mu_{A'}(y) = \ Sup_{x\epsilon X}(\mu_{A'\wedge(A\rightarrow B)}(x,y))$

Motor de inferencia borroso

Tenemos la siguiente base de conocimiento:

- R1: Si X es A_1, entonces Y es B_1.
- R2: Si X es A_2, entonces Y es B_2.
- ....
- Rn: Si X es A_n, entonces Y es B_n.
- Hecho: X es A'

Se aplica la RCI a cada regla, obteniendo:

Desborricificación

Denominado en inglés defuzzyfy. Con esto tratamos de convertir una distribución de posibilidad en un resultado más comprensible, bien para nosotros los humanos o bien para las máquinas. Existen dos formas de desborrificar:

- Valor numérico. Fácilmente comprensible por máquinas.
  - Para hallarlo utilizamos el método del **Centro de gravedad**.
- Etiqueta lingüística. Fácilmente comprensible por humanos. Tenemos dos métodos:
  - Distancia
  - Σ cuenta

Valor numérico

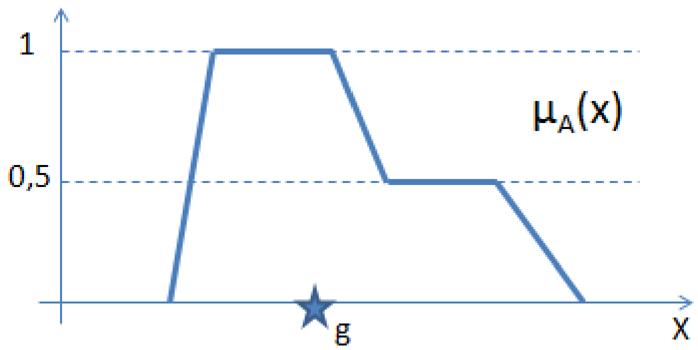
Método del centro de gravedad

Cálculo de la proyección en el eje x del centro de gravedad de la distribución. Se debe muestrear suficientemente fino para cubrir adecuadamente la función. Un muestreo excesivo exige gran carga computacional. Para generalizar decimos que:

$$\mu_A(x) = \mu_{B'}(y)$$

La fórmula sería como sigue:

$$g_{B'} = \frac{\sum_{i=1}^{N} x_i \cdot \mu_A(x_i)}{\sum_{i=1}^{N} \mu_A(x_i)}$$



Etiqueta lingüística

Método Σcuenta

Dado un conjunto de distribuciones de posibilidad: $\mu_{B1}(x)$, $\mu_{B2}(x)$,..., $\mu_{Bn}(x)$ nuestro objetivo es encontrar qué $\mu_{Bi}(x)$ se parece más a una dada $\mu_A(x)$.

La Σcuenta es una medida proporcional al área de la distribución. Se basa en la discretización

$$\Sigma cuenta\ (Bi/A) = \frac{\Sigma cuenta(A \wedge B_i)}{\Sigma cuenta(B_i)}$$

Método de la distancia

Dado un conjunto de distribuciones de posibilidad $\mu_{B1}(x)$, $\mu_{B2}(x)$,…, $\mu_{Bn}(x)$ nuestro objetivo será encontrar qué $\mu_{Bi}(x)$ **se parece más** a una dada $\mu_A(x)$. Se emplea una función distancia definida como:

$$d_i = \sqrt{\alpha \cdot (B_i - A)^2 + \beta \cdot (g_i - g)^2}$$

Donde:

- Bi y A son, respectivamente, las áreas de $\mu_{Bi}(x)$ y $\mu_A(x)$.
- g_i y g son, respectivamente, los centros de gravedad de $\mu_{Bi}(x)$ y $\mu_A(x)$.
- α y β son dos constantes: $\alpha + \beta = 1$.

El método elige como **más parecido** a $\mu_A(x)$ aquella $\mu_{Bi}(x)$ que **minimiza d_i**.

Para ver un ejemplo de cómo utilizar estos métodos conviene leer la segunda parte del ejemplo de RCI que se encuentra también en FIWIKI.

