



Bases de datos difusas

Pedro Manuel Gómez-Portillo





Resumen



Elevator pitch



En esta presentación se expondrá una visión general de las bases de datos difusas y cómo funcionan, se explicarán los cimientos en los que se apoyan y por último se presenta el lenguaje de consultas difusas FSQL.

Índice

- Introducción
- Bases de datos relacionales (y sus problemas)
- Teoría de bases de datos difusas
- Implementación de las bases de datos difusas
- Fuzzy SQL



Introducción



De dónde salen

El lenguaje es vago

- *Si llueve mucho cogeré el paraguas*
- *Sácalo del horno antes de que esté muy caliente*
- *No le des tan fuerte o lo romperás*

Lofti Zadeh (1964)

The concept of a linguistic variable and its applications to approximate reasoning.



313%

Aumento del número de artículos publicados de 2000 a 2010¹

¹*Patel, A. et al (2013) Application of Fuzzy Logic in Biomedical Informatics.*

Bases de datos relacionales

Y sus problemas

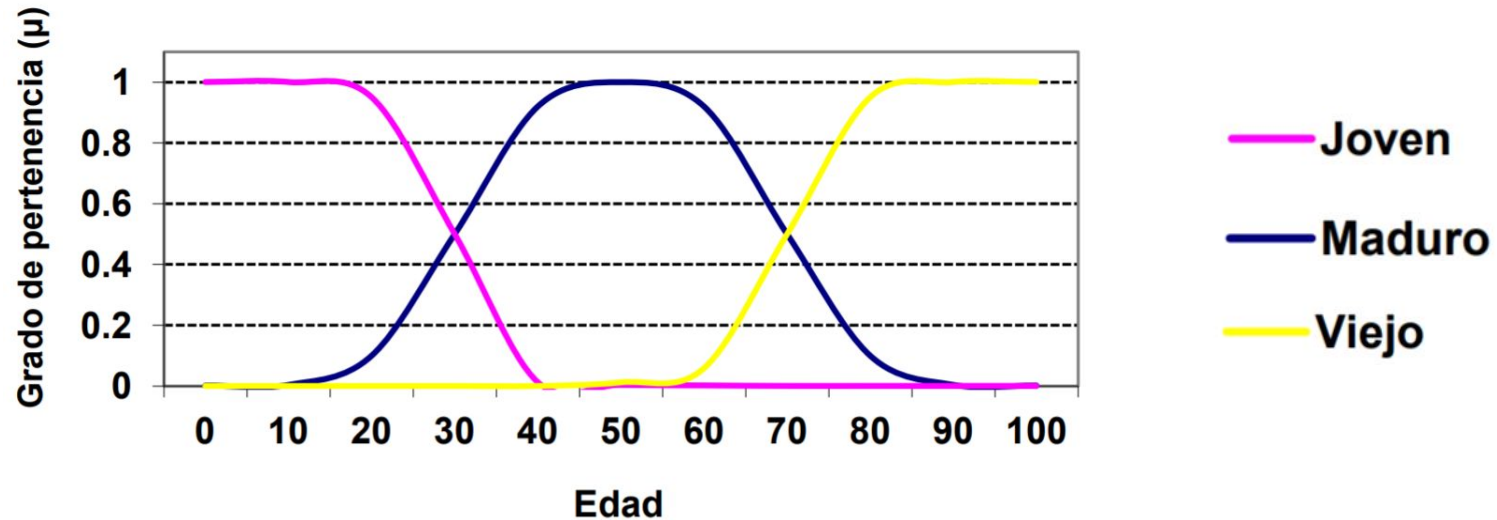
ENCUENTRA A LAS PERSONAS CON EDAD IGUAL A 20 AÑOS

Se obtendrá las personas que pertenecen al grupo

tener 20 años

¿Y si busco las de 19? ¿O las de 21?

ENCUENTRA A LAS PERSONAS **JÓVENES**



Cadenas Lucero, T. (2015) *Sistemas de bases de datos difusas sensibles al contexto*.

Teoría de bases de datos difusas

Conjuntos difusos I

- Un conjunto es una colección de elementos con una propiedad común
- Sin embargo, para un conjunto difuso hay distintos grados de pertenencia

Conjuntos difusos II

$$A = \{ (x, \mu_A) \mid x \in X \}$$

- Sea **X** una colección de objetos cuyos elementos se denominan **x**
- Sea **A** un conjunto difuso en **X**
- Sea **μ_A** la función de pertenencia del conjunto difuso **A**

Funciones de pertenencia

$$\mu_A(x) : X \rightarrow [0,1]$$

- μ_A es la función de pertenencia del conjunto difuso **A**, que asocia cada elemento **x** del conjunto **X** un grado de pertenencia al conjunto difuso en el intervalo **[0,1]**

Variables lingüísticas I

Quíntupla $(N, T(N), X, G, M)$

- N es el nombre de la variable lingüística
- $T(N)$ es el conjunto de términos o valores de N
- X es el Universo de Discurso donde se define $T(N)$
- G es una regla sintáctica que genera los valores de $T(N)$
- M es una regla semántica que asocia cada valor L su significado $M(L)$

Variables lingüísticas II

- Variable lingüística **$N = Edad$**
- Conjunto **$X = [0, 100]$ años**
- Conjunto de términos **$T(N) = \{joven, maduro, viejo\}$**
- Cada término definido por un intervalo difuso en **X**
- Regla sintáctica **$G = \{muy, poco\}$**

Implementación de las bases de datos difusas

Tipos de datos

- Datos precisos
- Datos imprecisos
 - Sobre dominios ordenados
 - Sobre dominios discretos
 - Indefinidos, desconocidos y nulos

Principales modelos

- Modelo relacional difuso básico
- Modelos relacional mediante distribuciones de posibilidad
- **Modelo de unificación mediante relaciones de similitud**

M. unificación por relaciones de simil.

	Rubio	Pelirrojo	Castaño	Moreno
Rubio	1	0.5	0.25	0
Pelirrojo		1	0.5	0.25
Castaño			1	0.5
Moreno				1

Nombre	Color de pelo
Álvaro	Castaño
Jaime	Moreno
Marta	Pelirroja
Ana	Rubia

ENCUENTRA LAS PERSONAS CASTAÑAS CON GRADO 0.5

M. unificación por relaciones de simil.

	Rubio	Pelirrojo	Castaño	Moreno
Rubio	1	0.5	0.25	0
Pelirrojo		1	0.5	0.25
Castaño			1	0.5
Moreno				1

Nombre	Color de pelo
Álvaro	Castaño
Jaime	Moreno
Marta	Pelirroja
Ana	Rubia

ENCUENTRA LAS PERSONAS CASTAÑAS CON GRADO 0.5

Lenguajes lógicos

FSQL

Fuzzy Structured Query Language

- Extensión de SQL
- Diseñado por el Dr. Juan Medina de la UGR en 1998
- Continuado por el Dr. José Galindo de la UMA
- Alternativas como SQLf

Valores lingüísticos

- Variable lingüística

Altura = { alto, medio, bajo }

- Llamada con ***\$alto, \$medio, \$bajo***
- Funciones de pertenencia trapezoidales

Comparadores difusos

Comparador FSQL		Equivalente SQL		Significado
FEQ	F=	EQ	=	Igual que
FDIF	F<>	DIF	<>	Diferente que
FGT	F>	GT	>	Mayor que
FGEQ	F>=	GEQ	>=	Mayor o igual que
MGT	F>>	X		Mucho mayor que
FLT	F<	LT	<	Menor que
FLEQ	F<=	LEQ	<=	Menor o igual que
MLT	F<<	X		Mucho menor que
INCL		X		Incluido en
FINCL		X		Difusamente incluido en

Umbrales I

- Palabra reservada ***THOLD*** x

```
SELECT nombre FROM Personas  
WHERE color_pelo F= $rubio THOLD 0.8  
AND altura FGEQ $alto THOLD 0.8
```

Umbrales II

- $[n, m]$ permite definir un intervalo
- $\$[a, b, c, d]$ permite construir un trapecio
- $\#n$ funciona como *aproximadamente n*

¿Preguntas?

gomezportillo@correo.ugr.es

 [/gomezportillo/fuzzy_databases](https://github.com/gomezportillo/fuzzy_databases)