ESTUDIO COMPARATIVO DE DATOS NÍTIDOS Y DIFUSOS EN ESTRUCTURAS DE DATOS

Fulbia Torres,
Universidad Nacional Experimental Politécnica "Antonio José de Sucre"
ftorres@unexpo.edu.ve
Ennodio Torres
Universidad Centroccidental Lisandro Alvarado
ennodiotorres@gmail.com

Resumen- El objetivo de la presente investigación es hacer un estudio comparativo de datos nítidos y difusos en estructuras de datos. La metodología utilizada consiste de un estudio documental en donde se observa y reflexiona sistemáticamente sobre realidades (teóricas o no) usando para ello diferentes tipos de documentos, con el fin de evaluar ambos tipos de datos. En los resultados se muestra todo una teoría sobre estructuras de datos, datos, tipos de datos, lógica difusa, teoría de conjuntos nítidos, sus operaciones, teoría de conjuntos difusos con sus operaciones. Una de las conclusiones más importantes es que los resultados mostraron que ambos tipos de datos son buenos, sin embargo para aplicaciones con esquemas complejos los tipos de datos difusos son los más adecuados.

Palabras claves: Estructuras de datos/ datos nítidos/ datos difusos.

COMPARATIVE STUDY DATA CRISP AND FUZZY DATA STRUCTURES

Abstract- The objet of the present investigation was to make a comparative study of data crisp and diffuse in data structures. The methodology consisted of a study in which notes and think systematically about realities (theoretical or not) taking into consideration different types of documents, in order to evaluate both types of data, the results are displayed in a theory on all structures data, data, data types, fuzzy logic, set theory sharp, its operations, fuzzy set theory with its operations. One of the most important findings is that the results showed that both types of data are good, however for applications with complex schemes fuzzy data types are best suited.

Key words- Data structures/ data crisp/ fuzzy data.

Este manuscrito fue recibido en Barquisimeto el 24/03/2010 y aprobado para su publicación13/05/2010. La MSc. Ing. Fulbia Torres es profesora de categoría asociado del Departamento de Ingeniería Electrónica de la Universidad Nacional Experimental Politécnica "Antonio José de Sucre" (UNEXPO). El Dr. Ennodio Torres Dr. En Ciencias Matemáticas es profesor jubilado de la Universidad Centroccidental "Lisandro Alvarado" (UCLA). Teléfono/fax +58 251 2591601. Correo electrónico: ennodiotorres@gmail.com

1. INTRODUCIÓN

Hoy en día Internet, los sistemas operativos, *DataWarehouse*, sistemas cliente-servidor, simuladores y otros han pasado a ser parte de la vida cotidiana. Éstos siempre se han considerado parte del campo de la informática la cual se caracteriza por su permanente dinámica. Éstos, tienen un común denominador, requieren el uso intensivo de estructuras de datos para almacenar, manipular y organizar la información con la que trabajan, Martínez y Quiroga [1]. Las estructuras de datos determinan la conexión lógica entre los datos, afectan el procesamiento físico de los mismos y es un elemento principal que determina el rendimiento de los programas.

Según Carrano y Prichard [2] las estructuras de datos se implementan a través de los lenguajes de programación y son un modelo que caracteriza y permite almacenar y utilizar una determinada organización de datos. En este sentido Loomis [3] señala que los datos pueden considerarse como materia prima, son agregados y sumados en diversas maneras útiles para constituirse en información. Existen diversidad de datos, entre los cuales están los datos nítidos (*data crisp*) y los datos difusos (*fuzzy data*), los mismos, tienen características propias y pueden ser almacenados y manipulados en las estructuras de datos.

El problema es que en la vida real existen hechos que no se puede definir como totalmente verdaderos o totalmente falsos, sino que tienen un grado de verdad o falsedad que puede variar de 0 a 1; la lógica clásica no es la más adecuada para tratar este tipo de razonamientos ya que excluye por completo una tercera posibilidad (o más) entre estos dos valores. La Lógica Difusa en cambio, permite utilizar conceptos relativos de la realidad, definiendo grados variables de pertenencia y siguiendo patrones de razonamiento similares a los del pensamiento humano (Kosko, 1995). La teoría de los conjuntos difusos, la lógica difusa constituyen modelos que resultan especialmente útiles para tratar con la incertidumbre de manera más "natural" y más "humana" que la lógica y la teoría de conjuntos clásicas. Los sistemas extraídos de la lógica clásica presentan las dificultades de la rigidez y la bivalencia, y resultan, por ello, inservibles para expresar la ambigüedad del significado que se da en el lenguaje natural, Por lo anterior, el presente trabajo se centra en la comparación de los datos mencionados, para conocer las características de cada uno y el ámbito de aplicación y utilidad de los mismos. En este orden de ideas, se plantea como objetivo de la presente investigación hacer un estudio comparativo de datos nítidos y difusos en estructuras de datos. La metodología utilizada responde a una investigación documental en donde se observa y reflexiona sistemáticamente sobre realidades (teóricas o no) usando para ello diferentes tipos de documentos, con el fin de evaluar ambos tipos de datos. Por lo que, el desarrollo del presente trabajo estuvo basado en la búsqueda, recuperación, análisis, crítica e interpretación de datos secundarios, es decir, los obtenidos y registrados por otros investigadores.

El presente trabajo está estructurado en cuatro(4) partes: la introducción que presenta una panorámica del problema, el propósito del trabajo y la metodología utilizada; el desarrollo del trabajo subdividido en la metodología utilizada para realizar la investigación, los resultados que muestran todo una teoría sobre estructuras de datos, datos, tipos de datos, lógica difusa, teoría de conjuntos nítidos con sus operaciones, teoría de conjuntos difusos con sus operaciones, y una aplicación utilizando base de datos relacionales donde se reflejan cada uno de los conceptos investigados y la discusión de resultados; las conclusiones del estudio y por último las referencias bibliográficas consultadas que dan sustento teórico a la investigación.