Controladores difusos (Volver)

Los controladores borrosos. Se activan cuando sucede un hecho previamente definido. Por ejemplo, un termostato se activa cuando baja la temperatura de un límite previamente definido. Es un sistema muy sencillo de construir pero es muy lento. No puede utilizarse para sistemas en tiempo real.

Si lo que necesitamos es reaccionar a tiempo, en un sistema de tiempo real (generalmente en cualquier sistema donde se monten personas) podemos utilizar controladores con ecuaciones diferenciales, mucho más complicados de diseñar pero mucho más rápidos.

Se emplean para controlar sistemas **inestables**. Si quiero mantener una variable en unos valores concretos y existen ciertas condiciones que desestabilizan el valor, necesito controladores para actuar en consecuencia y devolver el valor a la situación de equilibrio. El control tiene por objeto garantizar una salida en el sistema a pesar de las **perturbaciones** que le afectan. Ejemplos:

- Sistemas de **navegación**.
- Sistemas de **climatización**.
- Sistemas de **ventilación** (túneles, garajes).

Esquema de control difuso

Variables de estado

En qué situación me encuentro en cada momento. Por ejemplo, a que temperatura estoy. También cual es el desvío, la tendencia y el estado del elemento a controlar. Suelen ser:

- $s_1, s_2, \ldots, s_n$: estado del sistema. *Temperatura*.
- $e_1, e_2, \ldots, e_n$: desvíos (error) con respecto al valor de referencia. *Temperatura con respecto a 20 ºC*.
- $\Delta e_1, \Delta e_2, \ldots, \Delta e_n$: tendencia del error.
- $u_1, u_2, \ldots, u_m$: estado del elemento a controlar (actuador). *Apertura de válvula, caudal de gas*.

Pueden tomarse como entradas muchas variables. Siempre dependiendo del elemento concreto.

Acciones

v1, v2, …, vm: cambios e realizar en los actuadores. *Abrir o cerrar en un cierto grado válvula o caudal de gas*.

Modelo de control difuso

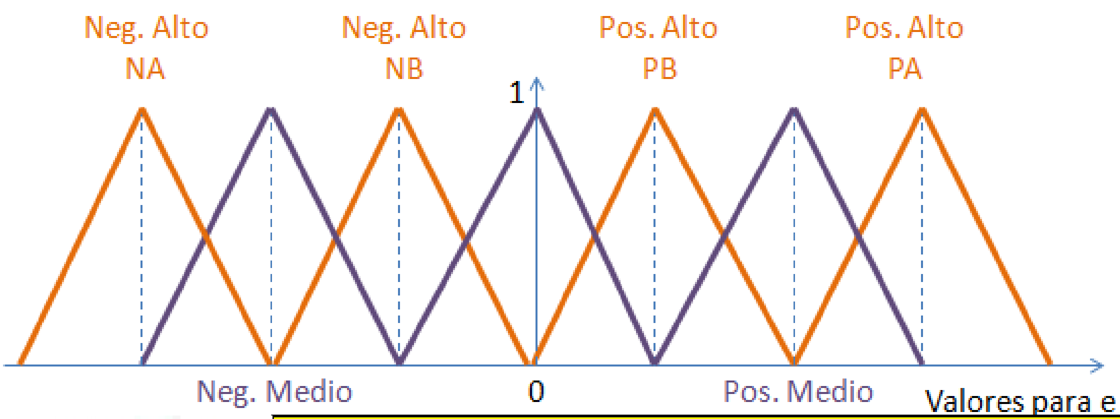
Los valores de las variables de estado y acciones pueden ser etiquetas lingüísticas (valores cualitativos) que son representados por funciones de posibilidad:

$$\text{Si } s_1=A_1,\ldots,\ e_1=B_1,\ldots,\ \Delta e_1=C_1,\ldots,\ u_1=D_1,\ldots,\ \text{entonces } v_1=E_1,\ldots$$

En lo que sigue, se utilizarán reglas del tipo:

$$\text{Si } e=A \text{ y } \Delta e=B, \text{ entonces } v = C$$

Cada valor de e, Δe y v serán etiquetas lingüísticas con su correspondiente función de posibilidad. Por ejemplo, para e:



La notación será la siguiente:

