



BASES DE DATOS DIFUSAS

José Alberto Gómez García

ÍNDICE DE CONTENIDOS

01

Introducción

02

Modelos de bases
de datos difusas

03

Fuzzy SQL



01

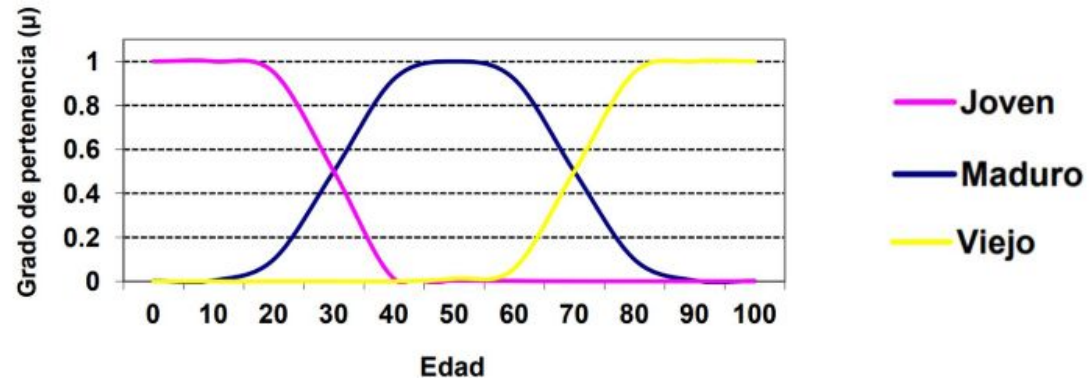
INTRODUCCIÓN

El lenguaje que usamos las personas es vago, impreciso y muchas veces subjetivo.

- Las prácticas de DSS no me salieron muy allá.
- Tenemos muchísimo trabajo.
- Esta semana ha hecho muchísimo frío.

Las bases de datos no permiten tratar bien estos datos que presentan imprecisiones.

¿Cómo busco en SQL gente joven?



Buscaremos unir bases de datos relacionales con teoría de conjuntos difusos de Zadeh.

Objetivos:

- Almacenar imprecisión.
- Operar con información difusa (consultas).



O2

Modelos de
bases de datos
difusas

La implementación de las Fuzzy DB siguen dos caminos:

- Base de datos “crisp” con capacidad de realizar consultas difusas.
- Base de datos capaz de almacenar información difusa y tratarla.

Intentamos construir sobre un SGBD ya existente

A la definición del formato interno de la base de datos difusa y su esquema de implementación se le llama FIRST.

FIRST = Fuzzy Attributes + Fuzzy Metaknowledge Base


¿Con qué tipos de datos trabajan las Fuzzy DB?

- Datos precisos
- Datos imprecisos
 - Sobre dominios continuos (ordenados)
 - Sobre dominios discretos.
 - Desconocidos, indefinidos o nulos.

Fuzzy Metaknowledge Base es el encargado de almacenar los atributos con capacidades difusas y metainformación asociada.

Estas pueden ser etiquetas lingüísticas, distancia para considerar dos valores muy separados, significado del grado (pertenencia, cumplimiento, posibilidad, etc)

Un sistema de base de datos relacional difuso debe:

- Poder representar información difusa.
 - Almacenar el significado de la información difusa.
 - Facilitar un número mínimo de operadores.
 - Satisfacer los requisitos de un modelo relacional.
- 
- The bottom of the slide features decorative abstract shapes. On the left, there is a light pink shape. In the center, there is a dark brown, irregular shape. On the right, there is a light orange shape. These shapes are layered and overlap, creating a modern, artistic background for the footer area.

Existen tres principales tipos de modelos:

- Relacional difuso básico.
- Unificación mediante relaciones de similitud
- Relacional mediante distribuciones de probabilidad.

Relacional difuso básico

Añade un grado de pertenencia difuso en intervalo $[0, 1]$

- A la tupla en general.
- A cada valor de un atributo en la tupla.
- A un conjunto de valores.

Unificación mediante relaciones de similitud

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

Para encontrar personas como mínimo 50% castañas, se observan los valores de la tabla.


- Resultado: Álvaro y Marta

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

Modelo de Buckles-Petry

Relacional mediante distribuciones de probabilidad.

Tienen en común que utilizan distribuciones de probabilidad para modelar la información conocida sobre el valor de un atributo.

The bottom of the slide features three large, overlapping, organic shapes in a muted color palette. From left to right, there is a light pink shape, a dark charcoal grey shape, and a light orange shape. These shapes are positioned at the bottom edge, partially cut off by the frame.

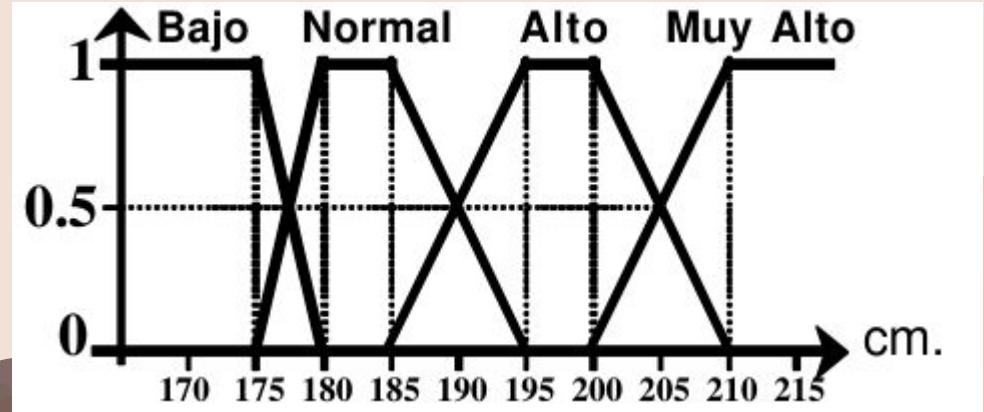
Algunos modelos que hacen uso de probabilidades.

- Umano - Fukami
- Prade - Testemale
- Zemankova - Kandel
- GEFRED (desarrollado por Medina-Pons-Vila; UGR)

UMANO - FUKAMI

Sobre el dominio X de un atributo, la distribución A indica que la posibilidad de que el atributo tome el valor x perteneciente a X es $A(x)$

Si hay que operar con probabilidad se usa el principio de extensión



PRADE – TESTEMALE

Añade un elemento al dominio de todos los atributos, para representar que el atributo no es aplicable.

Información	Modelo Prade-Testemale	Modelo Umamo-Fukami
Sabemos el dato y este es <i>crisp</i> : c	$A(e)=0; A(c)=1;$ $A(x)=0, \forall x \in X, x \neq c;$	$A(x) = \{1/c\};$
Desconocida (pero aplicable)	$A(e)=0;$ $A(x)=1, \forall x \in X;$	Unknown
No aplicable	$A(e)=1; A(x)=0, \forall x \in X;$	Undefined
Ignorancia total	$A(x)=1, \forall x \in X \cup \{e\};$	Null
Rango [m,n]	$A(x)=0, \forall x \notin [m,n] \text{ o } x = e;$ $A(x)=1, \forall x \in [m,n];$	$A(x)=0, \forall x \notin [m,n];$ $A(x)=1, \forall x \in [m,n];$
Distribución de Posibilidad B	$A(e)=0;$ $A(x)=B(x), \forall x \in X;$	$A(x)=B(x), \forall x \in X;$
Posibilidad de que no sea aplicable es λ y si lo es vale B .	$A(e)=\lambda;$ $A(x)=B(x), \forall x \in X;$	No representable

ZEMANKOVA - KANDEL

- Utiliza las medidas de posibilidad y certeza; en lugar de posibilidad y necesidad.
- La medida de certeza no tiene una interpretación clara.
- Define comparadores difusos poco intuitivos.