2. DESARROLLO

2.1. Metodología

El presente estudio responde a una investigación documental [4] donde se observa y reflexiona sistemáticamente sobre realidades (teóricas o no) usando para ello diferentes tipos de documentos. Dicha investigación indaga, interpreta, presenta datos e informaciones sobre un tema determinado de cualquier área de conocimiento, utilizando para ello, una metódica de análisis; teniendo como finalidad obtener resultados que pudiese ser base para el desarrollo de la producción de conocimientos. Por lo que, el desarrollo del presente trabajo estuvo basado en la búsqueda, recuperación, análisis, crítica e interpretación de datos secundarios, es decir, los obtenidos y registrados por otros investigadores. Se procede a elaborar fichas bibliográficas con los datos básicos de cada documento encontrado (el autor o los autores, la fecha, el título, los datos de publicación si fuere el caso, etc.). Una vez hecho esto, se realiza la lectura del material recopilado, para finalmente confrontar ambos datos y del cotejo de los mismos determinar las características propias de cada uno.

3. RESULTADOS

3.1 Algunas definiciones

Según Marzal [5] entre el nivel de los tipos de datos y el nivel de los tipos abstractos nos encontramos con las llamadas estructuras de datos. Las **estructuras de datos** son conjuntos de variables, quizás de tipos distintos, relacionadas (conectadas) entre sí de diversas formas y con un conjunto de operaciones definidas sobre dichas estructuras. En este sentido, Carrano y Prichard [2] señalan que las estructuras de datos se implementan a través de los lenguajes de programación, que son un modelo que permiten almacenar y utilizar una determinada organización de datos.

Marzal [5] expone algunos ejemplos de estructuras de datos que podemos encontrar en muchos ámbitos, desde las matemáticas (estructuras algebraicas como los grupos, anillos, cuerpos; o estructuras discretas como los grafos, árboles, autómatas) hasta el mundo de los negocios (estructura de una empresa). Los elementos de una estructura de datos dependen del lenguaje de programación a través de los tipos de datos que los definen. Sin embargo, la estructura en sí, que está definida por el tipo de relación entre sus elementos, no depende del lenguaje de programación empleado para el procesamiento de sus datos. La estructura de datos se caracteriza por el tipo de sus elementos, las relaciones definidas sobre tales elementos y las operaciones de las cuales está dotada esa estructura.

Para Martínez y Quiroga [1] una estructura de datos es cualquier colección o grupo de datos organizados de tal forma que tengan asociados un conjunto de operaciones para poder manipularlos. Al nivel de las estructuras de datos, ya no tienen relevancia las operaciones que se puedan realizar sobre los componentes individuales de la misma, solamente la tienen las operaciones que implican a la estructura globalmente. Al respecto, Sisa [6] describe las operaciones sobre una estructura de datos, que se muestran en la Tabla 1

Existen algunos aspectos fundamentales para el conocimiento de las estructuras de datos, como son las funciones básicas que permiten su creación, la forma de acceso y su destrucción, las cuales se presentan a continuación en la Tabla 2

Tal como afirma Sisa [6] se pueden considerar las estructuras de datos desde el punto de vista físico y desde el punto de vista lógico. La estructura de datos, desde el punto de vista físico, está enfocada básicamente a la manera como queda almacenada la información y desde el punto de vista lógico, enfoca el aspecto de las relaciones que existen entre los elementos.

Tabla 1. Operaciones que se realizan sobre una Estructura de Datos**.**

OPERACIÓN	**DEFINICIÓN**
Navegar por la estructura	Realiza un recorrido por la estructura con el propósito de recuperar la información almacenada.
Búsqueda	Determina si un elemento se encuentra o no en la estructura.
Consulta de la información	Obtiene información de uno o más elementos de la estructura.
Copia parcial o total	Adquiere total o parcialmente una estructura con características similares a la original.
Prueba	Verifica si uno o varios elementos cumplen con determinadas condiciones.
Modificación	Varía parcial o totalmente el contenido de la información de los elementos de la estructura.
Eliminación	Suprime elementos de la estructura.
Imprimir	Escribe la información contenida en la estructura.

Tabla 2. Funciones básicas para la creación, acceso y destrucción de las Estructuras de Datos.

FUNCIÓN	**DEFINICIÓN**
Función constructora	Crear la estructura, es decir, definir las características, la delimitación, las relaciones y asignan el espacio correspondiente, dejando a disposición del usuario la estructura para que proceda a colocar la información.
♦ *Función para acceso*	Facilitar la llegada de un elemento perteneciente a la estructura.
Función destructora	Devolver al sistema los recursos asignados a la estructura de datos para que queden a disposición de éste.

La elección de una estructura de datos adecuada es importantísima para resolver con éxito un problema. Por eso hay tantos tipos de estructuras. Los vectores y las matrices se adaptan muy bien a muchísimos problemas, pero para otros son completamente inadecuados. Y ahí es donde entran en juego las estructuras de datos: pilas, colas, árboles, grafos, etc. Los árboles son imprescindibles en muchos problemas de inteligencia artificial o, en general, de búsqueda exhaustiva de soluciones.

Del mismo modo, las pilas, las colas y las otras estructuras se adaptan como guantes a determinados problemas tan dispares como, por ejemplo, los analizadores sintácticos o el reparto de recursos en el sistema operativo

Del mismo modo, las pilas, las colas y las otras estructuras se adaptan como guantes a determinados problemas tan dispares como, por ejemplo, los analizadores sintácticos o el reparto de recursos en el sistema operativo.

Por ser las estructuras de datos fundamentales para el procesamiento de información, hay que considerar dos interrogantes: ¿qué son los datos? y ¿cómo apoyan a este procesamiento?. Sobre el primer aspecto, Millán [7] dice que la palabra *dato* proviene del latín *datum*, forma del verbo *dare*, "dar", que significa "lo que es dado". En inglés se usaba el latinismo directo *datum*, plural *data*, y a finales del XIX ya tenía el sentido (como en las otras lenguas cultas) de "hechos numéricos reunidos para referencia o información", como recoge el *Oxford English Dictionary*. Pero con el desarrollo de herramientas automáticas, en la década de 1940, la palabra se especializa en el sentido de "Representación de una información de manera adecuada para su tratamiento por un ordenador" (definición actual de la Academia Española). Para los autores Ferrís y Albert [8], los datos son las propiedades o atributos (cualidades o cantidades) asociados a hechos u objetos y que son procesados por el ordenador. El tipo de datos, en el contexto de un lenguaje de programación, define el conjunto de valores que una determinada variable puede tomar, así como las operaciones básicas sobre dicho conjunto, es decir, definen cómo se representa la información y cómo se interpreta. Los tipos de datos pueden variar de un lenguaje de programación a otro, tanto los tipos simples como los mecanismos para crear tipos compuestos. Los tipos de datos constituyen un primer nivel de abstracción, ya que no se tiene en cuenta cómo se representa realmente la información sobre la memoria de la máquina, ni cómo se manipula.