Si e = A(e) y Δe=B(Δe), entonces v = C(v)

Un controlador difuso posee una **base de conocimiento** de reglas del tipo:

- R1: Si e = A_1(e) y Δe=B_1(Δe), entonces v = C_1(v)
- R2: Si e = A_2(e) y Δe=B_2(Δe), entonces v = C_2(v)
- …
- Rn: Si e = A_n(e) y Δe=B_n(Δe), entonces v = C_n(v)

En un momento el estado del sistema es:

e = A(e) y Δe=B(Δe)

Tras un proceso de inferencia, el controlador responde para mantener el equilibrio: **v = C(v)**

Regla composicional de inferencia

El proceso de inferencia emplea la regla composicional de inferencia:

$$\mu_{B'}(y) = \operatorname*{Sup}_{x\epsilon\ X}\{T(\mu_{A'}(x), \mu_{A\to B}(x,y))\}$$

Donde:

y es v. La acción a tomar.

x es bidimensional: e y Δe.

$\mu_{B'}$(y) = C(v). Distrib. de posib. de la acción a tomar.

$\mu_{A'}$(x) = A(e) ∧ B(Δe). Estado actual del sistema.

$\mu_{A\to B}$(x,y) = Ri: Ai(e) ∧ Bi(Δe) → v = C_i(v)

$$C(v) = \underset{e,\Delta e}{Sup}\{T(A(e) \wedge B(\Delta e), A_i(e) \wedge B_i(\Delta e) \to C_i(v))\}$$

Inferencia en controladores

Las operaciones de los controladores borrosos son:

- T(x,y) = min(x,y)
- S(x,y) = máx(x,y)
- N(x) = 1-x
- J(x,y) = min(x,y) Mamdani

$$C(v) = \underset{e,\Delta e}{Sup}\left\{min\left(\underbrace{\min}_{T}\left(A(e)\wedge B(\Delta e)\right), \underbrace{\min}_{J}\left(min\left(A_i(e)\wedge B_i(\Delta e)\right), C_i(v)\right)\right)\right\}$$

Por asociatividad:

$$C(v) = \underset{e,\Delta e}{\text{Sup}}\left\{\min\left(\min\left(\min\left(A(e)\wedge B(\Delta e)\right), \min\left(A_i(e)\wedge B_i(\Delta e)\right)\right), C_i(v)\right)\right\}$$

$$C(v) = \min\left(\underset{e,\Delta e}{Sup}\left\{\min\left(\min\left(A(e), B(\Delta e)\right), \min\left(A_i(e), B_i(\Delta e)\right)\right)\right\}, C_i(v)\right)$$

$$C(v) = \min\left(min\left(\overbrace{\underset{e}{Sup}\left\{min\left(\underbrace{A(e), A_i(e)}_{Antecedente\ regla}\right)\right\}}^{Entrada\ al\ sistema}, \overbrace{\underset{\Delta e}{Sup}\left\{min\left(\underbrace{B(\Delta e), B_i(e)}_{Antecedente\ regla}\right)\right\}}^{Entrada\ al\ sistema}\right), C_i(v)\right)$$

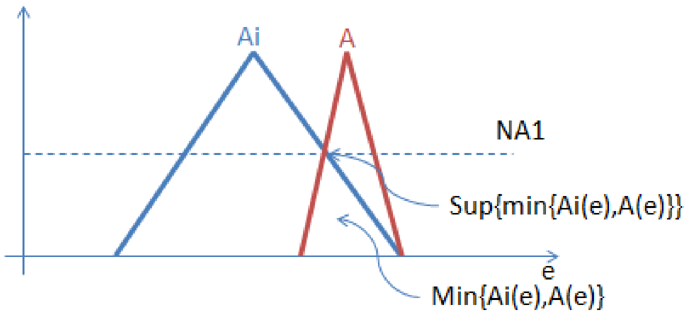
$$C(v) = \min\left(min\left(\underset{e}{Sup}\left\{min\left(A(e), A_i(e)\right)\right\}, \underset{\Delta e}{Sup}\left\{min\left(B(\Delta e), B_i(\Delta e)\right)\right\}\right), C_i(v)\right)$$