- **Medida de Posibilidad:** $p_A(F) = \sup_{x \in X} \{F(x) \bullet A_i(x)\};$
- **Medida de Certeza:** $c_A(F) = \max\{0, \inf_{x \in X} \{F(x) \bullet A_i(x)\}\};$

GEFRED

- Es un compendio de los anteriores.
- Permite dominios de cualquier tipo.
- Se tiene un “atributo de compatibilidad” que expresa en qué medida ha satisfecho la operación solicitada.
- Propone un comparador difuso genérico, pero sin definir.
- No propone otros comparadores difusos.

Galindo propone algunos para ciertas situaciones.

Estos modelos se diferencian en:

- Qué mide la distribución
(grado de pertenencia, cumplimiento, importancia, etc)
- Cómo gestionar valores desconocidos o no aplicables.
- Si permiten sólo valores fuzzy numéricos o no,
- ¿Qué hacer con las claves primarias y externas?
- Operadores de comparación que proporcionan,

¿Qué hacer con las claves primarias y externas?

- Modelos como Prade-Testemale usan un umbral de igualdad entre distribuciones de probabilidad.
 - Si dos valores son iguales, ¿cual almacenamos?
 - Si uno incluye a otro ¿qué pasa con contradicciones?
- Modelos como GEFRED exigen atributo “crisp” o difusos con criterio de redundancia (como ser exactamente iguales)



O3

FUZZY SQL

FSQL (Fuzzy Structured Query Language)

- Extensión de SQL (para SGBD Oracle)
- Diseñado por Juan Medina (UGR) en 1998
- Ampliado y continuado por José Galindo (UMA)

No es la única opción, hay otras como SQLf

Características de FSQL

- Las variables lingüísticas se definen como:
nombre = {valor1, valor2, valor3, ... }
- Se llama a estas variables con \$valor
- Las funciones de pertenencia son trapezoidales.
- [n, m] permite definir intervalos
- \$[a,b,c,d.] permite construir un trapecio.
- #n funciona como “aproximadamente n”
- UNKNOWN, UNDEFINED y NULL son constantes difusas

Operadores en FSQL

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

Ejemplo de consulta en FSQL

```
SELECT nombre FROM Personas WHERE  
Pelo FEQ $moreno THOLD 0.5 AND  
Edad FGT $Joven THOLD 0.8
```

- THOLD indica el valor mínimo de la función de pertenencia para satisfacer la condición



O4

CONCLUSIONES

- Las fuzzy DB permiten tratar los datos de manera más cercana al lenguaje natural.
- Se han analizado algunos modelos de fuzzy DB.
- Se ha abordado el lenguaje FSQL, el cual necesita de muchos parámetros.
- No hay una solución comercial potente, no parece despertar gran interés.
- ¿No relacional? ¿Grafos?

GRACIAS POR SU ATENCIÓN

¿ALGUNA PREGUNTA?

CREDITS: This presentation template was created by [Slidesgo](#), including icons by [Flaticon](#), infographics & images by [Freepik](#).

Please keep this slide for attribution



Bibliografía

- Base de datos relacionales difusas. Galindo. J
www.lcc.uma.es/~ppgg/FSS/FSS9.pdf
- Software FSQL. José Galindo
www.lcc.uma.es/~ppgg/FSQL
- TFG sobre bases de datos relacionales difusas. Sanchis. D
repositori.uji.es/xmlui/bitstream/handle/10234/149667/TFG_2014_SanchisMinguezD.pdf
- Fuzzy relational algebra for possibility distributions fuzzy relational model of fuzzy data.
Umano - Fukami
- A fuzzy representation for relational databases. Fuzzy Sets and Systems. Buckles - Petry.
- Sistemas de bases de datos difusas sensibles al contexto Cadenas Lucero, T.
- GEFRED. A Generalized Model of Fuzzy Relational Databases. Vila A., Medina, J. & Pons, O.
- New Characteristics in FSQL, a Fuzzy SQL for Fuzzy Databases. Galindo. J