Tipos de datos comunes son: números enteros, números de coma flotante (decimales), cadenas alfanuméricas, fechas, horas, colores, coches, valores lógicos, u otros datos que se nos ocurra. En general, los datos por su naturaleza pueden ser: cualitativos, nítidos, difusos. En este sentido, para Urrutia et. al. [9] los datos nítidos (numéricos, alfanuméricos, binarios) no permiten describir fenómenos que manifiesten cierta imprecisión y/o incertidumbre, tanto en su representación como en su consulta; pero existen otro tipo de datos, que contienen incertidumbre o imprecisión en su información, denominados *datos difusos*, pues están asociados a la teoría de conjuntos difusos, la cual formaliza conceptos vagos o difusos que las personas manejan de forma cotidiana y natural.

Lógica Difusa

Al respecto, Torres [10] presenta los contenidos que siguen.

Definiciones

Una *proposición difusa p* es un enunciado cuyo grado de verdad es un número real comprendido entre el 0 y el 1, que denotaremos con la expresión $g(p)$. Los grados de verdad serán expresados en términos absolutos mediante números reales o en términos relativos mediante porcentajes: $0 \leq g(p) \leq 1$ $0\% \leq g(p) \leq 100\%$.

Una *lógica difusa* es una pareja (U, g), donde U representa el universo del discurso de las proposiciones difusas, y $g : U \rightarrow \mathbf{I}$ es una función real de variable proposicional, siendo **I** el intervalo cerrado [0, 1]. En tal caso, decimos que g es la *función veritativa* de la lógica difusa (U, g).

Una lógica difusa (U, g) es *finita* si el rango $g(U)$ es un conjunto finito. En tal caso, a cada proposición $p \in U$, se asocia un grado de verdad

$g(p) \in \{0, 1/(n-1), 2/(n-1), 3/(n-1), \ldots, (n-2)/(n-1), 1\}$,

donde n es un número natural mayor o igual que 3.

Algunas lógicas finitas

Tabla 3. Lógicas finitas con su rango.

Lógica finita (U, g)	**Rango** $g(U)$
Bivalente	{0, 1}
Trivalente	{0, ½, 1}
Tetravalente	{0, ⅓, ⅔, 1}
Pentavalente	{0, ¼, ½, ¾, 1}

Diseñamos ejemplos de lógicas finitas teniendo a disposición en cada caso:

- el número irracional π;
- la sucesión de números racionales que constituyen valores aproximados de π;
- el intervalo cerrado $[3, \pi]$ como dominio de aproximación; y
- el orden de aproximación requerido.

Se puede apreciar en la tabla 4.

Tabla 4. Ejemplos de lógicas finitas utilizando el número irracional π

Proposición *p*	**Grado de verdad *g*(*p*)**
$\pi = 3$	0
$\pi = 3{,}1$	½
$\pi = 3{,}14$	1
Dominio de aproximación: [3, π]	**Orden de aproximación: 0,01**
$\pi = 3$	0
$\pi = 3{,}1$	⅓
$\pi = 3{,}14$	⅔
$\pi = 3{,}141$	1
Dominio de aproximación: [3, π]	**Orden de aproximación: 0,001**
$\pi = 3$	0
$\pi = 3{,}1$	¼
$\pi = 3{,}14$	½
$\pi = 3{,}141$	¾
$\pi = 3{,}1415$	1
Dominio de aproximación: [3, π]	**Orden de aproximación: 0,0001**

Observe que el enunciado $\pi = 3{,}1$ tiene un grado de verdad igual a ½ en el primer ejemplo, ⅓ en el segundo y ¼ en el tercero; lo cual pone de manifiesto el contexto donde se realiza la aproximación al número irracional π.

Operaciones en lógica difusa estándar. Dichas operaciones se muestran en la tabla V.

Algunas propiedades de la lógica difusa estándar

Sean *p* y *q* dos proposiciones en la lógica difusa estándar (*U, g*). Decimos que *p* y *q* son *equivalentes* si y sólo si los grados de verdad de ambas proposiciones son iguales. En lenguaje simbólico:

$p \equiv q \iff g(p) = g(q)$.

Las propiedades se presentan en la tabla 8.

Tabla 5 Operaciones en lógica difusa estándar

Operación	**Definición**
Negación	La *negación* de una proposición *p* es otra proposición $\sim p$, cuyo grado de verdad está definido por $g(\sim p) = 1 - g(p)$.
Conjunción	La *conjunción* de dos proposiciones *p* y *q* es una proposición compuesta $p \wedge q$, cuyo grado de verdad está definido por $g(p \wedge q) = mín\ \{ g(p), g(q)\}$.
Disyunción	La *disyunción* de dos proposiciones *p* y *q* es otra proposición $p \vee q$, cuyo grado de verdad está definido por $g(p \vee q) = máx\ \{ g(p), g(q)\}$.
Implicación	La *implicación* de dos proposiciones *p* y *q* es otra proposición $p \Rightarrow q$, cuyo grado de verdad está definido por $g(p \Rightarrow q) = mín\ \{1, 1 - g(p) + g(q)\}$.
Bicondicional	El *bicondicional* de dos proposiciones *p* y *q* es otra proposición $p \Leftrightarrow q$, cuyo grado de verdad está definido por $g(p \Leftrightarrow q) = 1 - \mid g(p) - g(q) \mid$.

Ejemplo de grados de verdad de las operaciones en una lógica pentavalente, se puede apreciar en las tablas 6 y 7.