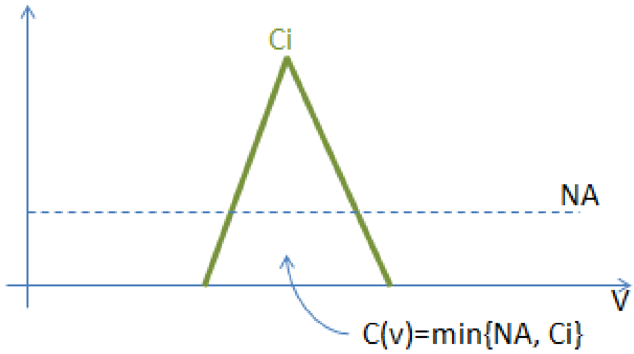
$$C(v) = \min\left(min(NA_1, NA_2), C_i(v)\right)$$

$$C(v) = \min\left(NA, C_i(v)\right)$$

Los que llevan subíndices provienen de las reglas, A_i es A de la regla i. Los que no llevan subíndices son los hechos que me dan como dato. NAx = Nivel de Ajuste del antecedente x de la regla. El nivel de ajuste me indica cómo se ajusta ese antecedente al hecho. Es un valor entre 0 y 1 donde 1 es un ajuste total del antecedente. Si es 0 no se puede lanzar la regla directamente por ser conjunciones.

Interpretación geométrica





Procedimiento general

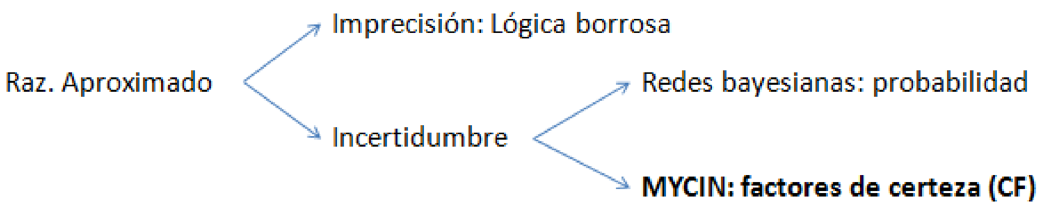
Dada una base de conocimiento:

- R1: Si e = $A_1(e)$ y $\Delta e=B_1(\Delta e)$, entonces v = $C_1(v)$
- ...
- Rn: Si e = $A_n(e)$ y $\Delta e=B_n(\Delta e)$, entonces v = $C_n(v)$
- Hecho: e = A(e) y $\Delta e=B(\Delta e)$.
- Objetivo: v = C(v)

Debe seguirse el siguiente procedimiento:

1. Para cada regla Ri, se obtiene:
    a. $NA=\text{mín}\{Sup_e[\text{mín}(A(e),A_i(e))],Sup_{\Delta e}[\text{mín}(B(\Delta e),B_i(\Delta e))]\}$
    b. La distribución del consecuente $C_{i'}(v)=\text{mín}(NA,C_i(c))$
2. Se obtiene la unión: de los $C_{i'}(v)$ de las reglas que se disparan:
    a. $C(v) = \cup_i \{C'_i(v)\} = \text{máx}_i \{C'_i(v)\}$
3. Desborrocificación de C(v)

Razonamiento con incertidumbre (volver)



La incertidumbre quiere decir que no sabemos exactamente qué va a salir como resultado.

es muy común cuando tratamos de representar la inferencia en el pensamiento de un profesional de la medicina que siempre trabaja con incertidumbre.

MYCIN

El problema de representar esta incertidumbre mediante probabilidad es que no diferencia entre la ignorancia y la igualdad de certeza. Nosotros representamos la ignorancia de un hecho dando a cada resultado la misma probabilidad (tiramos una moneda y, como no conocemos el resultado, hay un 50% de probabilidad para cada resultado. Sin embargo, si después de un proceso de estudio determinamos que la probabilidad es igual para cada resultado (no hay ignorancia) no sabremos si estamos en el caso anterior (parece ignorancia).

MYCIN, sin embargo, no utiliza probabilidades sino factores de certeza (CF), que es un conjunto de valores entre -1 y 1. Al contrario que las redes bayesianas, MYCIN no tiene un transfondo ni una demostración matemática pero funciona en la realidad. Veamos cómo funciona. Los factores de certeza son como siguen:

$CF \in [-1, 1]$ donde:

Si CF < 0 → Certeza en contra del hecho (mayor certeza cuanto más cerca del -1)
CF(Llueve) = -0,8 (si hace sol)

Si CF > 0 → Certeza a favor del hecho (mayor certeza cuanto más cerca del 0)
CF(Soleado) = +0,8 (si hace sol)

CF = 0 → ignorancia acerca de un hecho
CF(mañana_llueve) = 0

En una base de conocimiento también asignaríamos una CF a cada regla donde, si es positivo el antecedente apoya la hipótesis si es negativo los antecedentes no apoyan la hipótesis.

Así que tenemos dos tipos de certezas:

- CF(e) es una certeza que asignamos a un hecho.
- CF(h, e) es una certeza que asignamos a las reglas.

Como MYCIN es un lenguaje empírico podemos agregar etiquetas lingüísticas a cada valor de certeza. Por ejemplo:

CF = 1 Totalmente cierto		CF=-0,5 posiblemente falso
CF = 0,8 Casi seguro cierto	CF = 0 ignorancia	CF = -0,8 Casi seguro falso
CF = 0,5 Posiblemente cierto		CF = -0,5 Posiblemente falso

Un sistema MYCIN se basa en sistemas basados en reglas, como sistemas de producción. Tenemos, por tanto, una base de hechos y un motor de inferencias.

Ejemplo

Base de reglas

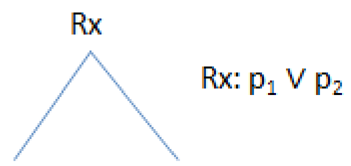
- **R1**: $h_2 \vee h_3 \rightarrow h_1$; CF = 0,5 (h_2 y h_3 están a favor de que se cumple h_1, es decir, si se cumple h_2 o h_3 posiblemente se cumpla h_1).
- **R2**: $h_4 \rightarrow h_1$; CF = 1
- **R3**: $h_5 \wedge h_6 \rightarrow h_3$; CF = 0,7
- **R4**: $h_7 \rightarrow h_3$; CF=-0,5 (h_7 está en contra de que se cumple h_3, es decir, si se cumple h_7 posiblemente es falso que se cumple h_3).

Base de hechos

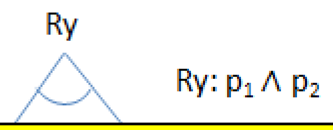
- CF(h_2) = 0,6 [Posiblemente se cumple h_2]
- CF(h_4) = 0,6
- CF(h_5) = 0,3
- CF(h_6) = 0,9
- CF(h_7) = 0,5

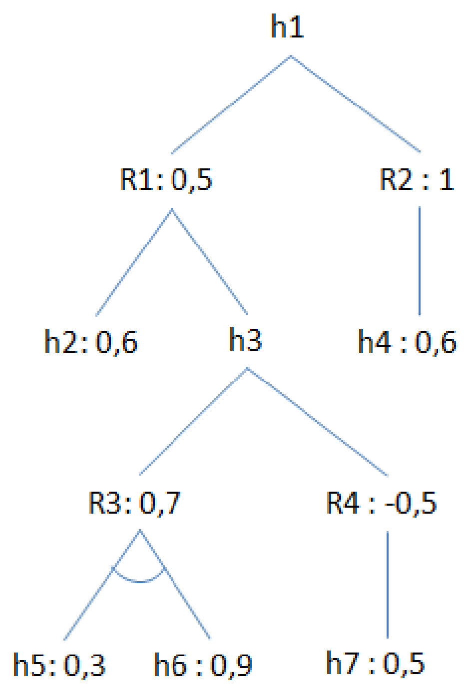
Para el resto de valores o bien su valor es 0 (ignorancia) o sabemos que se inferirá más tarde. La pregunta del ejemplo es: ¿Se cumple h_1?