Tabla 6. Operación de Negación

p	0	¼	½	¾	1
$\sim p$	1	¾	½	¼	0

Tabla 7. Operaciones de Conjunción, Disyunción, Implicación y Bicondicional

$\wedge$	0	¼	½	¾	1		$\vee$	0	¼	½	¾	1
0	0	0	0	0	0		0	0	¼	½	¾	1
¼	0	¼	¼	¼	¼		¼	¼	¼	½	¾	1
½	0	¼	½	½	½		½	½	½	½	¾	1
¾	0	¼	½	¾	¾		¾	¾	¾	¾	¾	1
1	0	¼	½	¾	1		1	1	1	1	1	1
$\Rightarrow$	0	¼	½	¾	1		$\Leftrightarrow$	0	¼	½	¾	1
0	1	1	1	1	1		0	1	¾	½	¼	0
¼	¾	1	1	1	1		¼	¾	1	¾	½	¼
½	½	¾	1	1	1		½	½	¾	1	¾	½
¾	¼	½	¾	1	1		¾	¼	½	¾	1	¾
1	0	¼	½	¾	1		1	0	¼	½	¾	1

Las proposiciones **1** y **0** se denominan *proposición segura* y *proposición imposible*, respectivamente. Las proposiciones equivalentes a la proposición segura, son llamadas *tautologías*; las equivalentes a la proposición imposible, *contradicciones*. Las leyes de contradicción y del tercio excluso de la lógica bivalentes no se cumplen en la lógica finita n–valente, cuando $n \geq 3$; en otras palabras, las proposiciones compuestas "$p \wedge \sim p$" y "$p \vee \sim p$" no son, respectivamente, ni una contradicción ni una tautología. En efecto, los siguientes contraejemplos, dados en una lógica pentavalente estándar, muestran la validez de tales enunciados. Los mismos se presentan a continuación en la tabla 9.

Tabla 8. Propiedades de la lógica difusa estándar con su representación en lenguaje simbólico

Propiedad	**Lenguaje simbólico**
Conmutatividad de la $\wedge$	$(p \wedge q) \equiv (q \wedge p)$
Asociatividad de la $\wedge$	$((p \wedge q) \wedge r) \equiv (p \wedge (q \wedge r))$
Conmutatividad de la $\vee$	$(p \vee q) \equiv (q \vee p)$
Asociatividad de la $\vee$	$((p \vee q) \vee r) \equiv (p \vee (q \vee r))$
Absorción $(\wedge,\vee)$	$(p \wedge (p \vee q)) \equiv p$
Absorción $(\vee, \wedge)$	$(p \vee (p \wedge q)) \equiv p$
Distributividad $(\wedge,\vee)$	$(p \wedge (q \vee r)) \equiv ((p \wedge q) \vee (p \wedge r))$
Distributividad $(\vee, \wedge)$	$(p \vee (q \wedge r)) \equiv ((p \vee q) \wedge (p \vee r))$
1ª Ley de De Morgan	$[\sim (p \wedge q)] \equiv (\sim p \vee \sim q)$
2ª Ley de De Morgan	$[\sim (p \vee q)] \equiv (\sim p \wedge \sim q)$
Elemento neutro de la $\wedge$	$(p \wedge \mathbf{1}) \equiv p$
Elemento neutro de la $\vee$	$(p \vee \mathbf{0}) \equiv p$

Tabla 9. Contraejemplos dados en la lógica pentavalente estándar**.**

p	$\wedge$	$\sim$	p	$\Leftrightarrow$	**0**		p	$\vee$	$\sim$	p	$\Leftrightarrow$	**1**
0	0	1	0	1	0		0	1	1	0	1	1
¼	¼	¾	¼	¾	0		¼	¾	¾	¼	¾	1
½	½	½	½	½	0		½	½	½	½	½	1
¾	¼	¼	¾	¼	0		¾	¾	¼	¾	¾	1
1	0	0	1	1	0		1	1	0	1	1	1

Predicados unitarios difusos

Dado un conjunto no vacío X y sea (U, g) la lógica estándar, donde el conjunto U representa el universo del discurso de las proposiciones difusas, y la función real de variable proposicional $g: U \to [0, 1]$ denota su función veritativa. Llamamos *predicado unitario difuso* en el universo X a una función proposicional $P : X \to U$, es decir, a una correspondencia que a cada elemento $x \in X$, asocia una única proposición $P(x) \in U$. La proposición $P(x)$ es de la forma sujeto–predicado "x es P", donde x es la variable sujeto y P es la variable predicado.

Ejemplos de predicados unitarios difusos

- La oración "x es un número natural mucho menor que 5", de variable individual x y de constante individual 5, es la especificación de un predicado unitario difusos $P : \{0, 1, 2, 3,\ldots\} \to U$, donde U es el universo del discurso de las proposiciones difusas. Los grados de verdad de las proposiciones $P(0)$, $P(1)$, $P(2)$, $P(3)$ y $P(4)$, en la lógica pentavalente estándar, son 1, ¾, ½, ¼ y 0, respectivamente. Estos grados de verdad forman una sucesión finita y decreciente de números racionales.

- El enunciado "x es un número racional aproximadamente igual al número irracional π", de variable individual x y de constante individual π, es la especificación de un predicado unitario difuso P del conjunto **Q** de todos los números racionales en el universo del discurso de las proposiciones difusas U. Tal especificación es denotada a través de la expresión $P(x)$: "$x \approx \pi$". Los grados de verdad de las proposiciones P (3), P (3,1), P (3,14), P (3,141), P (3,1415), ..., constituyen una sucesión, infinita y creciente, de números racionales entre el 0 y el 1.

Teoría de Conjuntos Difusos

Zadeh [11] en 1965 definió el concepto de conjunto difuso, basándose en la idea de que existen conjuntos en los que no está claramente determinado si un elemento pertenece o no al conjunto. A veces, un elemento pertenece al conjunto con cierto grado, llamado grado de pertenencia.

Definiciones

Torres [12] introduce el concepto de conjunto difuso en términos de predicados unitarios difusos.

Dado un predicado unitario difuso $P : X \rightarrow U$ y sea (U, g) la lógica estándar, donde el conjunto U representa el universo del discurso de las proposiciones difusas, y la función real de variable proposicional $g: U \rightarrow [0, 1]$ denota su función veritativa. En este contexto, el proceso para construir en el universo X un *conjunto difuso* A puede ser realizado mediante el *método de comprensión*, como sigue: el conjunto difuso A está formado por todos los elementos del universo X, y a cada uno de ellos se le asocia la condición numérica impuesta por la función compuesta $g \circ P : X \rightarrow [0, 1]$, la cual se denomina *función pertenencia* del conjunto difuso A y se denota μ_A. El número real $\mu_A(x)$ representa el grado de pertenencia del elemento $x \in X$ al conjunto difuso A. En tal caso, utilizamos la siguiente notación: $A = \{x|\mu_A(x) : (\forall x \in X)P(x)\}$. Esta expresión se lee "A es el conjunto difuso de elemento genérico $x \in X$, tal que el número $\mu_A(x)$ representa el grado de pertenencia del elemento x al conjunto A".

Por otra parte, el proceso para construir un conjunto difuso finito o infinito numerable puede ser realizado mediante el *método de extensión*, utilizando el siguiente patrón organizativo: enumeramos sus elementos con sus respectivos grados de pertenecia, separándolos por comas y encerrándolos dentro de una llave. Explícitamente, se presenta en la tabla X.

Tabla 10. Conjunto difuso finito e infinito numerable realizado mediante el método de extensión**.**

Conjunto difuso finito	Conjunto difuso infinito numerable
$A = \{a_1|\mu_A(a_1),\ a_2|\mu_A(a_2),\ a_3|\mu_A(a_3), \ldots,\ a_n|\mu_A(a_n)\}$.	$B = \{b_1|\mu_B(b_1),\ b_2|\mu_B(b_2),\ b_3|\mu_B(b_3), \ldots\}$.

Blanco [13] nos dice que los conceptos imprecisos se representan mediante el uso de variables lingüísticas. La interpretación de cada valor lingüístico de una variable se hace en este contexto a través de un conjunto difuso, el cual expresa una restricción sobre los valores que puede tener la variable base que se está expresando. Por ejemplo, el término difuso *joven* puede definirse mediante el conjunto difuso siguiente como se muestra en la tabla 11.

Tabla 11. Conjunto difuso para el término joven**.**

Edad	Grado de pertenencia
≤ 25	1.0
30	0.8
35	0.6
40	0.4
45	0.2
≥ 50	0

Es decir, la función de pertenencia del conjunto difuso *joven* viene dada por:

$\mu_A(x) = 1$ si $x \leq 25$, $\mu_A(30) = 0.8$, ..., $\mu_A(x) = 0$ si $x \geq 50$.

Que podemos representar en la gráfica mostrada en la figura 1.

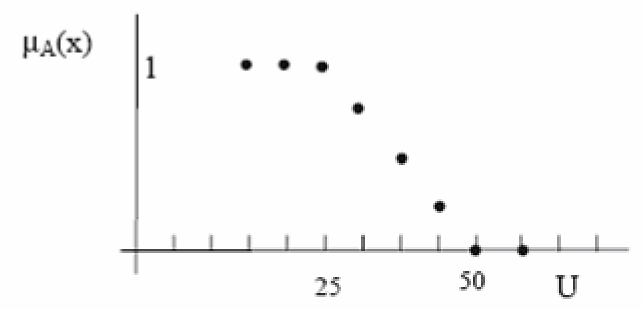


Figura 1 Función de Pertenencia de Conjunto Difuso Joven

Si el universo del discurso *U* es continuo, tendremos funciones de pertenencia continuas, como puede observarse en la figura 2.

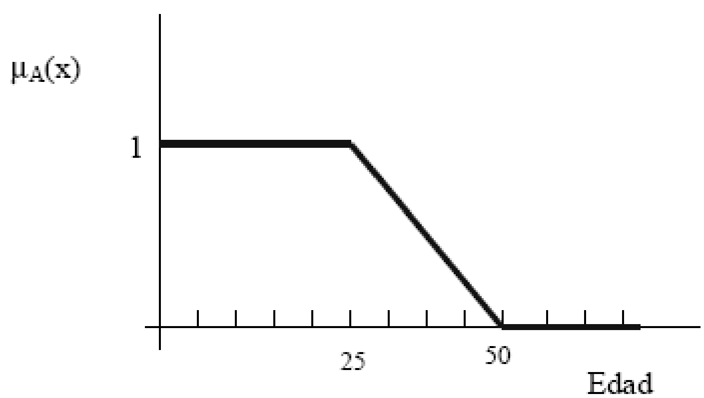


Figura 2 Función de Pertenencia de Joven si U es Continuo

Relaciones entre conjuntos difusos

Al respecto, Torres [12] define las relaciones de igualdad y de inclusión como se muestra en la tabla XII.

Tabla 12. Relaciones de igualdad y de inclusión entre conjuntos difusos.

Relación	Definición
Igualdad	El conjunto difuso A es *igual* al conjunto difuso B si y sólo si para cada elemento $x \in X$, el número $\mu_A(x)$ es igual al número $\mu_B(x)$. Se denota $A = B$. En forma compacta: $A = B \Leftrightarrow (\forall x \in X)[\mu_A(x) = \mu_B(x)]$.
Inclusión	El conjunto difuso A es un *subconjunto* del conjunto difuso B si y sólo si para todo $x \in X$, el número $\mu_A(x)$ es menor o igual al número $\mu_B(x)$. Se denota $A \subset B$. En forma compacta: $A \subset B \Leftrightarrow (\forall x \in X)[\mu_A(x) \leq \mu_B(x)]$.

Operaciones con conjuntos difusos

Como se ha visto, la función de pertenencia es la componente fundamental de un conjunto difuso; de ahí que las operaciones con tales conjuntos se definen a través de dicha función, se pueden apreciar en la tabla 13. Según Torres [12]:

Tabla 13. Operaciones con conjuntos difusos.

Operación	Definición
Complemento	El *complemento* de un conjunto difuso A es un segundo conjunto difuso B de elemento genérico $x \in X$, tal que el número $\mu_B(x)$ es igual al número $1 - \mu_A(x)$. Se denota A^c. En forma compacta: $A^c = B \Leftrightarrow (\forall x \in X)[\mu_B(x) = 1 - \mu_A(x)]$.
Intersección	La *intersección* de dos conjuntos difusos A y B es otro conjunto difuso C de elemento genérico $x \in X$, tal que $\mu_C(x)$ es igual al mínimo entre los números $\mu_A(x)$ y $\mu_B(x)$. Se denota $A \cap B$. En forma compacta: $A \cap B = C \Leftrightarrow (\forall x \in X)\ \mu_C(x) = mín\ \{\mu_A(x), \mu_B(x)\}$.
Unión	La *unión* de dos conjuntos difusos A y B es otro conjunto difuso C de elemento genérico $x \in X$, tal que $\mu_C(x)$ es igual al máximo entre los números $\mu_A(x)$ y $\mu_B(x)$. Se denota $A \cup B$. En forma compacta: $A \cup B = C \Leftrightarrow (\forall x \in X)\ \mu_C(x) = máx\ \{\mu_A(x), \mu_B(x)\}$.
Diferencia	La *diferencia* de dos conjuntos difusos A y B es un tercer conjunto difuso, denotado por $A–B$, y definido mediante la siguiente igualdad: $A - B = \mathrm{A} \cap \mathrm{B}^c$.
Diferencia simétrica	La *diferencia simétrica* de dos conjuntos difusos A y B es un tercer conjunto difuso, denotado por $A \Delta B$, y definido así: $A \Delta B = (A \cap B^c) \cup (B \cap A^c)$.

A continuación se expone un caso práctico [14] que indica el tratamiento de los datos, sean éstos precisos o imprecisos, en una estructura de datos. En este caso particular, se presenta una base de datos relacional, a través del álgebra relacional definida por Edgar Frank Codd de los laboratorios IBM y el álgebra relacional difuso definido por el Dr. Juan Medina. Para ello, se utiliza la aplicación FQ (Consultas Difusas) creada por el Dr. José Galindo, la cual permite realizar consultas de tipo nítidas, como el Lenguaje de Consulta Estructurado (SQL), y de tipo difusas, por ejemplo el Lenguaje de Consulta Estructurado Difuso (FSQL).

Se tiene a continuación la relación entre jugadores que almacena los datos de jugadores de básquetbol. La relación está compuesta por los atributos jugadores, equipo, altura y calidad, como se muestra en la tabla 14.

Se identifican valores nítidos y difusos, respecto de los atributos altura y calidad. En lo que sigue, se indicará cómo se encuentran los datos representados en la Base de Datos Relacional Difusa (BDRD), así como también se señalará cómo se almacenan los datos que poseen imprecisión o incertidumbre en ésta. Por ejemplo, el valor $[30, 35, 40, 45] se almacena [30, 35 – 30, 40 – 45, 45], es decir, [1°, 2°-1°, 3°-4°, 4°], tal como establece la Distribución Trapezoidal.

Se identifican valores nítidos y difusos, respecto de los atributos altura y calidad. En lo que sigue, se indicará cómo se encuentran los datos representados en la Base de Datos Relacional Difusa (BDRD), así como también se señalará cómo se almacenan los datos que poseen imprecisión o incertidumbre en ésta. Por ejemplo, el valor $[30, 35, 40, 45] se almacena [30, 35 – 30, 40 – 45, 45], es decir, [1°, 2°-1°, 3°-4°, 4°], tal como establece la Distribución Trapezoidal.

Tabla 14. La relación compuesta por los atributos jugadores, equipo, altura y calidad

Jugador	**Equipo**	**Altura**	**Calidad**
P1	Español	Bajo	$[30,38,40,45]
P2	Español	Muy Alto	$[2,7,10,15]
P3	U.Cat.	Normal	Regular
P4	U.Cat.	192	Regular
P5	U.Cat.	Alto	# 10
P6	U.Conc	198	Malo
P7	U.Conc	Muy Alto	# 35
P8	U.Conc	170	$[31,34,35,38]
P9	Petrox	Bajo	# 15
P10	Petrox	Normal	Bueno
P11	Austral	Muy Alto	# 25
P12	Austral	Bajo	Muy Bueno
P13	Bio Bio	Alto	Muy Bueno
P14	Bio Bio	Muy Alto	# 8
P15	Bio Bio	177	#6
P16	U.Chile	Alto	Muy Bueno
P17	U.Chile	Unknown	Unknown
P18	UCM	Unknown	$[8,12,15,25]
P19	UCM	Normal	#25

Las bases de datos relacionales nítidos imposibilitan el tratamiento apropiado de esta información, debido a la imprecisión o incerteza presente en el dato.

Se hace la siguiente consulta:

Consulta "Dame todos los datos de los jugadores que son posiblemente altos (grado mínimo 0.5) y que tienen una calidad posiblemente menor que buena (grado mínimo 0.25)". Respuesta FSQL.

R1 = Selección jugadores posiblemente altos;

SELECT * from jugadores;

WHERE ALTURA FEQ $Alto thold 0.5.

El resultado de la consulta se muestra en la figura 3.

FSQL (Lenguaje de Consulta Estructurado Difuso) nos entrega el grado de satisfacción de la *n*-upla con respecto a la condición.

A continuación se presenta la relación entre los jugadores que sean posiblemente menores que Bueno y la relación final efectuada en FSQL.

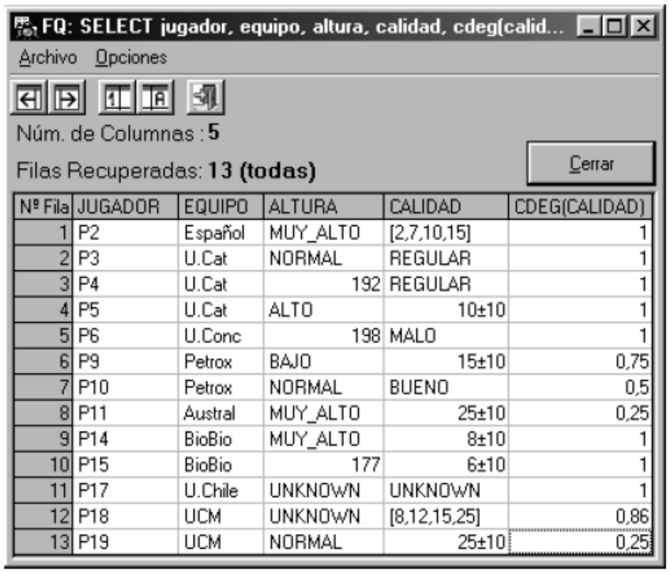
R2 = Selección jugadores que son posiblemente menores que Bueno;

SELECT jugador.% FROM jugadores;

WHERE Calidad FLT $Bueno;

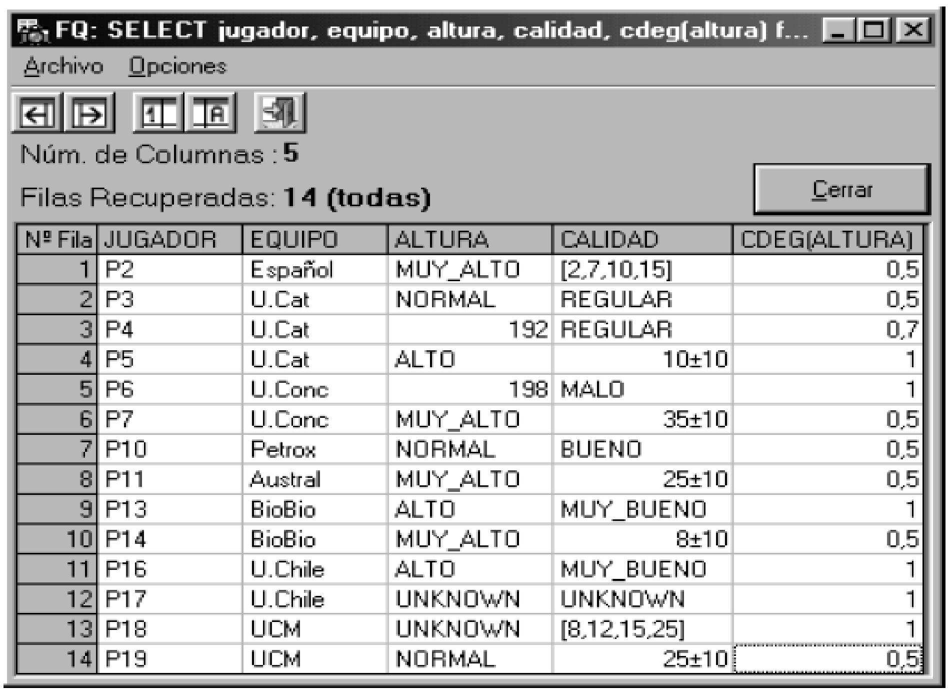
THOLD 0.25.

El resultado de la consulta se muestra en la figura 4



Nº Fila	JUGADOR	EQUIPO	ALTURA	CALIDAD	CDEG(CALIDAD)
1	P2	Español	MUY_ALTO	[2,7,10,15]	1
2	P3	U.Cat	NORMAL	REGULAR	1
3	P4	U.Cat	192	REGULAR	1
4	P5	U.Cat	ALTO	10±10	1
5	P6	U.Conc	198	MALO	1
6	P9	Petrox	BAJO	15±10	0,75
7	P10	Petrox	NORMAL	BUENO	0,5
8	P11	Austral	MUY_ALTO	25±10	0,25
9	P14	BioBio	MUY_ALTO	8±10	1
10	P15	BioBio	177	6±10	1
11	P17	U.Chile	UNKNOWN	UNKNOWN	1
12	P18	UCM	UNKNOWN	[8,12,15,25]	0,86
13	P19	UCM	NORMAL	25±10	0,25

Figura 3. R1 Jugadores posiblemente altos**.**



Nº Fila	JUGADOR	EQUIPO	ALTURA	CALIDAD	CDEG(ALTURA)
1	P2	Español	MUY_ALTO	[2,7,10,15]	0,5
2	P3	U.Cat	NORMAL	REGULAR	0,5
3	P4	U.Cat	192	REGULAR	0,7
4	P5	U.Cat	ALTO	10±10	1
5	P6	U.Conc	198	MALO	1
6	P7	U.Conc	MUY_ALTO	35±10	0,5
7	P10	Petrox	NORMAL	BUENO	0,5
8	P11	Austral	MUY_ALTO	25±10	0,5
9	P13	BioBio	ALTO	MUY_BUENO	1
10	P14	BioBio	MUY_ALTO	8±10	0,5
11	P16	U.Chile	ALTO	MUY_BUENO	1
12	P17	U.Chile	UNKNOWN	UNKNOWN	1
13	P18	UCM	UNKNOWN	[8,12,15,25]	1
14	P19	UCM	NORMAL	25±10	0,5

Figura 4. R2 Jugadores posiblemente menores que bueno

Luego se define el conjunto difuso que dará el grado de pertenencia a las etiquetas para los atributos altura y calidad, éstos son definidos para la relación, como se aprecia en las figuras 5 y 6.

ALTURA

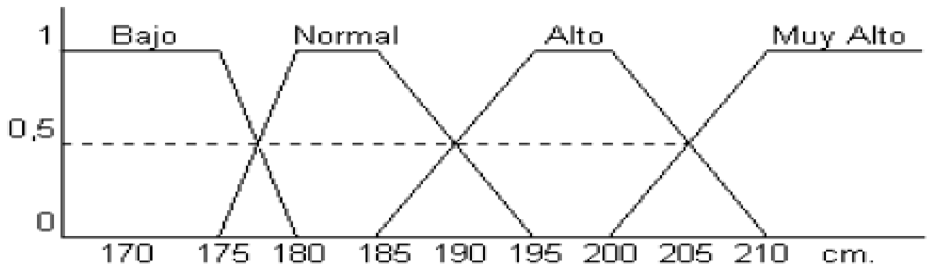


Figura 5. Etiqueta para altura

CALIDAD

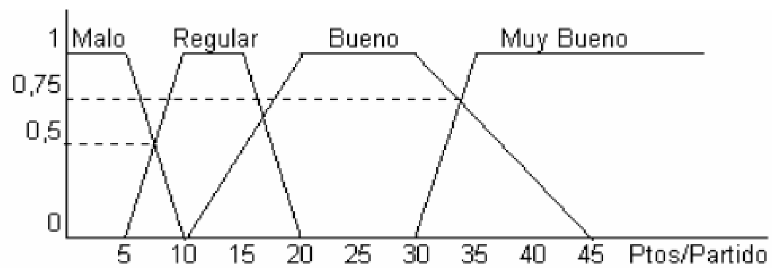


Figura 6. Etiqueta para calidad

También es importante, para un mejor entendimiento, realizar las gráficas por *n*-uplas, para representar las etiquetas (valores con signo $ delante) y las aproximaciones (valores con # delante). Solo se muestra un ejemplo, como se aprecia en la figura 7 y la tabla XV.

1: $[30,38,40,45]; 2: $[2,7,10,15]; 3:$[8,12,15,25].

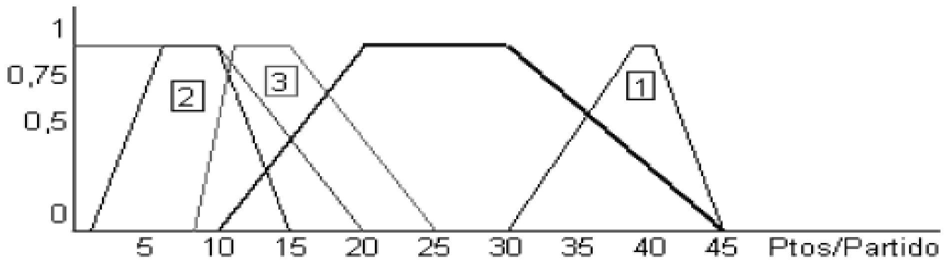


Figura 7. Etiqueta para Calidad de las *n*-uplas 1, 2 y 18

Tabla 15. Calidad de las *n*-uplas 1, 2 y 18.

JUGADOR	CALIDAD
P1	$[30,38,40,45]
P2	$[2,7,10,15]
P18	$[8,12,15,25]

4. DISCUSIÓN DE RESULTADOS

La importancia de los datos en cualquier actividad es vital, por lo que éstos deben representar la realidad y permitir que sean consultados de forma rápida y fácil. Debido a la gran cantidad de datos existente y a la diversidad de los mismos, es que se han generado las estructuras de datos, las cuales son fundamentales para el manejo de la información. Tomando en cuenta la diversidad de datos, hay dos tipos que son ampliamente estudiados en el área de las estructuras de datos, principalmente en las bases de datos relacionales como son los datos nítidos y los datos difusos.

Haciendo un análisis de la teoría estudiada se puede construir una tabla comparativa de dichos datos, la cual se presenta a continuación en la tabla XVI.

Tabla 16 Comparación de datos nítidos y datos difusos**.**

DATOS NÍTIDOS	DATOS DIFUSOS
Se representan mediante la teoría de conjuntos nítidos.	Se representan mediante la teoría de conjuntos difusos.
Utiliza lógica bivalente.	Utiliza la lógica difusa.
Pertenencia del dato representada por un grado del conjunto binario {0,1}.	Pertenencia del dato representada por un grado del intervalo cerrado [0, 1].
Son exactos, precisos o bien definidos.	Representan imprecisión e incertidumbre.
Se aplican en procesos que no estén bajo incertidumbre.	Se aplican a fenómenos o procesos que no son simples y que estén bajo incertidumbre.
Inservibles para expresar la ambigüedad del significado que se da en el lenguaje natural.	Resultan especialmente útiles para expresar la ambigüedad del significado que se da en el lenguaje natural.
Se aplican a modelos matemáticos convencionales.	Se aplican a modelos matemáticos que permite utilizar conceptos relativos a la realidad, siguiendo patrones de comportamiento similares al pensamiento de los humanos.

Fuente: Elaboración propia

5. CONCLUSIONES

- Las estructuras de datos son fundamentales para buscar, conservar, procesar, aplicar, difundir y gestionar la información.
- La teoría de conjuntos difusos permite acercar el funcionamiento de las estructuras de datos al modo de trabajo de los seres humanos, pues éstos manejan con gran frecuencia conceptos difusos (como “alto”, “bueno”, “malo”) que incluyen cierta imprecisión y que los sistemas informáticos tradicionales no entienden y por tanto, no pueden utilizar.
- Los sistemas extraídos de la lógica clásica presentan las dificultades de la rigidez y la bivalencia, y resultan, por ello, inservibles para expresar la ambigüedad del significado que se da en el lenguaje natural, base fundamental de nuestros procesos cognoscitivos en la toma de decisiones.
- Los modelos de Lógica Difusa son altamente flexibles, más tolerantes a la imprecisión de los datos y pueden trabajar con funciones no lineales de diversa complejidad, se les puede modificar fácilmente dependiendo de la solución requerida del problema.
- Todos los conjuntos nítidos tienen funciones de pertenencia bivalentes; en cambio, existen conjuntos difusos que tienen funciones de pertenencia cuyos rangos son conjuntos infinitos no numerables de números reales.
- Tanto los datos nítidos como los datos difusos son buenos para las aplicaciones en ingeniería; sin embargo, para las aplicaciones a fenómenos o procesos que no son simples, en ambientes de incertidumbre, los tipos de datos difusos son los más adecuados.

6. REFERENCIAS

[1] Martínez, R. y Quíroga, E. “**Estructuras de Datos**”. Mèxico D.F., Thomson, 2002, 269 pp.

[2] Carrano F. y Prichard J. “**Data Abstraction and Problem Solving with Java**”. Addisson Wesley, 2001.

[3] Loomis, M. “**Estructuras de Datos y Organización de Archivos”.** Segunda Edición**,** Naucalpan de Juárez, Prentice Hall, 1991, 516 pp.

[4] Universidad Pedagógica Experimental Libertador, Vicerrectorado de Investigación y Postgrado. **Manual de Trabajos de Grado de Maestría y Tesis Doctorales**. Caracas, 2006.

[5] Marzal, V. “**Estructuras de Datos**”. Curso. Material con fines didácticos, Disponible en: www.lawebdelprogramador.com/**cursos**, 2000.

[6] Sisa, A. “**Estructuras de Datos y Algoritmos**”. Bogota DC, Prentice Hall, 2002, 310 pp.

[7] Millán J., "**Vocabulario de ordenadores e Internet**", Disponible en: http://jamillan.com, 2001.

[8] Ferrís, R., Albert J. "**Algoritmos y Estructuras de Datos I**", Material con fines didácticos, Disponible en: http://informatica.uv.es/iiguia/AED, 2001.

[9] Urrutia A., Varas M., y Galindo J. "**Diseño de una Base de Datos Difusa Modelada con UML**" Disponible en: www.inf.udec.cl/~mvaras/papers, 2003.

[10] Torres, E. "**Lógica Formal**". Material con fines didácticos. Doctorado en Ciencias de la Ingeniería, mención Productividad, UNEXPO, Barquisimeto, 2006, pp.103.

[11] Zadeh L. A. "**Fuzzy Sets**". Information and Control, 8, 1965, pp. 338-353.

[12] Torres, E. "**Conjuntos y números borrosos**". Material con fines didácticos. Doctorado en Ciencias de la Ingeniería, mención Productividad, UNEXPO, Barquisimeto, 2007, pp. 60.

[13] Blanco M. I. J. "**Deducción en Bases de Datos Relacionales Difusas**". Tesis Doctoral. Universidad de Granada E.T.S. de Ingeniería Informática, 2001.

[14] Urrutia A., y Rodríguez C. "**Un Caso Práctico de Comparación del Álgebra Relacional**", Ingeniería informática, ISSN 0717-4195, Nº. 7, 2002, pp. 4-8.