Bien, lo primero que podemos hacer es pintar un grafo Y/O para deducir h_1 de forma gráfica y después utilizar encadenamiento hacia atrás apoyándote en el grafo. Un grafo Y/O es un grafo que pinta las reglas y sus antecedentes hacia detrás y marcando si los antecedentes se unen mediante una conjunción o mediante una disyunción. Si las premisas de una regla Rx se unen mediante una disyunción se genera un ángulo normal:



Si las premisas de una regla Ry se unen mediante una conjunción se genera un ángulo con arco:





Este sería nuestro árbol de deducción. Este árbol no es necesario, técnicamente, pero es muy buena ayuda para los siguientes pasos:

1. Cálculos de las certezas en los antecedentes de las reglas. Lo haremos de la siguiente forma.
    a. Si los antecedentes se unen mediante conjunción, el CF total es el mínimo de las CF de los antecedentes:

    $$CF(A \wedge B) = \min(CF(A), CF(B))$$

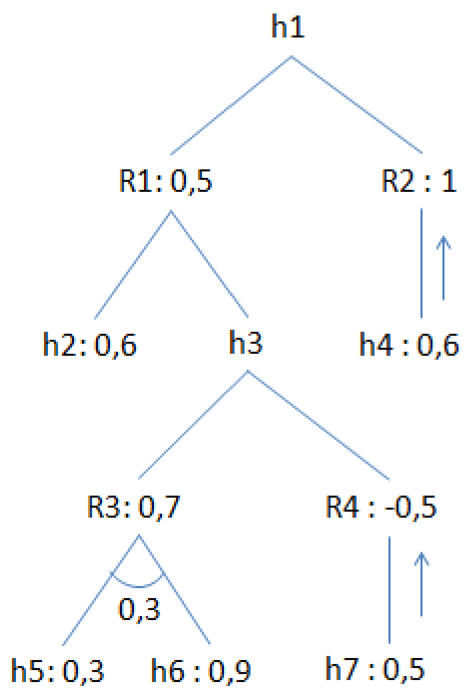
    b. Si los antecedents se unen mediante disyunción, el CF total es el máx. de los CF de los antecedentes:

    $$CF(A \vee B) = \text{máx}(CF(A), CF(B))$$

    c. Si solo hay un antecedente su CF será el CF del antecedente:

    $$CF(A) = CF(A)$$

    En nuestro grafo sería:



Del resto desconocemos su valor.

2. Propagación de certeza por las reglas. Se dan ciertos casos:
    a. CF(antecedente) $\leq$ 0 → La regla no se dispara. Puedo podar el árbol por esta rama.
    b. Caso contrario, $\underbrace{\mathrm{CF(h)}}_{\text{hipótesis}} = \overbrace{\mathrm{CF(e)}}^{\text{CF Antecedentes}} \cdot \underbrace{\mathrm{CF(h,\ e)}}_{\text{CF regla}}$

En nuestro ejemplo:

$$CF\underbrace{4}_{Regla}\underbrace{(h_3)}_{\substack{Hipótesis\\ desconocida}} = CF\underbrace{(h_7)}_{Antecedente} \cdot CF\underbrace{(h_3, h_7)}_{Regla\ 4} = -0{,}25$$

La regla 4 otorga certeza en contra de que h_3 se cumpla.

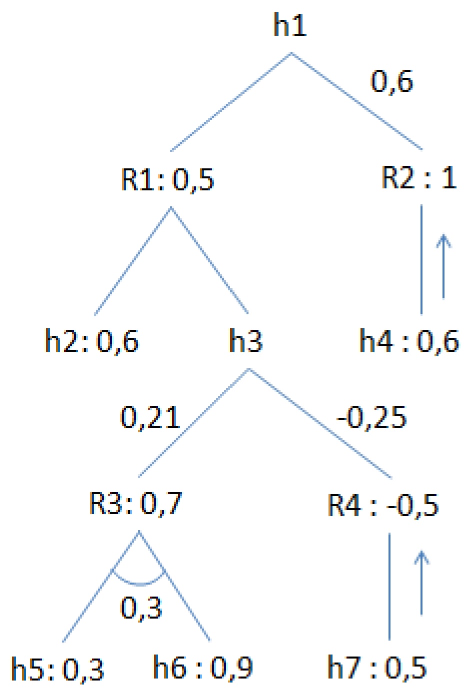
El antecedente de la regla R3 es, como vimos anteriormente:

$$CF_3(h_5 \wedge h_6) = \min\big(CF(h_5), CF(h_6)\big) = 0{,}3$$

Por tanto, la propagación del CF de la R3 es como sigue:

$$CF_3(h_3) = \underbrace{CF(h_5 \wedge h_6)}_{0{,}3} \cdot \underbrace{CF(h_3, h_5 \wedge h_6)}_{0{,}7} = 0{,}21$$

$$CF\ (h\ ) = CF(h\ \ h\ ) \cdot CF(h\ ) = 0{,}6 \cdot 1 = 0{,}6$$



Ahora necesitamos conocer el CF de h_3 para poder continuar.

3. Acumulación de certeza. Se dan tres casos: (consecuentes)
    a. CF_1, $CF_2 > 0$, ambos positivos → $CF = CF_1 + CF_2 - CF_1 \cdot CF_2$ Estos dos últimos nos impiden pasarnos de 1.
    b. CF_1, $CF_2 < 0$, ambos negativos → $CF = CF_1 + CF_2 + CF_1 \cdot CF_2$ Estos dos últimos nos impiden pasarnos de -1.
    c. $CF_1 \cdot CF_2 < 0$, uno negativo y otro positivo → $CF = \frac{CF_1 + CF_2}{1 - \min(|CF_1|, |CF_2|)}$

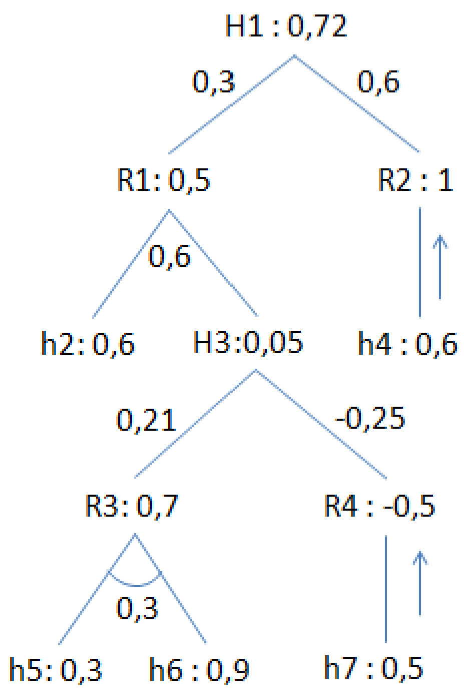
En nuestro ejemplo:

$$CF(h_3) = \frac{CF_3 + CF_4}{1 - \min(|CF_2|, |CF_4|)} = \frac{0{,}21 - 0{,}25}{1 - 0{,}21} = \frac{-0{,}04}{0{,}79} = -0{,}05$$

Entonces ya podemos continuar:

$$CF_1(h_2 \vee h_3) = máx\left(CF(h_2), CF(h_3)\right) = CF(h_2) = 0{,}6$$

$$CF_1(h_1) = \underbrace{CF(h_2 \vee h_3)}_{0{,}6} \cdot \underbrace{CF(h_1, h_2 \vee h_3)}_{0{,}5} = 0{,}3$$



Como hacer este tercer caso con tres antecedentes:

R5: CF = - 0,7 → h_1

$$CF(h_1) = CF_1(h_1) + CF_2(h_1) - CF_1(h_1) \cdot CF_2(h_2) = 0{,}72$$

$$CF(h_1) = \frac{CF_1 + CF_5}{1 - \min(|CF_1|, |CF_5|)} = \frac{0{,}72 - 0{,}7}{1 - 0{,}7} = \frac{0{,}02}{0{,}3} = 0{,}06 \rightarrow No\ podemos\ saberlo$$

Acumulamos positivos dos a dos y negativos dos a dos y luego los acumulamos entre sí dos a dos.

Evidencia independiente

En MYCIN se supone siempre **evidencia independiente.** No se puede utilizar la misma hipótesis formando parte de dos antecedentes de reglas distintas para devolver el mismo consecuente: