17
2 ej

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES
"ACATLAN"

USO DE CARITAS DE CHERNOFF PARA EL PROCESAMIENTO DE CONSULTAS DIFUSAS

T E S I S

QUE PARA OBTENER EL TITULO DE:

LICENCIADO EN MATEMATICAS APLICADAS Y COMPUTACION

P R E S E N T A:

CATALINA ZECUA FERNANDEZ

ASESOR DE TESIS:
DR. SERGIO V. CHAPA VERGARA

MEXICO, D. F. 1991.

INDICE

PAG.

INTRODUCCION

El objetivo de este trabajo es el proponer un sistema gráfico que pueda plantear consultas difusas a una base de datos. Es un trabajo cuya contribución original consiste en proponer nuevos iconos que permitan un procesamiento más inteligente con una clasificación automática de datos. Los iconos son caritas con una semántica muy rica que asocian a operadores de alto orden que permiten formular consultas difusas, con respuestas de datos clasificados. Los iconos de caritas de Chernoff asociado a un método de programación automática en variables lingüísticas y teoría de conjuntos es una nueva modalidad como sistema de consulta difusa a una base de datos.

El principal antecedente Herman Chernoff en 1971 propone el uso de caritas para la representación gráfica de datos multivariados, a partir de esto, un cierto número de proyectos de investigación de análisis exploratorio de datos han sido llevados a cabo. La representación icónica de este método de carita con: cejas, nariz, boca y orejas; para significar un agrupamiento de variables o valores dentro de un proceso automático de operación. Las caritas son editadas gráficamente para formar un icono que encapsula un operador para agrupar datos, considerando que los rasgos influyen en las asignaciones y sesgos con métricas para el escalamiento de datos, y con esto realizar una consulta difusa a una base de datos.

Es conveniente comentar que bajo la misma temática existen proyectos que se han tomado como base para este trabajo de tesis, como son: "Lenguajes de Flujograma para la consulta en una Base de Datos (un estudio comparativo)"[6], y "Caritas como Representación Icónica para el Procesamiento de Consultas Difusas en el Lenguaje de Flujograma"[33], cuya participación se dio lugar en la quinta Conferencia Internacional de UNISYS y la UNAM llevada a cabo en noviembre de 1989.

El presente trabajo se desarrolla sobre dos ejes principales: el computacional con metodología de programación visual y gráfica, y fundamentación matemática en Lógica Difusa. Trato en este trabajo las representaciones gráficas como una herramienta cognoscitiva en computación modelos de abstracción alta. El uso se plantea como representación de datos, o bien, para la creación de lenguajes visuales que permitan a través de programación automática resolver problemas. El tipo de problemas que se quieren resolver son consultas de tipo difuso mediante el planteamiento con caritas de Chernoff que deben ser interpretadas a un lenguaje de variables lingüísticas con un sistema de computación que basado en métodos formales de lógica difusa transforma a variables difusas. Las respuestas a la formulación de preguntas son una clase de tuplas de relaciones de una base de datos estructurada en el modelo relacional.

En el capítulo 1, se dan los antecedentes de las representaciones gráficas dado que han sido de gran interés a lo largo de la historia, su gran importancia como comunicación visual ya que han servido como un medio de comunicación para expresar los pensamientos e ideas del hombre a sus semejantes; y por tal razón el interés aún persiste en mejorar las representaciones gráficas como un medio de divulgación de la información, presentando entonces las distintas formas de gráficas como representación de datos, como gráficas de programación y finalmente las herramientas gráficas por computadora.

En el capítulo 2, se presentan en primer lugar los conceptos y definiciones sobre los que está soportada la teoría de conjuntos difusos y las operaciones básicas que pueden realizarse con ellos. Teniendo en cuenta que la teoría de los conjuntos difusos representa un intento para construir un esquema conceptual para el trato sistemático tanto en el aspecto cuantitativo como cualitativo, lo cual nos conduce a la lógica difusa.

En el capítulo 3, con base en la teoría de conjuntos difusos del capítulo 2, se trata la semántica difusa cuantitativa al considerar las etiquetas lingüísticas como operadores. Todo esto como bases matemáticas para el procesamiento de una consulta difusa que se basa en variables lingüísticas como una exploración natural donde el uso de palabras o declaraciones más que números son caracterizaciones lingüísticas, que en general son menos específicas que las numéricas.

En el capítulo 4, se menciona la idea inicial u original de las caritas de Chernoff como una técnica de desplegado de datos multivariados, así como sus ventajas y desventajas contra otras técnicas y se mencionan algunas aplicaciones.

Por último en el capítulo 5, se describen los aspectos formales del procesamiento de una consulta difusa que es basado en la teoría de conjuntos difusos y variables lingüísticas que se describen en el capítulo 2 y 3. Los iconos basados en caritas de Chernoff que se mencionan en el capítulo 4 en combinación con esta teoría de conjuntos, los presento como una nueva modalidad en el intento del manejo de consultas difusas en una base de datos. En esta parte se mencionan las técnicas conceptuales para el procesamiento de una consulta difusa y se plantea el modelo de caritas como una nueva sintaxis a este tipo de consulta, la cual estará representada por medio de los rasgos faciales, así mismo se hace una comparación con un lenguaje de consulta artificial llamado SEQUEL.

Finalmente se hace una breve conclusión de lo que es esta propuesta de uso de caritas para el procesamiento de consultas difusas, como una nueva forma de sintaxis gráfica para elaborar preguntas a una base de datos, pretendiendo que esta sintaxis sea más ambigua ya que contendrá términos lingüísticos.

CAPITULO 1

GRAFICAS, HERRAMIENTA COGNOSCITIVA EN COMPUTACION

1.1 ANTECEDENTES.

El desarrollo del hombre y el crecimiento de las civilizaciones ha dependido en gran parte de los progresos en algunas actividades; el descubrimiento del fuego, la domesticación de animales, la división del trabajo, el desarrollo del lenguaje; todas en la evolución de pensar para recibir, para comunicarse y para registrar sus conocimientos.

El hombre es esencialmente un ser comunicativo; la comunicación es una de sus actividades importantes, él domina el único poder de hablar y escribir; él ha contactado con sus antecesores y sus descendientes en un sentido de historia y tradición. Tales poderes de comunicación han hecho posible su organización en las sociedades más complejas y manteniendo un estado de cambio continuo.

La comunicación se logra a través de un lenguaje o un simbolismo; ya sea hablar un dialecto, grabados en piedra, señales en algún código y hasta una cadena de pulsaciones de números binarios en una computadora moderna.

Las escrituras primitivas de las civilizaciones del Mediterráneo fueron pictografías, jeroglíficos y símbolos; simplemente representaciones pictóricas para la reproducción de objetos y también para asociación de nombres, acciones e ideas de todo tipo; pero el paso de gran importancia fué la escritura fonética con la cual los sonidos dieron origen a los símbolos, ya que la escritura y el lenguaje forman un proceso, a pesar de las dificultades que presentaron las escrituras y papiros por la mezcla de signos fonéticos y pictogramas, al igual que muchos signos ociosos y adornos. Por tal razón a través de toda la historia el hombre ha puesto gran interés en mejorar la forma de comunicar sus pensamientos e ideas.

El simbolismo matemático moderno ilustra un sistema de lenguaje, procesando un alto grado de comprensión de información. El triunfo de Galileo fué el reconocimiento de las matemáticas como un lenguaje universal para la descripción de sistemas físicos, posterioremente Descartes y su aplicación de fórmulas para geometría. También Leibnitz trabaja los símbolos matemáticos, que influyen en forma decisiva en el desarrollo de las matemáticas modernas.

Russell trató las bases matemáticas en 1910 como una generalización de la lógica simbólica que fué escrita casi completamente sin palabras, de ahí que Leibnitz no estuviera tan lejos de pensar en un simbolismo matemático como un "lenguaje de lógica" universal.

Uno de los sueños de Leibnitz fué el de crear un lenguaje universal con símbolos pictóricos para ser interpretados lo más directamente posible y sin traducción alguna.

A partir de aqui podemos establecer un punto de referencia del cual se tiene el interés en lenguajes de tipo visual.

En la actualidad cada vez más se vislumbran lenguajes nuevos de tipo visual y modelos con representaciones icónicas. Esto se debe a que las imágenes ofrecen una percepción rápida y con frecuencia con significado completo en su comunicación. La creación de una variedad amplia de nuevas formas de lenguaje simbólico ha sido motivado por ciertos problemas de la vida cotidiana, permitiendo su adecuada solución con su herramienta visual respectiva. Por ejemplo, las señales en las carreteras, en los aeropuertos, o en almacenes de servicio, entre otros.

El resurgimiento de este tipo de comunicación visual también ha incursionado en disciplinas del conocimiento, en atención a que las imágenes son una herramienta útil y poderosa. En particular, para la administración y la computación se han creado elementos pictóricos con la finalidad de cumplir con ciertos requerimientos de enfoque. De aqui la importancia por la creación de una comunicación de naturaleza visual.

1.2 ICONOS, GRAFICAS Y DIAGRAMAS COMO COMUNICACION VISUAL.

El tipo de comunicación visual representa uno de los cuatro conocimientos intelectuales básicos : literal, numérico, articulado y gráfico. El literal incluye los conocimientos básicos de lectura y escritura (es decir, comunicación con palabras escritas), numérico expresa la habilidad para comunicar en números y otras notaciones matemáticas, articulado el arte de la comunicación hablada y el gráfico que implica la habilidad para la comunicación visual.

Los símbolos pictóricos, gráficas y diagramas son manifestaciones visuales que tienen por objeto el representar, ilustrar y comunicar objetos e ideas.

En general, en la comunicación visual se involucran primeramente símbolos que dependen de las necesidades que se pretenden satisfacer, con una clasificación que se presenta en seguida:

* De representación.
* Abstractos y
* Arbitrarios.

Un símbolo de representación reproduce en forma simple y fidedigna la imagen de un objeto o una acción. Los símbolos abstractos reducen elementos de un mensaje a términos gráficos los cuales pueden ser representativos, pero también haberse simplificado por el diseñador o cambiado al transcurrir el tiempo. Y finalmente, los símbolos arbitrarios que son aquellos que se inventan para que tengan que ser aprendidos.

SIMBOLOS PICTORICOS O ICONOS.

Los símbolos pictóricos denominados también ICONOS son manifestaciones del universo visual que contienen forma y color. Los ICONOS tienen asociada una semántica y con reglas sintácticas es posible crear un lenguaje pictórico. De todo esto tenemos en la actualidad una área de interés denominada comunicación icónica.

La palabra ICONO viene del griego IKON y significa un modo de comunicación usando imágenes visuales que permitan a las personas percibir de forma natural; la forma y señal antes que símbolos alfabéticos los cuales son definidos en términos de convenciones arbitrarias y requieren de una educación especial para interpretarse.

En años recientes el interés por un tipo especial de representaciones pictóricas se ha desarrollado originalmente como parte de una técnica estadística analítica. Las tentativas son hechas para adaptar su innovación para la presentación y comunicación de los datos estadísticos.

En general los símbolos pictóricos forman las llamadas gráficas pictóricas que pueden ser clasificadas en cuatro tipos básicos:

(1) Las gráficas en donde el tamaño de los símbolos pictóricos son hechos en proporción a los valores representados.

(2) Las gráficas de unidades pictóricas en las cuales cada símbolo representa un valor uniforme y finito.

(3) Las gráficas de caricatura y boceto en donde la forma gráfica básica, como una curva o barra es representada como una imagen; y

(4) Las gráficas con embellecimientos pictóricos.

Entre los símbolos usados en las técnicas de gráficas analíticas se tiene a "glyphs", "estrellas", "árboles" y "caras de caricaturas". Los símbolos más frecuentemente usados para los propósitos de presentación y comunicación son las "caritas de caricatura" inventadas por Herman Chernoff.

GRAFICAS.

Las gráficas son una herramienta poderosa de la comunicación visual, que resumen una gran cantidad de información para que el lector de una sola mirada, perciba y examine los datos, estimulando el análisis para encontrar relaciones escondidas.

Las gráficas las podemos clasificar[1,2] en:

(a) Gráficas que describen relaciones teóricas tales como: funciones de probabilidad, funciones de densidad y funciones de riesgo .

(b) Gráficas no numéricas como: monogramas, caritas de Chernoff , tablas y mapas.

(c) Gráficas que despliegan datos y resultados de análisis como: diagramas de series de tiempo, histogramas, resultados del método de Monte Carlo.

(d) Gráficas analíticas como: media, normal y otras gráficas probabilísticas donde las conclusiones son dibujadas directamente en la gráfica.

En general las gráficas son representaciones de datos por medio de la combinación de puntos, lineas, sistemas de coordenadas, números, símbolos, palabras, sombras y color.

DIAGRAMAS.

Los diagramas son dibujos geométricos que tienen por objeto representar una determinada situación, resolver o expresar gráficamente un problema, a fin de comprender facilmente el procedimiento y la solución.

Los diagramas son un nuevo paradigma que de forma gradual estan reemplazando a los lenguajes de texto. En general ayudan al diseñador de un sistema de "software" para analizar claramente el sistema y ejecutar un buen diseño estructural.

Por otra parte, como instrumentos para documentación, estos facilitan la comunicación entre los diseñadores y usuarios, permitiendo hacer depuraciones en un fácil mantenimiento.

También es común usar técnicas de diagramación en las etapas de desarrollo de software propuestas como un lenguaje de programación visual; estas técnicas pueden ser divididos en tres diferentes categorias:

(1) Estructura de árbol como: descomposicion funcional, y diagramas estructurados.

(2) Diagrama estructurado como: diagrama de Nassi-Shnederman.

(3) Diagrama no estructurado como: redes de Petri, diagrama de flujo de datos.

Las gráficas/diagramas se evalúan como buenas o malas en términos de la calidad y desempeño funcional de su representación visual que se establece en función de las condiciones especificas, las condiciones de planeación y ejecución; o bien, algunos términos abstractos.

Los criterios que generalmente se aplican para clasificar a las gráficas de buenas o malas, o aceptables o no aceptables son:

Precisión.- La dimensión y otros aspectos físicos de una buena gráfica pueden reflejar el alto grado de precisión, una gráfica no debe ser engañosa o distorsionada para no tener una mala intepretación y como consecuencia la inexactitud.

También se debe seleccionar los patrones o símbolos adecuados para no crear ilusiones ópticas.

Simplicidad.- El diseño de una gráfica debe ser simple y directo, no cargarla con irrelevancias , símbolos triviales o superfluos y adornos. Un objetivo importante del diseño es maximizar la información comunicada con los límites de simplicidad y concisión. Una buena gráfica nunca debe poseer características complejas que requieran mucho tiempo y esfuerzo para ser descifradas.

Claridad.- En la evaluación de una gráfica la calidad, el valor y la efectividad son algunas de las características más importantes. Si la gráfica es visualmente ambigua y confusa, el proceso de comunicación puede ser rápidamente interrumpido. Los elementos de la gráfica deben estar en una armoniosa combinación ocupando lugares correctos; esto ayuda al lector en la intepretación de hechos, entendiendo más rápidamente, de mejor manera y con menos esfuerzo.

Apariencia.- La apariencia de una gráfica tiene mucho que ver con el interés o atracción. Una "buena" gráfica está diseñada y construida para atraer y retener la atención, reflejando una hábil y digna apariencia profesional. Una buena gráfica es artística, incluye una composición armoniosa, en unidad, con proporción, y tintes de contraste que deben balancearse adecuadamente.

Estructura Bien Diseñada.- La estructura de una gráfica bien diseñada puede conformar ciertos principios básicos; uno es el reconocimiento completo de que todos los elemetos diseñados sean interdependientes. Otro principio considerablemente más complicado es la importancia visual y la distinción asociada a los elementos variados de una gráfica, debiendo ser adecuado con el significado de las ideas presentadas.

Entonces un modelo gráfico en su representación ofrece una atractiva atención al lector proporcionando facilidades para recordar las relaciones. Los diagramas y las gráficas tienen un alto significado que resume una gran cantidad de datos que se visualizan de una sola mirada, permitiendo estimular y ayudar a pensar analiticamente para encontrar relaciones escondidas.

1.3 GRAFICAS PARA REPRESENTACION DE DATOS.

En la actualidad y en los diferentes campos como: Sociología, Economía, Ingeniería, entre otros, es imposible prescindir del recurso de la estadística cuyo conocimiento permite representar, explorar, examinar y analizar datos. Las técnicas matemáticas son usadas como una base y con la finalidad de reducir y determinar los datos más sobresalientes.

La presentación de datos estadísticos puede hacerse en tres formas: textual, tabular y gráfica; esto es, puede presentarse en forma escrita, en tablas o en cuadros estadísticos y mediante gráficas. Esta última forma permite mostrar, explicar, interpretar y analizar de manera más clara, sencilla, y efectiva, los datos estadísticos mediante formas geométricas tales como líneas, áreas, volúmenes; permitiendo la comparación de magnitudes, tendencias y relaciones entre los valores. Las gráficas son especialmente útiles como medio de divulgación del análisis estadístico, ya que las relaciones visuales se captan con facilidad y resulta sencillo de recordarlas.

El uso abstracto de imágenes para mostrar números es de reciente creación, quizas de la diversidad de habilidades requeridas -lo artistico visual, la estadística y la matemática-. Esto no fue hasta los años de 1750 a 1800 cuando aparecieron las gráficas estadísticas para mostrar series de tiempo, medidas de dispersión y desplegados multivariados. William Playfair (1759-1823) desarrolló o casi mejoró todos los diseños gráficos fundamentales, reemplazando las tablas convencionales de números con las representaciones visuales sistemáticas de su "aritmética lineal".

Hoy en dia las gráficas modernas pueden hacer mucho más que simplemente sustituir las tablas estadísticas; siendo éstas un instrumento para razonar un poco más la información cuantitativa, un camino efectivo para describir, explorar y resumir un conjunto de números muy grande.

Tal ha sido el interés de la representación de datos, que para llevar a cabo este propósito se han generado varias técnicas, las que pueden clasificase de la siguiente forma:

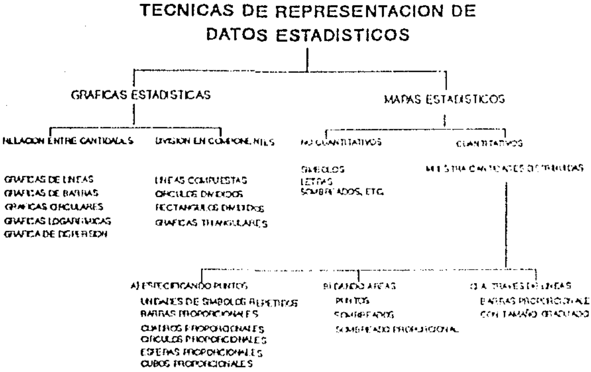


Figura 1.1 Técnicas de Representación de Datos Estadísticos.

De ninguna manera se puede decir que la clasificación[8] presentada en la Figura 1.1, esta completa, pueden existir otras clasificaciones con más técnicas sin embargo, esta nos muestra un panorama de las técnicas estadísticas más comunes y las divide en dos grandes categorías Gráficas Estadísticas y Mapas Estadísticos.

Gráficas Estadísticas.

Son las técnicas que pertenencen a la rama de las matemáticas conocida como estadística, son distribuciones con simples métodos capaces de ser entendidos sin conocimientos de matemáticas. Varias técnicas comprendidas en esta categoría pueden ser divididas en dos grandes grupos:

(a) Técnicas de relación entre cantidades.

Las gráficas son un medio para mostrar los cambios en una cantidad; por ejemplo la temperatura o la altura del barómetro, estos son cambios relacionados a una cantidad, en ambos casos al tiempo. En matemáticas a estos cambios en cantidades se les llaman variables en donde una es la variable dependiente y la otra independiente. Para mostrar este tipo de información la podemos presentar en varias gráficas como son:

Gráfica de lineas.

El nombre de gráficas de lineas es derivado de la relación entre los valores trazados y la unión de cualquier punto por medio de una linea curva uniforme, la cual cambia de dirección sin cambios arbitrarios en los puntos que se trazan como la Figura 1.2, ó altenativamente una serie de lineas rectas (continuas) dibujadas con los valores trazados Figura 1.3.

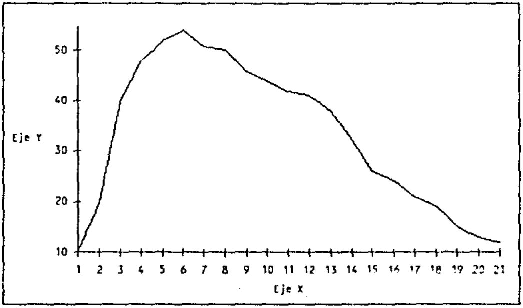


Figura 1.2 Gráfica de Líneas.

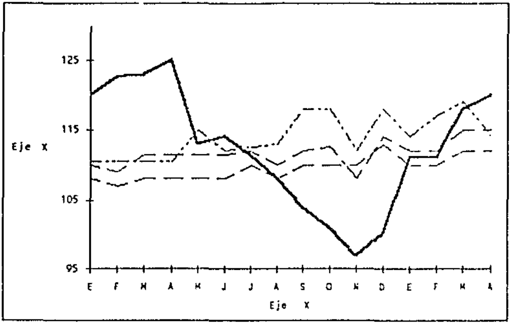


Figura 1.3 Gráfica de Líneas múltiples.

Gráfica de barras

Similar a la gráfica de líneas es la gráfica de barras, cuyos valores son dibujados sobre los ejes; pero en este caso, en lugar de estar unidos los valores por una línea, son representados por una serie de barras verticales u horizontales. Usualmente en cada barra se observa un límite, pero esto no es absolutamente necesario; el tope de las barras señala el total, por otro lado la presencia de las barras atraen la atención y esto ciertamente es una ventaja. Existen diferentes tipos de estas gráficas:

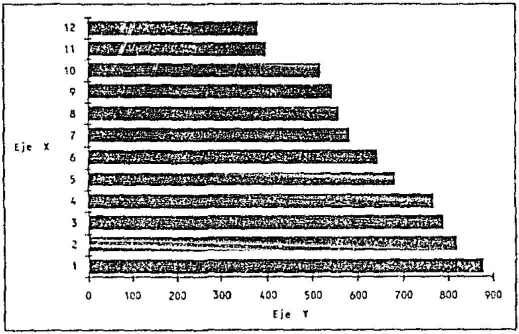


Figura 1.4 Gráfica de Barras Horizontales.

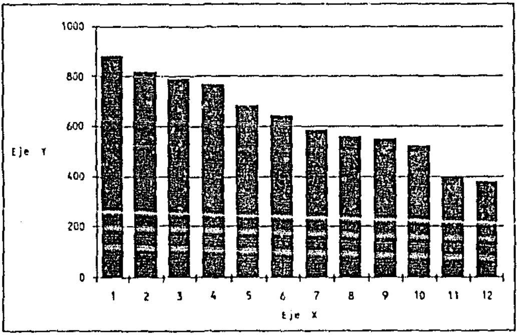


Figura 1.5 Gráfica de Barras Verticales.

Gráficas Circulares.

En la gráfica circular también llamada gráfica de reloj o gráfica polar, la idea básica es simplemente tomar la gráfica normal extenderla, doblarla y cortarla en dos partes iguales. Los valores se incrementan radialmente hacia afuera y es usual marcar un cero en un círculo pequeño. Una gráfica circular no es muy usada; este tipo de gráficas son comúnmente usadas para describir la mortalidad; este tipo de gráfica es confiable en igualdades lógicas y matemáticas pero visualmente es ineficiente.

Gráficas Logarítmicas.

Existen gráficas semilogarítmicas que tienen una gran utilidad para mostrar los cambios relativos de las magnitudes en el tiempo. Las gráficas aritméticas enfatizan los cambios absolutos, las gráficas semilogarítmicas ponen de relieve a los cambios relativos expresados como proporciones y porcentajes. Por el énfasis que estas gráficas dan a los cambios relativos se le conoce también como gráfica de razones; el nombre de las escalas semilogarítmicas proviene de que sólo uno de los ejes tiene escala logarítmica , cuando ambas escalas son logarítmicas se le conoce como gráfica logarítmica. Esta gráfica nunca tiende a cero ya que los valores se dan de acuerdo a los logaritmos.

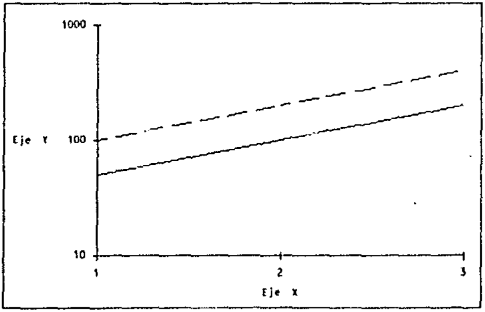


Figura 1.6 Gráfica Logarítmica.

Gráfica de Dispersión.

La gráfica de dispersión es usada para investigar el orden de relación, si existe entre las dos variables en toda el área. Si los puntos presentan una dispersión aleatoria es obvio que no tienden a agruparse; por lo tanto no existe una relación sistemática entre las dos variables. Pero puede ocurrir la tendencia de formar grupos, la tendencia más usual es como la que se ilustra en la Figura 1.7 para los valores trazados a lo largo de una línea o más comúnmente dentro de la estrecha zona lineal. Donde ocurre esto se puede expresar la relación entre las dos variables por medio de una fórmula matemática la cual es relacionada con una línea recta.

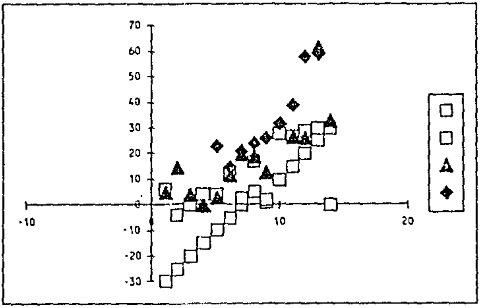


Figura 1.7 Gráfica de Dispersión.

(b) Técnicas para mostrar División en Componentes.

Estas técnicas son aplicadas para indicar que en una parte completa se mueven varios componentes.

Gráficas de líneas y barras compuestas.

Cuando a lo largo de la escala vertical de una gráfica de línea simple o de barras es posible hacer la división de áreas bajo la línea o el contenido en la barra dentro de algún número de partes que la componen; el resultado puede ser llamado una gráfica de líneas o barras compuestas como en la Figura 1.8 y Figura 1.9 respectivamente.

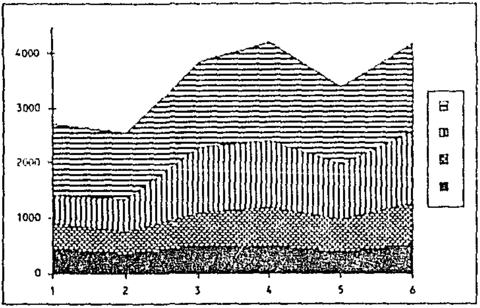


Figura 1.8 Gráfica de Areas.

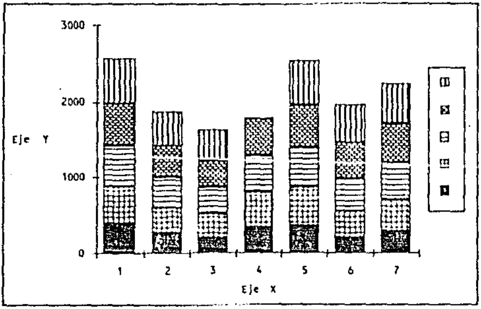


Figura 1.9 Gráfica de Barras Compuestas.

Gráfica de Pastel o círculos divididos.

La gráfica de pastel, generalmente llamada así, es una de las gráficas estadísticas más comunes. La cantidad es representada por un círculo dividido en segmentos proporcionales en tamaño a los componentes, cuyas comparaciones se hacen entre las fluctuaciones como puede observarse en la Figura 1.10. Las gráficas de pastel siguen normalmente tres puntos:

1. Un círculo debe ser dividido proporcionalmente en áreas para la cantidad total que representa.

2. Cada parte del componente debe ser expresado como un decimal o un porcentaje completo.

3. El ángulo al cual corresponde este decimal o porcentaje está dado de acuerdo a 360 grados.

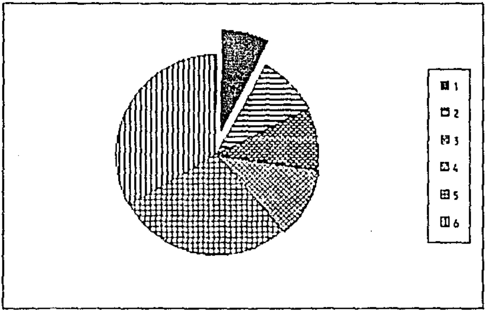


Figura 1.10 Gráfica de Pastel.

Gráfica de rectángulos divididos.

Un rectángulo dividido surge de la idea de tener barras divididas en la gráfica de barras compuestas. En este caso un rectángulo (donde el área puede ser proporcional a la cantidad total) reemplaza el círculo y puede ser dividido en capas donde cada una representa una de las componentes.

Gráfica triangular.

La gráfica triangular es un dispositivo estadístico que en ciertos aspectos puede ser considerada como una gráfica para mostrar tres variables en vez de las dos usuales, alternativamente las tres variables deben ser expresadas como porcentajes, un 100% en total.

La Figura 1.11, es una gráfica que consiste de un triángulo equilátero donde la escala va 0 a 100 en los tres lados, usualmente se incrementa en sentido de las manecillas del reloj a lo largo de cada lado.

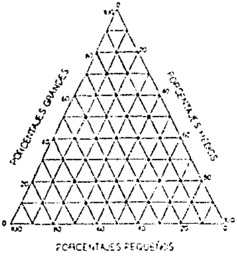


Figura 1.11 Gráfica triangular.

Mapas Estadísticos.

Los mapas estadísticos constituyen uno de los mejores procedimientos para representar relaciones especiales en mapas geográficos. Afortunadamente existe un rango variado de técnicas disponibles:

(a) Mapas estadísticos no cuantitativos.

Estos son llamados simplemente apariencia de igualdad, para marcar los lugares o áreas con alguna característica de interés sin la necesidad de diferenciar de acuerdo al tamaño o importancia sí alguna característica ocurre más de una vez.

(b) Mapas Cuantitativos.

A pesar de la aceptación de los mapas no cuantitativos, los aspectos de cantidad son mucho más importantes en los mapas estadísticos. Existen tres tipos básicos de distribuciones estadísticas:

Una serie de puntos.- este es el que causa menos complicaciones y pueden ser presentados como símbolos repetitivos, barras, circulos, esferas y cubos todos estos proporcionales por un rango de símbolos graduados.

Cantidades contenidas en una área.- estos agrupan cantidades en una área; quizás son los más comunes y los podemos encontrar representados como mapas punteados, sombreados, sombreados proporcionales, lineas iguales, isolíneas.

A través de líneas.- son líneas proporcionales con tamaño graduado.

1.4 GRAFICAS EN PROGRAMACION.

La introduccion de técnicas estructuradas en programación ha sido un gran avance, lo que es llamado como "la revolución estructurada", estas técnicas estructuradas necesitan diagramación estructurada, lo cual ha conducido a desarrollar una diversidad de técnicas de diagramación del "Software".

A principios de la década de los setentas aparece la Programación Estructurada donde se presentan: las conversiones de codificación estructurada, la programación de arriba-abajo (TOP-DOWN) y niveles de abstración de Dijkstra; a la mitad de esta década aparece el Diseño Estructurado con las metodologías de diseño de Jackson y Warnier-Orr; y para finales de la misma se presenta el Análisis Estructurado, donde toman auge los lenguajes de diseño de requerimientos y las técnicas de base de datos.

En la década de los ochentas surgen las Técnicas Automáticas con los modelos de datos automatizados, modelos de datos inteligentes, los lenguajes no procedurales, los diagramas de acción y los generadores de código; y aparece el Diseño Asistido por Computadora en donde se presentan las herramientas de gráficas interactivas para analistas y gráficas para generación de código.

Las gráficas como herramienta cognoscitiva en computación juegan un papel muy importante para los analistas y constructores de sistemas, estas herramientas las podemos clasificar como:

1) Herramientas para dibujar propósitos generales; con las cuales se crean imágenes estáticas de todo tipo.

2) Herramientas para dibujos dinámicos; un dibujo estático es aquel que no tiene un mecanismo incorporado, pero si el dibujo es cambiado o movido y además es definido con conexión entre componentes o relaciones entre iconos es llamado dibujo dinámico.

3) Herramientas para las metodologías existentes; que algunas de éstas son construidas para implementar las metodologías desarrolladas como son diagramas de flujo de datos, diagramas estructurados, diagramas Jackson entre otros. Estos proporcionan los iconos con los cuales estos diagramas pueden ser creados y modificados rápidamente en la pantalla. Algunas de estas herramientas emplean menús, cuadros blancos y llenos u otras figuras geométricas que ayudan a crear los diagramas.

4) Herramientas de diseño gráfico para generar código ejecutable.

La técnica de diagramación debe ser diseñada de manera que el humano dibuje líneas en una computadora de tal forma que pueda realizar un diagrama elegante y claro.

Los diagramas juegan un papel importante en el diseño de sistemas complejos y desarrollo de programas. Cuando un número de personas trabajan en un sistema o programa, los diagramas son una herramienta esencial para la comunicación. Una técnica formal de diagramación es necesaria para desarrollar e intercambiar ideas y hacer componentes separados, adecuados y al mismo tiempo con precisión.

Los diagramas, entonces son un lenguaje esencial para un claro pensamiento y la comunicación humana. Por ejemplo los arquitectos, ingenieros y diseñadores de circuitos tienen herramientas con las cuales pueden dibujar y manipular diagramas en una pantalla de computadora, tales como: CAI, CAD, CAM (Instrucciones Asistidas por Computadora, Diseño Asistido por Computadora y Manufactura Asistida por Computadora) y también CASA y CAP (Análisis de Sistemas Asistido por Computadora y Programación Asistida por Computadora).

Una vez definida la técnica apropiada de diagramación, es mucho más fácil describir actividades complejas y procedimientos en diagramas que en texto. Una imagen es mucho mejor que mil palabras porque la imagen es concisa, precisa y clara.

En general podemos decir que las técnicas de diagramación estructuradas son muy importantes como documentación de programa , dan una visión de su estructura y sus funcionamientos internos. Existen varias técnicas de diagramación :

a) Diagramas de Descomposición.

Los diagramas de descomposición son una herramienta básica para el diseño y análisis estructurado; son árboles simples estructurados los cuales muestran secuencias, una o muchas descomposiciones, opciones y condiciones.

Opciones.- Se representan por medio de un punto en la ramificación; y esto indica que puede o no existir.

Condiciones.- Se representan como la opción sólo que éste muestra un bloque cuando es omitido.

Exclusividad Mutua.- Esto es cuando dos partes y otra son permisibles, pero no ambas; o una parte de un grupo es permisible. Esto es indicado en la línea de bifurcación.

Secuencia.- La secuencia se representa con una flecha en la línea la cual conecta los procesos. Una flecha es usada para indicar la secuencia, que normalmente es de izquierda a derecha o de arriba hacia abajo (ver Figura 1.12).

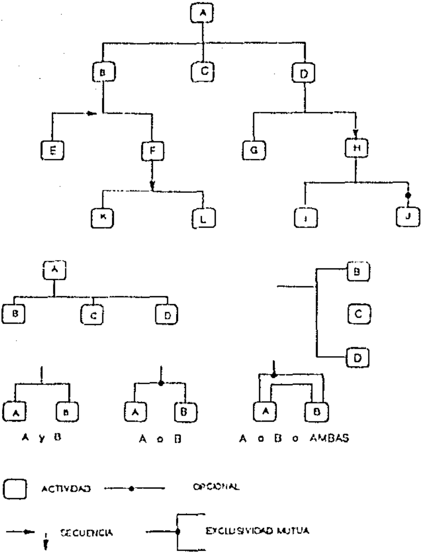


Figura 1.12 Operaciones Básicas de Diagramas de Descomposición.

b) Diagramas de Dependencia.

Un diagrama dependiente tiene bloques para representar las actividades y las flechas entre bloques muestran que una actividad es dependiente de otra; es decir una vez existente la dependecia entre dos actividades una no puede llevarse a cabo hasta que la otra sea completada. Los diagramas dependientes pueden contener:

Opciones.- La opción se muestra con un punto, conectando procesos o procedimientos.

Cardinalidad.- Es cuando un proceso puede ser ejecutado muchas veces.

Bifurcación.- Cuando una actividad es dependiente de muchas otras actividades o muchas actividades son dependientes de una actividad.

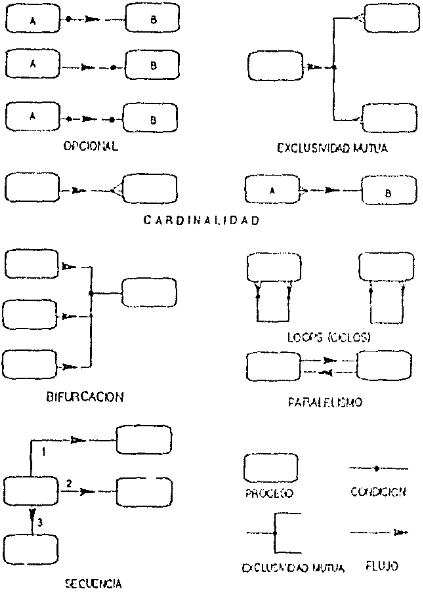
Exclusividad Mutua.- Cuando una u otra actividad deben ser ejecutadas pero no ambas.

Ciclos.- Cuando un proceso depende de él mismo.

Paralelismo.- Es cuando la unión de dos procesos va en ambas direcciones; aquí también ocurre un ciclo.

Eventos.- Es cuando algunos procesos son disparados por otros procesos, se usa una flecha para mostrar la ocurrencia del evento.

Secuencia.- La secuencia es indicada por flechas que por lo general son numeradas (ver Figura 1.13).



Fifura 1.13 Operaciones Básicas de Diagramas de Dependencia.

c) Diagramas de Flujo de Datos

Un diagrama de flujo de datos muestra procesos y el flujo de datos entre éstos; primeramente son usados como herramienta de análisis de sistemas para dibujar los componentes básicos y los datos del procedimiento. Estos diagramas muestran el flujo de datos a través de un sistema lógico, pero no dan el control o secuencia de información.

Los diagramas de flujo de datos son construidos por cuatro componentes básicos: el flujo de datos, los procesos, el almacenamiento de datos y el terminador.

Flujo de datos.- La dirección del flujo de datos es indicado por una flecha con el nombre de los datos escrito sobre la flecha.

Proceso.- Cada proceso es representado por un círculo o una caja con las esquinas redondeadas, el nombre del proceso es escrito dentro de la caja.

Almacenamiento de datos.- El almacenamiento de datos representa un archivo lógico, es dibujado como un par de líneas paralelas algunas veces cerradas y con el nombre del almacenamiento de datos escrito dentro.

Terminador.- Un terminador muestra el origen(fuente) de los datos usados por el sistema y el almacenamiento de memoria, esto es representado por una caja rectangular o cuadro doble (ver Figura 1.14).

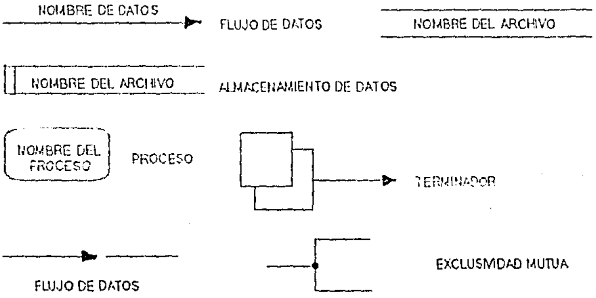


Figura 1.14 Operaciones Básicas de Diagramas de Flujo de Datos.

d) Diagramas Estructurados

Los diagramas estructurados son una representación de árbol jerárquico de un programa estructurado, donde se define la arquitectura completa de un programa mostrado en módulos y sus relaciones.

Los componentes de este tipo de diagramas son cajas rectangulares y flechas que conectan las cajas, donde cada caja rectangular representa un módulo y el nombre del módulo es escrito dentro de la caja.

Estos diagramas muestran las interrelaciones por módulos en niveles y los niveles conectados por flechas; estos son llamados módulos comunes.

Módulos de biblioteca.- La biblioteca predefinida es indicada en el diagrama estructurado por medio de una caja rectangular con líneas verticales dobles.

Transferecia de datos.- Los datos pueden ser transferidos entre una y otra dirección de los módulos. La dirección es mostrada por una pequeña flecha y los nombres de los datos son escritos al lado de la flecha (ver Figura 1.15).

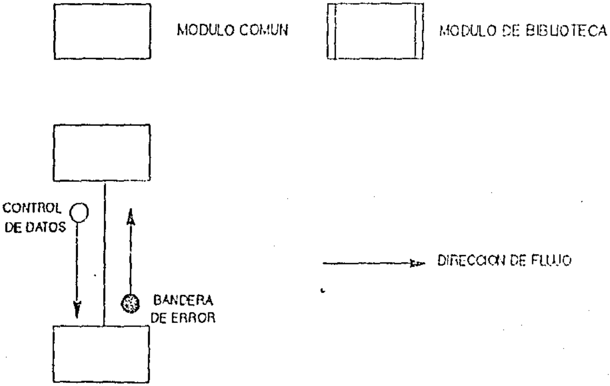


Figura 1.15 Operaciones Básicas de Diagramas Estructurados.

e) Diagramas Warnier-Orr

Estos diagramas se representan mediante paréntesis (llaves) para mostrar la descomposición jerárquica de actividades o datos. Esta descomposición puede representar la estructura o la lógica detallada de un programa.

El diagrama es leído de izquierda a derecha y de arriba hacia abajo con las llaves. Las llaves encierran campos relacionados lógicamente y cada nivel separado jerárquicamente. Los cambios son listados verticalmente, el número de ocurrencias del campo es escrito en paréntesis abajo del nombre.

Los diagramas Warnier-Orr son usados ampliamante para diseñar sistemas nuevos y para documentar sistemas existentes; pero el defecto que tiene este tipo de diagramas es que no muestra condiciones lógicas y otro defecto mayor es que no está orientado a base de datos (ver Figura 1.16).

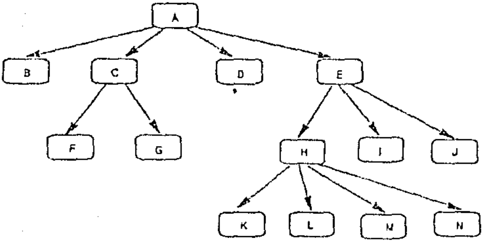
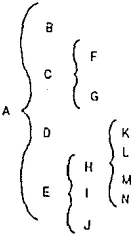


DIAGRAMA DE DESCOMPOSICION



DIAGRAMAS WARNIER-ORR

Figura 1.16 Diagramas de Warnier-Orr.

f) Diagramas Michael Jackson

M. Jackson creó una descomposición de árbol estructurado de la entrada y salida de datos de un programa localizando la correspondencia entre éstos.

Al igual que los diagramas de Warnier-Orr los diagramas Jackson tienen la ventaja de representar tanto la estructura de datos como la estructura del programa; es decir, ambos enfatizan que la estructura del programa puede ser derivado de la estructura de datos.

Jackson ve la estructura de un programa y la estructura de datos como estructuras jerárquicas y usa diagramas de árbol para representarlas. Un diagrama de árbol estructurado esta compuesto de cuatro componentes básicos: secuencia, selección, iteración y elemental.

Secuencia.- Es la especificación del orden, el cual es leído de izquierda a derecha; donde el primer nivel es llamado el componente y el segundo nivel es el listado de estas partes.

Selección.- Un componente de selección esta formado de dos o más partes, exactamante uno ocurre para el componente seleccionado; esto es representado mediante un pequeño círculo dentro de la caja.

Iteración.- Un componente de iteración consiste de cero, una o más ocurrencias. Un asterisco en la esquina superior derecha es usado para indicar la iteración.

Elemental.- Es una parte elemental en un diagrama de estructura y es representada por una caja rectangular.

Estos diagramas no son de gran ayuda para programas lógicos complejos, sin embargo son valiosos para tratar a los sistemas de base de datos que son esencialmente vistos como sistemas de archivos (ver Figura 1.17).

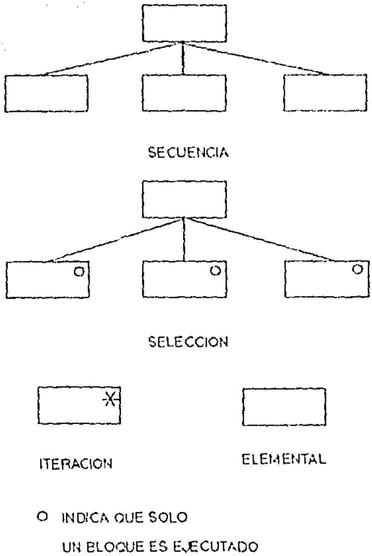


Figura 1.17 Diagramas de Michael Jackson.

g) Diagramas de Flujo

Los diagramas de flujo durante mucho tiempo han sido el método predominante de la representación lógica de un programa. Sin embargo, actualmente han empezado a caer en desuso debido a que la inclusión del GOTO no es muy favorable dentro de la designación estructurada.

Los diagramas de flujo han sido los más usados por los analistas y programadores antes de la era de las técnicas estructuradas. Estos diagramas pueden ser de dos tipos: diagramas de flujo de sistemas y diagramas de flujo de un programa.

Diagramas de Flujo de Sistemas.- muestran la entrada básica, la salida y los componentes del procesamiento para un sistema o programa; mostrando cómo son ejecutadas las operaciones del sistema (ver Figura 1.18).

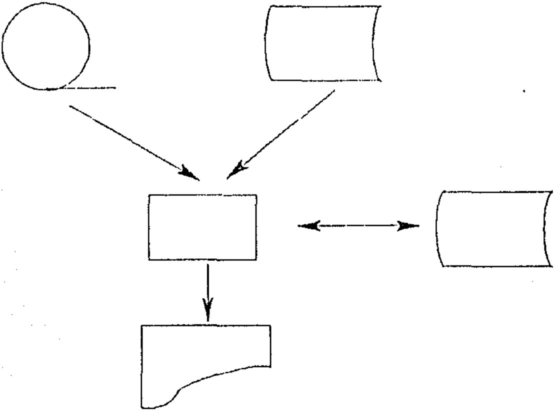


Figura 1.18 Diagrama de Flujo de Sistemas.

Diagramas de Flujo de un Programa.- Es primeramente una herramienta de codificación. La secuencia es mostrada en forma gráfica en la cual las instrucciones o bloques de procesos son ejecutados de acuerdo a una lógica de ejecución. Tradicionalmente los diagramas de flujo han servido para dos propósitos: primero son usados como una herramienta de diseño de programa para planear la lógica detallada y complicada de un programa, segundo son usados como documentador de programas (ver Figura 1.19).

Figura 1.19 Símbolos de Diagramas de Flujo.

h) Diagramas Nassi-Shneiderman (N-S)

Estos diagramas sirven para dibujar el control detallado de la estructura de un programa. A diferencia de muchas de las técnicas para dibujar el panorama de la estructura de un programa, los diagramas Nassi-Shneiderman pueden dibujar condiciones y contructores CASE, DO WHILE y DO UNTIL.

I. Nassi y B. Shneiderman proponen reemplazar los diagramas de flujo tradicionales con un diagrama que ofrece una estructura y una visión jerárquica de la lógica del programa, para diseñar detalladamente un programa y documentarlo.

Estos diagramas representan la estructura de un programa mediante constructores de control de secuencias, selección y repetición; mostrando de esta forma anidamientos y recursión. Estos diagramas son fácilmente convertibles a código estructurado.

Un diagrama N-S consiste de una caja rectangular representando la lógica de un módulo de programa. La caja de un diagrama N-S es propuesta para ser dibujada en una página; sin embargo no puede tener más que de 15 a 20 subdivisiones y cuando un diagrama se hace muy grande sus subfunciones son separadas y dibujadas en otro diagrama N-S.

Secuencia.- La secuencia es mostrada por un apilamiento vertical de cajas llamadas proceso.

Selección.- La selección (IF-THEN-ELSE) es mostrada por la caja de proceso dividida en cinco partes: dividir la caja a la mitad; la mitad de arriba es dividida en tres triángulos, el triángulo superior contiene la condición a ser comprobada; los triángulos de abajo indican la parte "verdadera" y la parte "falsa" de IF-THEN-ELSE. La mitad de abajo es dividida en cajas para el proceso verdadero y falso.

CASE.- Es la estructura en la cual una selección es hecha de múltiples opciones mutuamente exclusivas.

Repetición.- La repetición es indicada por una estructura DO WHILE o DO UNTIL. La estructura de condición DO WHILE es primeramente comprobada y si la condición es verdadera el proceso es ejecutado; en la estructura DO UNTIL el proceso es ejecutado primeramente y entonces la condición es comprobada.

Los diagrama Nassi-Shneiderman son una técnica de diagramación usada principalmente para el diseño detallado de un programa; siendo éstos fáciles de leer y fáciles de convertir a código de programa. Sin embargo ésto es solamente una herramienta de diseño de procedimiento y no puede ser usado para diseñar estructura de datos. En general aunque son fáciles de leer, no siempre son fáciles de dibujar.

Otro defecto de este diagrama es que no está orientado a base de datos; es decir no liga a un modelo de datos o a un diccionario de datos (ver Figura 1.20).

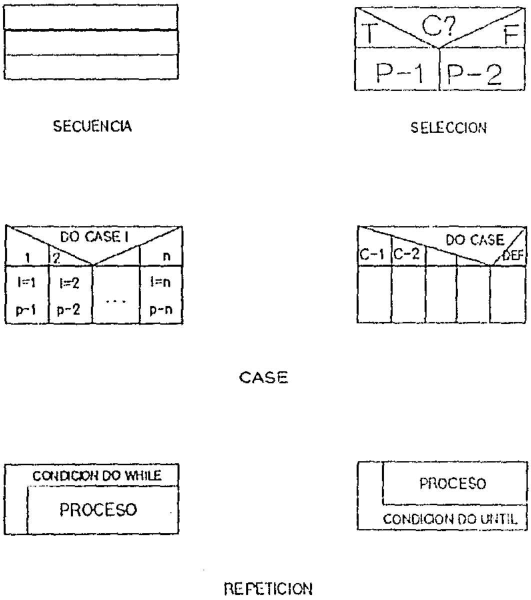


Figura 1.20 Operaciones Basicas de Diagramas Nassi-Shneiderman.

i) Diagramas de Acción.

Esta es una técnica simple para dibujar el panorama de la estructura un programa y el control de su estructura (condición CASE y los constructores DO

WHILE y DO UNTIL). Muchas técnicas no pueden dibujar el panorama de la estructura de un programa y el control estructurado detallado; los diagramas de acción lo pueden hacer.

Estos diagramas son rápidos y fáciles para dibujar y modificar, dibujan todos los constructores de programación estructurada tradicional, su representación en más gráfica que pseudocódigo y están diseñados para ligar un modelo de datos (ver Figura 1.21).

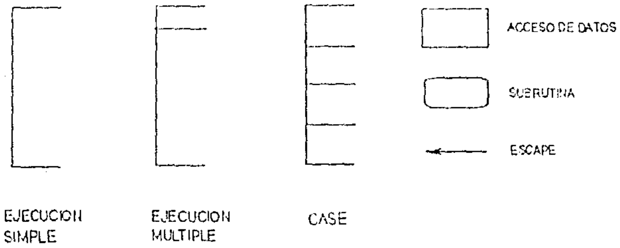


Figura 1.21 Diagramas de Acción.

j) Tablas o árboles de decisión

Los árboles de decisión y las tablas de decisión no se originan como técnicas de diagramación de computadora, éstas tienen una amplia aplicación. Existen cuatro campos de aplicación:

- Taxonomía, diagnóstico y reconocedor de patrones.

- Diseño de circuitos (lógicos) y prueba de rehabilitación.

- Análisis de algoritmos.

- Programación de tablas de decisiones y base de datos.

Las tablas y árboles de decisión pueden ser usados como herramienta de diseño detallado para la lógica de un programa complejo. Estos son raramente usados para mostrar el control de la estructura de un programa, aunque en general no pueden ser usadas como una herramienta de diseño (ver Figura 1.22).

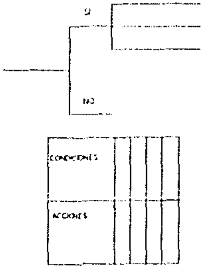


Figura 1.22 Diagrama de Arbol y Tabla de Decisión

k) Diagramas de Estados de Transición.

Es una técnica para dibujar la lógica compleja que envuelve muchas transiciones entre los estados, basandose en la notación de la máquina de estados finitos; estos diagramas son una herramienta útil en el análisis de sistemas, son fáciles para dibujar y ayudan a aclarar situaciones donde ocurren múltiples cambios de estado (ver figura 1.23).

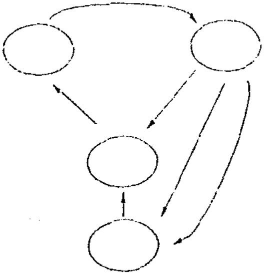


Figura 1.23 Diagrama de Estados.

De las técnicas mencionadas podemos decir que los diagramas estructurados de Warnier-Orr y Jackson dan un panorama de arriba a abajo de un programa estructurado pero la lógica detallada no. Los diagramas de flujo, y los diagramas Nassi-Shneiderman dan la lógica detallada pero no un amplio panorama de la estructura. Los diagramas de acción dan tanto el alto nivel y la visión detallada.

Todas estas técnicas pueden quedar divididas en:

1) Técnicas para mostrar el panorama de la estructura de un programa.

   - Diagramas de descomposición
   - Diagramas estructurados
   - Diagramas de Flujo de Datos
   - Diagramas de Warnier-Orr
   - Diagramas de Jackson
   - Diagramas de Acción

2) Técnicas para mostrar la lógica detallada de programas (condiciones, estructuras de bifurcación CASE, estructuras cíclicas).

   - Diagramas Warnier-Orr
   - Diagramas de Acción
   - Diagramas de Flujo
   - Diagramas Nassi-Shneiderman
   - Arboles de Decisión
   - Tablas de Decisión
   - Diagramas de estados finitos.

Estos diagramas son usados para mostrar la organización estructurada, sistemas estructurados, archivos estructurados y reportes estructurados

1.5 HERRAMIENTAS DE PROGRAMACION EN COMPUTADORA.

1.5.1. Lenguajes de Programación.

La primera generación de los lenguajes de computadora fué el lenguaje de máquina (principos de la década de los cincuentas); las computadoras fueron progranadas con notación binaria.

Por ejemplo:

011011 000000 000000 000001 110101

esto es muy difícil de programar en la computadora sin tener errores, pero la situación pudo mejorar ligeramante usando códigos mnemotécnicos para representar operaciones, entonces una simple instrucción podía ser

CLA 000000 000000 000001 110101

los números pueden ser escritos para las localidades de almacenamiento o registros y entonces la instrucción quedaba.

CLA 00117

Los lenguajes de la segunda generación se empezaron a usar en la mitad de la década de los cincuentas, podemos citar el lenguaje simbólico ensamblador. La escritura simbólica fué usada fisicamente en la máquina.

Por ejemplo:

CLA Salario

La tercera generación tomó auge en la década de los sesentas y fueron llamados lenguajes de alto nivel; como ALGOL y FORTRAN para trabajos científicos, COBOL para trabajos comerciales y PL/I y ADA que incluyen computación cientifica y comercial. Con esta generación el lenguaje se hizo extensamente independiente del hardware; un programador puede codificar sin algún conocimiento de instrucciones de máquina y registros. Una instrucción de esta generación es compilada (traducida) en varias instrucciones de lenguaje de máquina, entonces la cantidad de trabajo necesario para escribir programas es por lo tanto reducido; estos lenguajes son un paso hacia los lenguajes para usuarios.

En la década de los ochentas los lenguajes de cuarta generación son creados para satisfacer los siguientes objetivos:

- Tener alta velocidad en la construcción de procesos de aplicación.

- Hacer aplicaciones fáciles y rápidas, reduciendo el costo de mantenimiento.

- Minimizar los problemas de depuración.

- Generar código de requerimientos de expresiones de alto nivel, y

- Tener lenguajes accesibles para que los usuarios puedan resolver problemas en la computadora.

En general estos lenguajes emplean declaraciones (instrucciones) secuenciales de lenguajes de la tercera generación, empleando una diversidad de mecanismos, tales como el llenado de formas, pantallas interactivas y gráficas de ayuda; también muchos de estos lenguajes dependen de una base de datos y un diccionario o directorio de datos.

Actualmente se está trabajando en lo que es llamado lenguajes de quinta generación, usualmente este término se refiere a sistemas que usan disciplinas que se originan en el campo de la Inteligencia Artificial, específicamente:

- Sistemas basados en el conocimiento.

- Sistemas expertos.

- Máquina de inferencia y

- Procesamiento de lenguajes humanos.

Los sistemas de quinta generación codifican conocimientos complejos como una máquina que puede dibujar inferencias, el proceso de estas inferencias es usado para ejecutar tareas complejas y que aparentemente son triviales para los humanos; tales como entendimiento del idioma, visión y el lenguaje humano.

Los lenguajes que pueden ser considerados de quinta generación están en proceso todavía, tal vez podemos considerar PROLOG como tal; también exite una gran variedad de "shells" para construir sistemas expertos y proporcionar una forma de lenguaje de alto nivel frecuentemente con el lenguaje LISP como base fundamental, consultores de base de datos, generadores de reportes y el diseño de técnicas graficas con inteligencia artificial.

De la diversidad de lenguajes de la primera a lo que podría llamarse quinta generación encontramos que cada uno de estos caen dentro de lo que son llamados lenguajes procedurales y lenguajes no procedurales.

1.5.2. Lenguajes Procedurales y no procedurales.

Algunos lenguajes de cuarta generación son referenciados como lenguajes "No Procedural". Los lenguajes Procedurales son lenguajes que especifican ampliamente el cómo realizar el proceso , mientras un lenguaje no procedural especifica solamente el qué debe realizar sin describir el cómo.

Los lenguajes como COBOL son procedurales, sus programadores dan instrucciones precisas y detalladas para que cada acción sea realizada. Sin embargo, en un generador de aplicaciones, un generador de reportes y un paquete gráfico el usuario sólo llena formas y no se preocupa por el detalle de los procedimientos, a esto se le considera como un lenguaje no procedural.

Algunos lenguajes no procedurales emplean gráficas, los usuarios visualizan los desplegados en la pantalla de la computadora y pueden posicionarse, llenarlos y manipularlos y los resultados son instrucciones de requerimientos que son traducidas a código ejecutable. Recurriendo nuevamente al probervio antiguo chino de que una imagen vale más que mil palabras, entoces las representaciones gráficas de requerimientos o lógica pueden hacer fácilmente su uso y manipulación dado que este tipo de lenguajes dan un gran poder semántico para las instrucciones y que se hacen para favorecer a los usuarios.

1.5.3. Lenguajes Gráficos (Icónicos).

Contrario al ambiente de programación que prevalece en lenguajes de texto como Pascal, Cobol, Lisp, etc.; los lenguajes icónicos -a veces llamados visuales- emplean gráficas. En sí, las dos categorías no son completamente ajenas debido a que ambos tienen involucrados aspectos gráficos y de texto. Algunas veces como elementos decorativos, o algunas otras, integrados más definitivamente de acuerdo a sus reglas sintácticas y semánticas. Sin embargo, Glinert[5] precisa "para que un lenguaje o sistema de programación esté diseñado en terminos visuales, -opuesto a uno de texto- debe cumplir con las siguientes condiciones:

(1) Las entidades gráficas deberán ser diseñadas a un nivel alto, de manera que estén disponibles para los usuarios como átomos. Los átomos podrán ser manejados en la

programación o en el tiempo de ejecución. Las entidades gráficas de alto nivel son objetos geométricos (círculos y cuadros, o cualquier otra imagen que represente "tokens") del alfabeto estándar de programación.

(2) Los elementos gráficos pueden contener texto y números como componentes, formando una parte integral del lenguaje y no sólo como parte decorativa."

Un sistema de programación es denominado icónico si el proceso de programación se basa esencialmente en seleccionar y/o arreglar apropiadamente iconos en una pantalla. Los iconos son imágenes que significan algo y pueden ser ordenados con reglas sintácticas, con la finalidad de representar un programa. Los sistemas icónicos datan de los setentas con Sutherland, quien introduce un sistema que de forma significativa usa gráficas interactivamente. A partir de éste, se desarrollaron algunos otros trabajos, pero no es hasta mediados de la década de los setentas cuando una mayor cantidad de sistemas de tipo gráfico empiezan a ser implantados,como:

PYGMALION.- escrito en SmallTalk por la Xerox Alto, es un sistema para la demostración y programación de tipo procedural. El concepto primordial del sistema son los iconos, los cuales ofrecen como contribuciones: 1) el ambiente de programación se basa en una comunicación hombre-máquina a través de pantallas gráficas, 2) el sistema proporciona al usuario un modelo interactivo y 3) el énfasis en la programación es de procedimiento. La programación en Pygmalion es un proceso de diseño y edición en términos icónicos. Los programadores interactúan con el sistema a nivel de pantalla con imágenes que ellos crean. Con la ayuda de un ratón el programador de Pygmalion demuestra las etapas de cómputo usando valores de datos muestra.

THINGLAB.- es otro sistema gráfico que tiene como objetivo representar gráficamente experimentos de laboratorio orientados a una simulación con restricciones. El lenguaje Thinglab implementado en SmallTalk, incorpora objetos gráficos y simbólicos con un enfoque a dos diferentes clases de usuarios.

OMEGA.- tiene iconos que son abstracciones visuales que representan objetos almacenados con varias alternativas de representación. También el usuario tiene la posibilidad de leer objetos en forma de sintaxis libre con el señalamiento de iconos. Aunque Omega tiene un enfoque gráfico interactivo, mantiene un cierto predominio en la convencional vista de programación de texto.

PECAN.- es un sistema para desarrollar programas que requieren de pantallas de alta resolución para presentar información acerca de la sintaxis, semántica y ejecución del programa. Cada sistema puede presentar diferentes vistas del programa en una sola pantalla: diagrama de Nassi-Shneiderman, diagrama de flujo, lista de programa, árbol del "parser", tabla de símbolos, pila de ejecución y una muestra del diálogo de entrada-salida.

El sistema START de Xerox es un claro ejemplo en donde se utilizan iconos para representar varias funciones y objetos de datos que se encuentran comunmente en los lugares de la administración. En Start es posible procesar registros con el uso de formas que definen los datos. Sin embargo, para el desarrollo de programas de aplicación el usuario debe esperar usar el lenguaje de programación CUSP, el cual es usado para realizar aplicaciones en un ambiente de programación con procedimientos, variables y parámetros.

Tal vez uno de los más populares lenguajes para la automatización de las oficinas de tipo icónico es Query-By-Example (QBE), fué inventado por Zloof en 1974, como lenguaje de consulta basado en el modelo relacional[22]. QBE es un lenguaje de no procedimiento con una programación en dos dimensiones, en donde las operaciones de consulta, actualización y definición en la base de datos se efectua con el llenado de un "ejemplo", que es un objeto gráfico que semeja formas, cartas, tablas y reportes.

FORMAL.- es uno de los más recientes sistemas para desarrollo de aplicaciones. La finalidad del proyecto fué desarrollar un lenguaje de muy alto nivel con la capacidad para computarizar un amplio rango de actividades de procesamiento de datos. El sistema FORMAL para el desarrollo de aplicaciones con una apariencia visual y orientado a formas consiste de tres componentes principales: El modelo de datos de formas, lenguaje orientado a formas y la usual aproximación de llenado de formas. El modelo de datos de formas fué primero sugerido por Nan C. Shu, como una representación en dos dimensiones de datos jerárquicos, con la idea de proporcionar al usuario una ayuda visual para entender las operaciones de las formas de alto nivel. La segunda componente es el lenguaje de programación orientado a formas, llamado FORMAL que es un acronismo de Forms-ORiented MAnipulation Language. Lo fundamental en este lenguaje consiste en llevar las actividades de procesamiento de datos en términos de un procesamiento de formas. En general, cada proceso de formas toma una o dos formas como entrada y produce otra forma de salida. El elemento final del sistema consiste en una comunicación entre el sistema y el usuario determinada con el llenado de una forma.

LENGUAJE DE FLUJOGRAMA.- Un último sistema icónico que vamos a mencionar es el Lenguaje de Flujograma (LF) como lenguaje de muy alto nivel de tipo Gráfico creado por Sergio Chapa V.[6] en base al concepto de Flujograma propuesto por Adolfo Guzmán en [7]. En el lenguaje se definen símbolos gráficos que representan operaciones asociadas a procesos básicos que se ejecutan afectando una colección de datos referentes a la información. Los elementos gráficos son iconos que tienen asociada a su imagen una semántica y con ciertas reglas sintácticas se disponen en una carta gráfica para formar un FLUJOGRAMA.

Los programas en el LF se escriben con un conjunto de iconos que representan procesos y arcos dirigidos de entrada y salida. En ellos fluyen objetos de datos que se identifican con relaciones. Los iconos en el LF están asociados fundamentalmente a: operadores del álgebra relacional, operadores transformadores de archivos o de manejo de datos (ordena, mezcla, actualiza, etc.), procesos que modifican o aumentan un dominio bajo la declaración de una expresión algebráica y funciones que operan sobre un dominio de un atributo; todos ellos con el punto de vista unificado de abstracción de datos en los objetos de relación y dominios.

Actualmente, se están diseñando nuevos iconos basados en caritas de Chernoff (motivo de este trabajo), que en combinación con conjuntos difusos y variables lingüísticas tratados en los dos siguientes capítulos permitan una programación automática para la clasificación de datos mediante un procesamiento de consultas difusas.

CAPITULO 2

CONJUNTOS DIFUSOS

2.1 GENERALIDADES.

Uno de los fundamentos de la ciencia moderna es representar un fenómeno, que aunque no bien entendido, si puede ser caracterizado en términos cuantitativos. Lord Kelvin en 1883 precisa:

" En la ciencia física un primer paso esencial en la dirección del aprendizaje de cualquier sujeto es adquirir los principios del cálculo numérico y métodos prácticos para medir alguna propiedad conectada con este. Frecuentemente cuando decimos que podemos medir, que podemos hablar y expresar con números alguna situación, nosotros conocemos acerca de esto; pero cuando no podemos medirlo, y no podemos expresarlo con números, nuestros conocimientos resultan pocos e insuficientes; y esto puede ser el inicio del conocimiento pero con carencias"

Examinada esta perspectiva, mucho de lo que constituye el fondo de los conocimientos científicos puede ser considerado como un depósito de conceptos y metodologías, los cuales pueden ser representados en modelos matemáticos para producir información cuantitativa y observar su conducta.

Con la llegada de las computadoras y su rápida expansión en el uso de métodos cuantitativos en la mayor parte de los campos de inteligencia humana; indudablemente han resultado ser efectivas en el proceso de sistemas mecánicos, esto es, con sistemas inanimados cuyo comportamiento está regido por leyes de la mecánica, física, química y electromagnetismo.

Desafortunadamente no podemos decir lo mismo con los sistemas humanísticos[1], los cuales han resultado ser resistentes al análisis matemático y simulación por computadora. Desde luego, esto coincide con el uso de las computadoras, que no han dado mucha luz en resultados en filosofía, literatura, derecho, política, sociología y otras

---

1 Por un sistema humanístico entendemos que es aquél cuyo procedimiento es frecuentemente influenciado por la decisión, percepción o emoción humana. Ejemplos de sistemas humanísticos son: sistemas económicos, sistemas políticos, sistemas legales, sistemas educacionales, etc. Un simple individuo y su manera de proceder puede ser considerado como un sistema humanístico.

áreas de orientación humana; tal vez exceptuando algunos ejemplos del proceso del pensamiento humano por medio de Inteligencia Artificial y áreas relacionadas a un conocimiento de naturaleza más cuantitativa y simbolica.

La inefectividad de las computadoras en el comportamiento de los sistemas humanisticos es una manifestación de lo que puede ser llamado el **principio de incompatibilidad** -un principio en el cual se afirma que la alta precisión es incompatible con la alta complejidad. Esto puede verse en el caso de técnicas convencionales de análisis de sistemas y simulación en computadora basados en la manipulación precisa de datos numéricos- siendo incapaces de comprender la complejidad de los procesos del pensamiento humano y la toma de decisiones.

En cuanto a la precisión en el aspecto de complejidad predomina la exploración natural con el uso de lo llamado **variables lingüísticas**, que son variables donde los valores no son números, son palabras o declaraciones en un lenguaje natural o artificial. La motivación para el uso de palabras o declaraciones más que números, son caracterizaciones lingüísticas que en general son menos especificas que las numéricas. Por ejemplo hablar de **edad**; podemos decir "Juan es joven", esto es menos preciso que decir "Juan tiene 25 años"; en este caso la etiqueta joven puede ser considerada como un **valor lingüístico** de la variable **edad**, entendiendo que éste juega el mismo papel que el valor numérico 25 pero es menos preciso y por consiguiente menos informativo. Lo mismo es cierto para los valores lingüísticos muy joven, no joven, extremadamente joven, no muy joven, etc., en contraste con los valores numéricos 20, 21, 22, 23, ...

Ahora bien, el total de valores de una variable lingüística constituye el **conjunto de términos**, el cual en principio puede tener un número infinito de elementos. Por ejemplo, el conjunto de términos de la variable lingüística **edad** puede ser:

T(edad)= joven+no joven+muy joven+no muy joven+mucho muy joven+...+viejo+no viejo+muy viejo+no muy viejo+...+no muy joven y no muy viejo+...+edad media+no edad media+...+no viejo y no edad media+...+extremadamente viejo+...,

en donde + es usado para denotar la unión no como la suma aritmética. De forma similar el conjunto de términos de la variable lingüística Apariencia puede ser:

T(Apariencia)=hermoso+bonito+lindo+bello+atractivo+no hermoso+muy bonito+mucho muy hermoso+más o menos bonito + completamente bonito +completemente hermoso+mediamente hermoso+ no muy atractivo y no muy inatractivo+...

En el caso de la variable lingüística **edad**, la variable numérica **edad** toma valores numéricos de 0,1,2,3,...,100 que pueden ser asociados a una **variable base** para edad. En términos de esta variable un valor lingüístico tal como joven debe ser interpretado como una etiqueta para una restricción **difusa** en los valores de la variable base. Esta restricción difusa es tomada para hacer el significado de joven.

Una restricción difusa en los valores de la variable base es caracterizada por una función **compatibilidad**, la cual asocia a cada valor de la variable base un número dentro del **intervalo** [0,1], que representa la **compatibilidad** con la restricción difusa. Por ejemplo, las compatibilidades de las edades numéricas 22, 28 y 35 con la restricción difusa etiquetada joven puede ser 1, 0.7 y 0.2 respectivamente. El significado de joven, entonces puede ser representado por una gráfica como se muestra en la Figura 2.1.

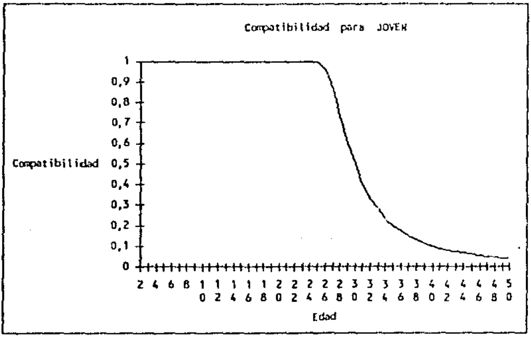


Figura 2.1 Función de compatibilidad para joven

Tomando el ejemplo de "Juan es joven", el nombre "Juan" es visto como el nombre de una variable lingüística compuesta cuyos componentes son llamados variables lingüísticas como Edad, Estatura, Peso, Apariencia, etc. Entonces la declaración "Juan es joven" es interpretada como una ecuación de asignación Figura 2.2, la cual asigna el valor joven a la variable lingüística edad.

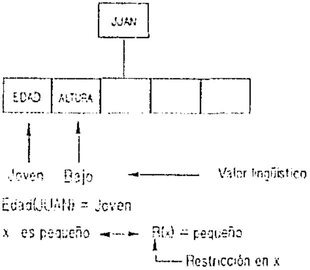


Figura 2.2 Asignación de valores lingüísticos para atributos de Juan y x.

Edad = joven

En cuanto a el valor joven es interpretado como una etiqueta para una restricción difusa en la variable base edad, cuyo significado de esta restricción difusa es definida por la función de compatibilidad. En la Figura 2.3 muestra la estructura jerárquica entre la variable lingüística edad, la restricción difusa la cual representa el significado de estos valores y los valores de la variable base edad.

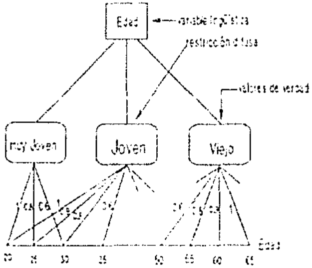


Figura 2.3 Estructura jerárquica de una variable lingüística.

Algunos aspectos básicos del concepto de una variable lingüística son:

Lo primero es comprender que el concepto de compatibilidad es distinto al de probabilidad. Esto es la declaración de compatibilidad de joven con 28 puede ser 0.7, y no relaciona a la probabilidad de tener el valor de edad 28; la interpretación correcta del valor de compatibilidad 0.7 es meramente una indicación subjetiva de la extensión para el valor edad 28, que forma concepciones de la etiqueta joven.

Segundo, nosotros usualmente asumimos que una variable lingüística es **estructurada** en el sentido que es asociada con dos reglas:

Regla 1 (Regla Sintáctica).- Especifica la manera en la cual los valores lingüísticos están en el conjunto de términos de la variable, que puede ser generada. En esta regla asumimos que los términos del conjunto de términos de la variable son generados por un libre contexto gramático.

Regla 2 (Regla Semántica).- Especifica un procedimiento para calcular el significado de algún valor lingüístico dado. Un valor típico de una variable lingüística, por ejemplo no muy joven y no muy viejo envuelve lo que puede ser llamado **términos primarios**, donde los significados de joven y viejo son subjetivos y de contextos independientes.

En general para los términos primarios un valor lingüístico puede envolver conectivos tales como y, o, cualquiera, ninguno, etc.; la negación no (not) y las uniones tales como muy, más o menos, completamente, totalmente, bellamente, extremadamente, algo, etc. Las conectivas, intersecciones y la negación pueden ser tratados como operadores los cuales modifican el significado de sus operandos especificados en forma de contexto independiente. Esto es, si la definición de joven está dada por la función de compatibilidad como es mostrada en la Figura 2.1, entonces el significado de muy joven puede ser obtenido por el cuadrado de la función de compatibilidad de joven, mientras que de no joven puede estar dado por la substración de la función de compatibilidad de joven Figura 2.4.

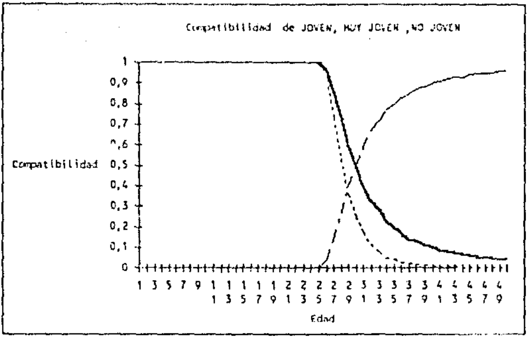


Figura 2.4. Compatibilidad de joven, no joven y muy joven.

Tercero, cuando hablamos de una variable lingüística tal como **edad**, lo fundamental es la variable base **edad** de naturaleza numérica. Esto es, podemos definir el significado de una variable lingüística tal como joven por una función de compatibilidad la cual es asociada con cada edad en el intervalo [0,100] y un número en el intervalo [0,1], que presenta la compatibilidad de la edad con la etiqueta joven.

Por otro lado, en el caso de la variable lingüística Apariencia, no podemos hacer una buena definición de la variable base, es decir no sabemos cómo expresar el grado de belleza y entonces suponemos una función con alguna medida física (a criterio dar valores a un grupo de personas). Los valores de la función de compatibilidad pueden basarse en impresiones las cuales pueden no ser aptas para articularse o formalizarse en términos explícitos; en otras palabras, nosotros definimos la función de compatibilidad, no como un conjunto de objetos matemáticamente bien definidos, pero sí como un conjunto de impresiones etiquetadas, de tal forma que las definiciones son significativas para un humano pero no para la computadora.

Ahora bien, las variables lingüísticas y la lógica difusa son herramientas básicas para tratar de representar el hecho que la mayor parte del razonamiento humano es más aproximado que exacto y que los valores que toman muchas de las variables usadas en el discurso humano son usualmente, palabras y no números por tal motivo el Razonamiento Aproximado es una área importante de aplicación que usa el concepto de una variable lingüística como parte de su método de razonamiento que involucra falta de precisión. Este concepto da como resultado tratar a **verdadero** como una variable lingüística donde los valores verdaderos de un conjunto de términos puede ser:

T(Verdad)=verdadero+ no verdadero+ muy verdadero+ completamente verdadero+ más o menos verdadero+ totalmente verdadero+ esencialmente verdadero+ ...+ falso+ muy falso+ ni verdadero ni falso+...

La variable base correspondiente entonces, asume un número en el intervalo [0,1], y el significado de un término primario tal como verdadero es identificado como una restricción difusa en los valores de una variable base, entonces la restricción es caracterizada por una función de compatibilidad que asocia un número en el intervalo [0,1] con cada valor numérico verdadero.

El concepto de variable lingüística proviene de la caracterización aproximada de fenómenos complejos los cuales resultan difíciles de describirse en términos cuantitativos convencionales. En particular tratar a verdadero como una variable lingüística con valores tales como verdadero, muy verdadero, completamente verdadero, falso, etc., induce a lo que es llamado como **Lógica Difusa**. Dicha Lógica tiene por objeto un razonamiento de aproximación básico, que no es exacto ni inexacto, de tal modo que es más realista para el razonamiento humano, que la lógica tradicional de dos valores.

Una **Variable Lingüistica** se entiende como una variable cuyos valores son palabras o declaraciones en un lenguaje natural o artificial y cuya definición es:

Una quíntupla **(æ, T(æ), U, G, M)**, en donde **æ** es el nombre de la variable; **T(æ)** es el conjunto de términos de **æ**, que es una colección de valores lingüísticos; **U** es el universo de discusión; **G** es la regla sintáctica la cual genera los términos de **T(æ)**; y **M** es la regla semántica la cual está asociada con cada valor lingüístico **x** designado en **M(x)**, donde **M(x)** denota el subconjunto de **U**.

La designación de un valor lingüístico **x** está caracterizado por una función compatible **c:U --> [0,1]**, la cual asocia a cada **u** en **U** compatible con **x**.

A manera de conclusión, quiero mencionar que la capacidad que el hombre tiene para manejar información aproximada o imprecisa y que pueda razonar a partir de ésta, da como consecuencia el desarrollo de un modelo conceptual basado en variables lingüísticas y lógica difusa. La capacidad humana para almacenar del mundo real información vaga o difusa, así como, dar a este mismo información de tal naturaleza; conlleva a una caracterización de los procesos del pensamiento que no es posible modelar con el uso de la lógica tradicional bi-valuada o multi-valuada, para tal caso se tiene una extensión denominada Lógica Difusa que usa valores de verdad difusos, conectivos de lógica y reglas de inferencia Difusa.

Por difusidad se entiende un tipo de imprecisión que se asocia con el uso de conjuntos difusos; anteriormente se había entendido como vaguez, ambigüedad o incertidumbre, sin embargo, desde que Zadeh en 1965 propuso el concepto de conjuntos difusos[26], se ha estudiado en forma amplia y ha tenido gran aceptación, que la aplicación principal de las aproximaciones lingüísticas es en los sistemas humanísticos, especialmente en los campos de inteligencia artificial, lingüística, procesos de decisión humana, reconocedor de patrones, conducta, psicología, diagnósticos médicos y áreas relacionadas.

2.2 CONCEPTO DE UNA VARIABLE DIFUSA.

Primero que todo, definimos el concepto de una variable difusa.

DEFINICION 2.2.1. Una variable difusa es caracterizada por **una triada** $(X, U, R(X;u))$, en donde X es el nombre de la variable; U es un universo de discusión (conjunto finito o infinito); u es el nombre genérico para los elementos de U (un simple nombre para todos los elementos del conjunto); y $R(X;u)$ es un subconjunto de U, el cual representa una restricción en los valores de u, impuesta por X. Usualmente se puede abreviar a $R(X;u)$ como $R(X)$ o $R(u)$ o $R(x)$, donde x denota el nombre genérico para los valores de X y se puede referir a $R(X)$ simplemente como la restricción en u o la restricción impuesta por X.

En general una variable es asociada con una **ecuación de asignación**

$$x=u:R(X) \qquad 2.2.1$$

o equivalentemente

$$x=u, \quad x \in R(X) \qquad 2.2.2$$

la cual representa la asignación de un valor u a x, sujeto a la restricción R(X). Esto quiere decir que la ecuación de asignación es satisfecha si y sólo si $u \in R(X)$.

Ejemplo 2.2.1. Como simple ilustración considerar una variable llamada **edad**. En este caso U debe tomar los enteros del conjunto 0,1,2,3,..., y R(X) debe ser el subconjunto 0,1,2,3,...,100.

Más generalmente, dadas $X_1,\dots,X_n$ variables con respectivos universos de discusión $U_1,\dots,U_n$, la n-tupla ordenada $X=(X_1,\dots,X_n)$ puede ser referenciada como una n-aría variable compuesta (o unión). El universo de discusión para X es el producto cartesiano

$$U=U_1 \times U_2 \times\dots\times U_n \qquad 2.2.3$$

y la restricción $R(X_1,\dots,X_n)$ es una n-aría relacion en $U_1 x\dots x U_n$. Esta relación puede ser definida por la función característica (pertenencia) $\mu_R:U_1 \times\dots\times U_n \dashrightarrow \{0,1\}$ donde

$$\mu_R(u_1,\dots,u_n) = 1 \quad \text{si } (u_1,\dots,u_n) \in R(X_1,\dots,X_n)$$

$$= 0 \quad \text{en otro caso} \qquad 2.2.4$$

u_i es un nombre genérico para los elementos de U_i, $i=1,\dots,n$ correspondientes, la n-aría ecuación de asignación asume la forma

$$(x_1,\dots,x_n) = (u_1,\dots,u_n):R(X_1,\dots,X_n) \qquad 2.2.5$$

la cual es escrita como

$$x_i=u_i, \quad i=1,\dots,n \qquad 2.2.6$$

sujeto a la restricción $(u_1,\dots,u_n) \in R(X_1,\dots,X_n)$ con x_i, $i=1,\dots,n$ denotando un nombre genérico para los valores X_i.

Ejemplo 2.2.2. Suponemos que $X_1 \hat{=}$ edad del padre [2], $X_2 \hat{=}$ edad del hijo, y $U_1 \hat{=} U_2 =\{1,2,\dots,100\}$. Además, suponemos que $X_1 \geq X_2+20$ (X_1 y X_2 son nombres genéricos para los valores de X_1 y X_2). Entonces $R(X_1,X_2)$ pueden ser definidos por

---

2 El símbolo $\hat{=}$ significa "denota" o es "igual por definición"

$$\mu_R(u_1,u_2) = 1 \quad \text{Para } 21 \leq u_1 \leq 100, \quad u_1 \geq u_2+20$$

$$= 0 \quad \text{en otro caso} \qquad 2.2.7$$

2.2.1 Restricciones Condicionadas y Marginales.

Como en el caso de las distribuciones probabilísticas, la restricción $R(X_1,...,X_n)$ impuesta por $(X_1,...,X_n)$ induce a restricciones marginales $R(X_{i1},...,X_{ik})$, impuesta por variables compuestas de la forma $(X_{i1},...,X_{ik})$, donde el índice secuencial $q=(_{i1},...,_{ik})$ es una subsecuencia del índice secuencial $(1,2,...,n)$.

$R(X_{i1},...,X_{ik})$ es una restricción impuesta por $(X_{i1},...,X_{ik})$ la cual satisface la implicación

$$(u_1,...,u_n) \in R(X_1,...,X_n) \Longrightarrow$$

$$(u_{i1},...,u_{ik}) \in R(X_{i1},...,X_{ik}) \qquad 2.2.8$$

dado por la k-tupla, donde $u_{(q)} \triangleq (u_{i1},...,u_{ik})$ es un elemento de $R(X_{i1},...,X_{ik})$, si y sólo si existe una n-tupla $u \triangleq (u_1,...,u_n) \in R(X_1,...,X_n)$, donde los $i_1,...,i_k$ componentes son iguales a $u_{i1},...,u_{ik}$ respectivamente.

Ahora expresado en términos de funciones características de $R(X_1,...,X_n)$ y $R(X_{i1},...,X_{ik})$, esta declaración se transforma en la ecuación

$$\mu_{R(X_{i1},...,X_{ik})}(u_{i1},...,u_{ik}) = \bigvee_{u_{(q')}} \mu_{R(X_1,...,X_n)}(u_1,...,u_n)$$

$$2.2.9$$

en su forma compacta

$$\mu_{R(X_{(q)})}(u_{(q)}) = \bigvee_{u_{(q')}} \mu_{R(X)}(u) \qquad 2.2.10$$

donde q' es el complemento del índice secuencial $q=(i_1,...,i_k)$ relativo a $(1,...,n)$, $u_{(q')}$ es el complemento de la k-tupla $u_{(q)} \triangleq (u_{i1},...,u_{ik})$ relativo a la n-tupla $u \triangleq (u_1,...,u_n)$, $X_{(q)} = (X_{i1},...,X_{ik})$, y el operando $\bigvee_{u(q')}$ denota lo más grande (supremo) sobre las u's, las cuales están en $u_{(q')}$. (los símbolos $\bigvee$ y $\bigwedge$ denotan el Max y Min respectivamente). Esto es, para cualquier real **a,b**

$$a \vee b = \text{Max}(a,b) = a \quad \text{si } a \geq b$$

$$= b \quad \text{si } a < b$$

y $\qquad 2.2.11$

$$a \wedge b = \mathrm{Min}(a,b) = a \quad \text{si } a \leq b$$
$$= b \quad \text{si } a > b$$

En esta notación el símbolo $\bigvee_z$ puede ser leído como "supremo sobre los valores de z", donde z es cualquier elemento. Entonces μ_R puede tomar solo dos valores - 0 o 1- de aquí que $\mu_{R(X_{(q)})}(u_{(q)})$ es 1 si y sólo si existe un $u_{(q')}$ tal que $\mu_{R(X)}(u)=1$.

Comentario 2.2.1. Una simple analogía para aclarar la noción de una variable y los conceptos mencionados. Una variable difusa cuyo significado es formalizado en la Definición 2.2.1., puede ser semejante a una maleta de lados rígidos (duros) rótulada con el nombre X, U representa una lista de objetos para ser puestos en la maleta y R(X) representa una sublista de U que contiene los objetos, que pueden ser puestos dentro de la maleta X. (Por ejemplo, un objeto parecido a un elefante puede no estar en U, mientras que un objeto parecido a una silla puede estar en U y no en R(X), pero si un libro pueden estar en R(X)).

Ahora bien, el poner un objeto en la maleta corresponde a la asignación de un valor a la variable y la restricción corresponde a un subconjunto del universo de discusión, el cual contiene los puntos que pueden ser asignados como valores a la variable. Para esta interpretación la ecuación de asignación

$$x=u:R(X)$$

significa que un objeto u, que satisface la restricción R(X) es puesto en X (es decir una lista de objetos puede ponerse en X). Figura 2.5.

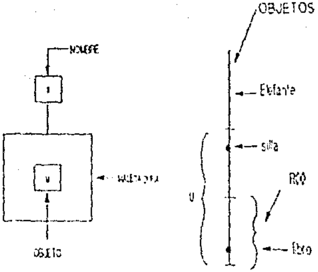


Figura 2.5 Ilustración de la analogía de la maleta para una variable no difusa.

Una n-aría variable compuesta $X=(X_1,\ldots,X_n)$ corresponde a una maleta llevando el rótulo-nombre X, la cual tiene n compartimientos llamados $X_1,\ldots,X_n$ con particiones ajustables entre éstos. Las restricciones $R(X_1,\ldots,X_n)$ corresponden a una lista de n-tuplas de objetos $(u_1,\ldots,u_n)$, tales que u_1 puede ponerse en el compartimiento X_1, u_2 en el compartimiento $X_2,\ldots$,y u_n en el compartimiento X_n simultáneamente (ver Figura 2.6). Por ejemplo si n=2, entonces para una colocación particular de la partición, podemos poner una cubierta en el compartimiento X_1 y acomodar el compartimiento X_2 o bien la cubierta en el compartimiento X_2 y una caja de zapatos en el compartimiento X_1; entonces este evento (cubierta, acomodo) y (zapatos, cubierta) pueden ser incluídos en la lista de pares de objetos, los cuales pueden ser puestos en X simultáneamente.

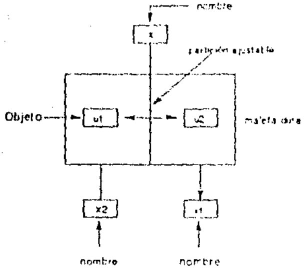


Figura 2.6 Analogía de maleta para una variable binaria no difusa.

En términos de la analogía de la maleta, la n-aría ecuación asignada

$$(X_1,\ldots,X_n)=(u_1,\ldots,u_n):R(X_1,\ldots,X_n)$$

representa la acción de poner u_1 en $X_1,\ldots,$ y u_n en X_n al mismo tiempo, bajo la restricción que la n-tupla de objetos $(u_1,\ldots,u_n)$ que deben estar en la lista $R(X_1,\ldots,X_n)$. Además, una restricción marginal tal como $R(X_{i1},\ldots,X_{ik})$ puede ser interpretada como una lista de k-tuplas de objetos los cuales pueden ser puestos en los compartimientos $X_{i1},\ldots,X_{ik}$ simultáneamente, en conjunción con todas las colocaciones permitidas de objetos en los compartimientos restantes.

Comentario 2.2.2 Como puede notarse (2.2.9) es análoga a la expresión para una distribución marginal de una distribución probabilística, con $\bigvee$ correspondiente a suma (o integración). Sin embargo, esta analogía puede no ser construida para implicar que $R(X_{i1},\ldots,X_{ik})$ es un factor a una distribución probabilística marginal.

Visualizando la parte derecha de (2.2.9) como la función característica de la proyección de $R(X_1,\ldots,X_n)$ en U_{i1} x...x U_{ik}, que en símbolos es:

$$R(X_{i1},\ldots,X_{ik}) = \text{Proj } R(X_1,\ldots,X_n)$$

$$\text{en } U_{i1} \times\ldots\times U_{ik} \qquad 2.2.12$$

o más simple,

$$R(X_{i1},\ldots,X_{ik}) = \text{Pq } R(X_1,\ldots,X_n)$$

donde Pq denota la operación de proyección en U_{i1} x...x U_{ik} con $q=(i_1,\ldots,i_k)$.

Ejemplo 2.2.3. Tomando el ejemplo 2.2.2. tenemos

$$R(X_1) = P_1\ R(X_1,X_2) = \{21,\ldots,100\}$$
$$R(X_2) = P_2\ R(X_1,X_2) = \{1,\ldots,80\}$$

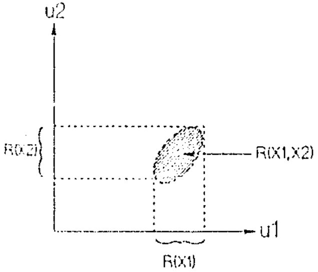


Figura 2.7 Restricciones marginales inducidas por $R(X_1,X_2)$.

Ejemplo 2.2.4. La Figura 2.7. muestra las restricciones en u_1 y u_2 inducidas por $R(X_1,X_2)$.

Una alternativa de describir proyecciones es considerando $R(X_1,\ldots,X_n)$ como una relación en U_1 x ... x U_n, dado $q'=(j_1,\ldots,j_m)$ denota el índice de secuencia complementario para:

$q=(i_1,\ldots,i_k)$, y $R(X_{i1},\ldots,X_{ik}|U_{j1},\ldots,U_{jm})$ -o más compacto $R(X_{(q)}|u_{(q')})$ - denota la restricción en U_{i1} x...x U_{ik}, la cual es condicionada en $U_{j1},\ldots,U_{jm}$. La función característica de esta restricción condicionada es definida por

$$\mu_{R(X_{i1},\ldots,X_{ik}|u_{j1},\ldots,u_{jm})}(u_{i1},\ldots,u_{ik}) =$$
$$\mu_{R(X_1,\ldots,X_n)}\ (u_1,\ldots,u_n) \qquad 2.2.13$$

o más compacta

$$\mu_{R(X_{(q)}|u_{(q')})}(u_{(q)}) = \mu_{R(X)}(u)$$

entendiendo que el argumento $u_{j1},\ldots,u_{jm}$ en (2.2.13) son tratados como parámetros. En consecuencia la restricción condicionada de la función característica es numéricamente igual para $R(X_1,\ldots,X_n)$, y define una relación en U_{i1} x...x U_{ik} lo contrario a U_1 x...x U_n. Entonces por (2.2.9), (2.2.12) y (2.2.13) la proyección de $R(X_1,\ldots,X_n)$ en U_{i1} x...x U_{ik} puede ser expresada como

$$PqR(X_1,\ldots,X_n) = U_{\,u(q')} R(X_{i1},\ldots,X_{ik}|u_{j1},\ldots,u_{jm}) \qquad 2.2.14$$

donde $U_{\,u(q')}$ denota la unión de la familia de restricciones $R(X_{i1},\ldots,X_{ik}|u_{j1},\ldots,u_{jm})$ parametrizado por $u_{(q')} \triangleq (u_{j1},\ldots,u_{jm})$. En consecuencia (2.2.14) implica que la restricción marginal $R(X_{i1},\ldots,X_{ik})$ en U_{i1} x...x U_{ik} puede ser expresada como la unión de restricciones condicionadas $R(X_{i1},\ldots,X_{ik}|u_{j1},\ldots,u_{jm})$ es decir,

$$R(X_{i1},\ldots,X_{ik}) = U_{\,u(q')} R(X_{i1},\ldots,X_{ik}|u_{j1},\ldots,u_{jm}) \qquad 2.2.15.$$

o más compactamente

$$R(X_{(q)}) = U_{\,u(q')} R(X_{(q)}|u_{(q')})$$

Ejemplo 2.2.5. Para ilustrar (2.2.15), tenemos que $U_1=U_2 \triangleq \{3,5,7,9\}$ y que $R(X_1,X_2)$ está caracterizada por una relación matricial.

R	3	5	7	9
3	0	0	1	0
5	1	0	1	0
7	1	0	1	1
9	1	0	0	1

en este caso

$$R(X_1,X_2 \mid u_1=3) = \{7\}$$

$$R(X_1,X_2 \mid u_1=5) = \{3,7\}$$

$$R(X_1,X_2 \mid u_1=7) = \{3,7,9\}$$

$$R(X_1,X_2 \mid u_1=9) = \{3,9\}$$

y de aquí

$$R(X_2) = \{7\}\ U\ \{3,7\}\ U\ \{3,7,9\}\ U\ \{3,9\} = \{3,7,9\}$$

2.2.2 Interacción y no Interacción.

Un concepto básico es la interacción entre dos o más variables, el cual es análogo a la dependencia de variables aleatorias. Más específicamente la variable $X=(X_1,...,X_n)$ es asociada con la restricción $R=(X_1,...,X_n)$, la cual induce las restricciones $R(X_1),...,R(X_n)$ en U1,...,Un respectivamente.

DEFINICION 2.2.2. $X_1,...,X_n$ son variables no interactivas bajo $R(X_1,...,X_n)$, si sólo si $R(X_1,...,X_n)$ son separables, es decir

$$R(X_1,...,X_n) = R(X_1) \times ... \times R(X_n) \qquad 2.2.16$$

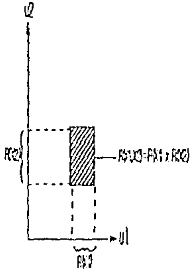
donde $i=1,...,n$.

$$R(Xi)=\mathrm{Proj}R(X_1,...,X_n) \text{ en } U_i$$

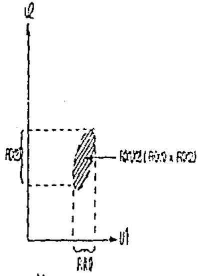
$$= U_{u(q')}R(X_i|u_{(q')}) \qquad 2.2.17$$

con $u_{(q)} \triangleq u_i$ y $u_{(q')} \triangleq$ complemento de u_i en $(u_i,...,u_n)$.

Ejemplo 2.2.6 La Figura 2.8(a) muestra dos variables no interactivas X_1 y X_2 cuyas restricciones $R(X_1)$ y $R(X_2)$ son intervalos; en este caso $R(X_1,X_2)$ es el producto cartesiano de los intervalos en cuestión. En la Figura 2.8(b) $R(X_1,X_2)$ es un propio subconjunto de $R(X_1)$ x $R(X_2)$, y de aquí X_1 y X_2 son interactivos. (En el ejemplo 2.2.3. X_1 y X_2 son interactivos).



a)



b)

Figura 2.8 (a) X_1 y X_2 son no interactivos (b) X_1, X_2 son interactivos.

Si $X_1,\ldots,X_n$ son no interactivos, entonces la n-aría ecuación de asignación

$$(x_1,\ldots,x_n) = (u_1,\ldots,u_n):R(X_1,\ldots,X_n) \qquad 2.2.18$$

puede ser descompuesta en una secuencia de n ecuaciones asignadas

$$
\begin{aligned}
x_1 &= u_1 : R(X_1),\\
x_2 &= u_2 : R(X_2),\\
&\;\vdots\\
x_n &= u_n : R(x_n),
\end{aligned}
\qquad 2.2.19
$$

donde $R(X_i)$, $i=1,\ldots,n$ es la proyección de $R(X_1,\ldots,X_n)$ en U_i y por la Definición 2.2.2., tenemos que

$$R(X_1,\ldots,X_n) = R(X_1) \times\ldots\times R(X_n) \qquad 2.2.20$$

En el caso donde $X_1,\ldots,X_n$ son interactivos, la secuencia de n ecuaciones asignadas asume la forma siguiente

$$
\begin{aligned}
x_1 &= u_1 : R(X_1),\\
x_2 &= u_2 : R(X_2|u_1),\\
&\;\vdots\\
x_n &= u_n : R(X_n|u_1,\ldots,u_{n-1})
\end{aligned}
\qquad 2.2.21
$$

donde $R(X_i|u_1,\ldots,u_{i-1})$ denota la restricción inducida en u_i con $u_1,\ldots,u_{i-1}$. La función característica de ésta restricción condicionada es expresada por

$$\mu_{R(X_i|u_1,\ldots,u_{i-1})}(u_i) = \mu_{R(X_1,\ldots,X_i)}(u_1,\ldots,u_i) \qquad 2.2.22$$

entendiendo que los argumentos $u_1,\ldots,u_{i-1}$ en (2.2.22) juegan el papel de parámetros.

Comentario 2.2.3. Usando la notación (2.2.21), en el caso de variables interactivas, una vez asignado un valor u_1 a x_1, la restricción en u_2 se hace dependiente en u_1. Entonces, la restricción en u_3 se hace dependiente en los valores asignados a x_1 y x_2, y finalmente la restricción en u_n se hace dependiente en $u_1,\ldots,u_{n-1}$. Además (2.2.22) implica que la restricción en u_i dada en $u_1,\ldots,u_{i-1}$ es esencialmente la misma como la restricción marginal en $(u_1,\ldots,u_i)$, con $u_1,\ldots,u_{i-1}$ tratada como parámetros. Esto es ilustrado en la Figura 2.9.

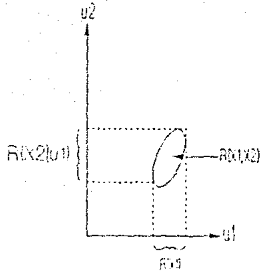


Figura 2.9 $R(X_2|u_1)$ es la restricción en u_2 condicionada en u_1.

En términos de la analogía de la maleta, $X_1,\ldots,X_n$ son no interactivas si las particiones entre los compartimientos llamados $X_1,\ldots,X_n$ son no ajustables. En este caso, puede ser que un lugar en un compartimiento X_i no influye en los objetos que pueden estar en el lugar de otros compartimientos.

En el caso donde las particiones son ajustables, $X_1,\ldots,X_n$ se hacen interactivos en el sentido que la colocación de un objeto u_i en X_i afecta un lugar en los compartimientos complementarios. Desde este punto de vista, la secuencia de las ecuaciones de asignación en (2.2.21) describen el camino, en donde la restricción en los compartimientos X_i son influenciados por la colocación de objetos $u_1,\ldots,u_{i-1}$ en $X_1,\ldots,X_{i-1}$.

2.3 CONJUNTOS DIFUSOS Y EL PRINCIPIO DE EXTENSION.

2.3.1 Notación de Conjuntos Difusos y Terminología.

Un subconjunto difuso A de un universo de discusión U es caracterizado por una función de pertenencia $\mu_A: U \dashrightarrow [0,1]$, la cual asocia a cada elemento de u en U un número $\mu_A(u)$ en el intervalo $[0,1]$, con $\mu_A(u)$ representando el grado de pertenencia de u en A. El soporte de A es el conjunto de puntos en U en donde $\mu_A(u)$ es positivo. Lo grande de A es el supremo de $\mu_A(u)$ sobre U. A puede denotar el conjunto de pares ordenados $(\mu_A(u),u)$. Un punto que atraviesa A es un punto en U donde el grado de pertenencia en A es 0.5.

Es decir un conjunto difuso A se considera como un conjunto de pares ordenados, en los que el primer componente es un número en el rango [0,1] que denota el grado de pertenencia de un elemento u de U en A, y el segundo componente especifica precisamente quien es ese elemento u. En general los grados de pertenencia son subjetivos en el sentido de que su especificación es una cuestión de definición más que un análisis o experimentación objetiva. Se debe aclarar que aunque $\mu_A(u)$ puede interpretarse como el grado de verdad de que la expresión "$u \in A$" sea cierta, es más natural considerarlo simplemente como un grado de pertenencia.

Además se puede notar que:

a) Mientras más próximo esté $\mu_A(u)$ a el valor 1, se dice que u pertenece más A (de modo que 0 y 1 denotan la no pertenencia y la pertenencia completa, repectivamente).

b) Un conjunto en el sentido usual es también difuso pues su función característica

$$\mu_A: u \longmapsto \begin{cases} 1 & \text{si} \quad u \in A \\ 0 & \text{si} \quad u \text{ no pertenece a } A \end{cases}$$

es también una función μ_A: u--->[0,1]; o sea que los conjuntos difusos son generalización de los conjunto usuales.

Ejemplo 2.3.1. El universo de discusión es el intervalo [0,1], con u interpretada como la edad. Un subconjunto difuso de U etiquetado viejo puede ser definido por una función de pertenencia (característica) tal como

$$\mu A(u) = 0 \qquad \text{Para} \quad 0 \le u \le 50$$

$$\mu A(u) = \left[1 + \left(\frac{u-50}{5} \right)^{-2} \right]^{-1} \qquad \text{Para } 50 \le u \le 100 \tag{2.3.1}$$

en este caso, el soporte de viejo es el intervalo [50 , 100]; lo grande de viejo es efectivamente la unidad; y el punto que atraviesa viejo es 55.

Ejemplo 2.3.2. Sea $U=\{1,2,3,4,5,6,7,8,9,10\}$, entonces los siguientes conjuntos definidos son difusos:

Pocos = {(0.4,1), (0.8,2), (1,3), (0.4,4)}

Varios = {(0.5,3), (0.8,4), (1,5), (1,6), (0.8,7), (0.5,8)}

Muchos = {(0.4,6), (0.6,7), (0.8,8), (0.9,9), (1,10)}

se puede notar que el elemento 4 pertenece en grado 0.4 al conjunto Pocos, en grado 0.8 a Varios y en grado 0.0 a Muchos.

Para simplificar la representación de conjuntos difusos empleamos la notación siguiente:

Un conjunto finito no difuso tal como

$$U=\{u_1,\ldots,u_n\} \qquad 2.3.2$$

puede ser expresada como

$$U=u_1 + u_2 +\ldots+ u_n \qquad 2.3.3$$

o

$$U = \sum_{i=1}^{n} u_i, \qquad 2.3.4$$

entendiendo que + denota la unión y no la suma aritmética. Esto es, la notación (2.3.3) puede ser vista como una representación de U, siendo la unión de esta simple constitución.

Como una extensión de (2.3.3), un subconjunto difuso A de U puede ser expresado como

$$A=\mu_1 u_1 +\ldots+ \mu_n u_n \qquad 2.3.5$$

o

$$A = \sum_{i=1}^{n} \mu_i u_i \qquad 2.3.6$$

donde μ_i, $i=1,\ldots,n$, es el grado de pertenencia de u_i en A. En casos donde los u_i son números, éstos pueden dar una relación de ambigüedad en la identidad de μ_i y u_i componetes de la cadena $\mu_i u_i$. En tales casos podemos emplear un símbolo / como separador para quitar lo ambiguo, entonces podemos escribir

$$A = \mu_1/u_1 + \ldots + \mu_n/u_n \qquad 2.3.7$$

o

$$A = \sum_{i=1}^{n} \mu_i/u_i \qquad 2.3.8$$

Ejemplo 2.3.3. Dado U = (a,b,c,d), o equivalentemente,

$$U = a+b+c+d \qquad 2.3.9$$

en este caso un subconjunto difuso A de U puede ser representado sin ambigüedad como

$$A = 0.3a + b + 0.9c + 0.5d \qquad 2.3.10$$

por otro lado, si

$$U = 1 + 2 + \ldots + 100 \qquad 2.3.11$$

entonces podemos escribir

$$A = 0.3/25 + 0.9/3 \qquad 2.3.12$$

esto es para evitar la ambigüedad.

Ejemplo 2.3.4. El universo de discusión comprende los enteros 1, 2,..., 10, es decir

$$U = 1 + 2 + \ldots + 10, \qquad 2.3.13$$

el subconjunto difuso etiquetado varios puede ser definido como

$$\text{varios} = 0.5/3 + 0.8/4 + 1/5 + 1/6 + 0.8/7 + 0.5/8 \qquad 2.3.14$$

Ejemplo 2.3.5. Ahora el universo de discusión

$$U = 0 + 1 + 2 + \ldots \qquad 2.3.15$$

el conjunto difuso etiquetado pequeño puede expresado como

$$\text{pequeño} = \sum_{0}^{\infty} \left[1 + \left(\frac{u}{10} \right)^2 \right]^{-1} / u \qquad 2.3.16$$

De este modo (2.3.3), (2.3.5) pueden ser interpretadas como una representación de un conjunto difuso mediante la unión (como se verá más adelante en 2.3.34), si en la representación de A tenemos que $u_i=u_j$, entonces podemos hacer la sustitución expresada por

$$\mu_i u_i + \mu_j u_i = (\mu_i \vee \mu_j) u_i \qquad 2.3.17$$

Por ejemplo:

$$A = 0.3a + 0.8a + 0.5b \qquad 2.3.18$$

puede ser escrita

$$\begin{aligned} A &= (0.3 \vee 0.8)a + 0.5b \\ &= 0.8a + 0.5b \end{aligned} \qquad 2.3.19$$

Cuando el soporte de un conjunto difuso es continuo y perfectamente contable o es un conjunto finito, podemos escribir:

$$A = \int_U \mu_A(u)/u \qquad 2.3.20$$

con la condición que $\mu_A(u)$ es el grado de pertenencia de u en A, y la integral denota la unión de los simples elementos difusos $\mu_A(u)/u$, $u \in U$.

Ejemplo 2.3.6. El universo de discusión está dado por el intervalo [0 , 100], u = edad, el subconjunto difuso etiquetado viejo (la función de pertenencia está dada por la notación 2.3.1.) puede ser expresada como:

$$\text{viejo} = \int_{50}^{100} \left[1 + \left(\frac{u-50}{5} \right)^{-2} \right]^{-1} / u \qquad 2.3.21$$

evaluando la función de pertenencia en un punto u de este conjunto, tenemos que

$$\mu_{viejo}(u)=0.5 \qquad 2.3.22$$

cuando u=55

Un conjunto difuso A es citado normal si la unidad es lo más grande, esto es ,si

$$\sup_u \mu_A(u) = 1 \qquad 2.3.23$$

en otro caso A es **subnormal**. En este sentido, el conjunto viejo definido por (2.3.21) es normal, como lo es el conjunto de varios definida por (2.3.14). Por otro lado, el subconjunto de U = 1 + 2 +...+ 10 etiquetado no pequeño y no grande definido por

$$\text{no pequeño y no grande} = 0.2/2+0.3/3+0.4/4+0.5/5+0.4/6 +0.3/7+0.2/8 \qquad 2.3.24$$

es subnormal. Puede ser notado que este es un conjunto difuso subnormal, que puede ser **normalizado** por la división μ_A entre el $\operatorname{Sup}_u \mu_A(u)$.

Un subconjunto difuso de U puede ser un subconjunto de otro subconjunto difuso o no difuso de U. Más específicamente, A es un subconjunto de B, o está contenido en B si sólo si $\mu_A(u) \leq \mu_B(u)$ para todo u en U. En símbolos,

$$A \subset B \Longleftrightarrow \mu_A(u) \leq \mu_B(u), \quad u \in U \qquad 2.3.25$$

Ejemplo 2.3.7. Si U = a + b + c +d y

$$A = 0.5a + 0.8b + 0.3d$$
$$B = 0.7a + b + 0.3c + d \qquad 2.3.26$$

entonces $A \subset B$

2.3.2 Niveles de un Conjunto Difuso.

Si A es un subconjunto difuso de U, entonces un **conjunto nivel-α** de A es un conjunto no difuso denotado por A_α, el cual contiene todos los elementos de U donde el grado de pertenencia en A es mayor o igual a α. Entonces tenemos la notación

$$A_\alpha = \{ u \mid \mu_A(u) \geq \alpha \} \qquad 2.3.27$$

Un conjunto difuso A puede ser descompuesto en conjuntos-nivel, por medio de la Identidad de Resolución

$$A = \int_0^1 \alpha A_\alpha \qquad 2.3.28$$

$$A = \sum_\alpha \alpha A_\alpha \qquad 2.3.29$$

donde αA_α es el producto de un escalar α con el conjunto A_α (como se verá en 2.3.36), y

$\int_0^1$ (o Σ) es la unión de la A_α con α en el rango de 0 a 1.

La identidad de resolución puede ser vista como el resultado de la combinación con los términos de (2.3.5), los cuales caen en algún conjunto-nivel. Por ejemplo, suponemos que A es representada de la forma:

$$A = 0.1/2 + 0.3/1 + 0.5/7 + 0.9/6 + 1/9 \qquad 2.3.30$$

Entonces usando (2.3.17), A puede ser escrita como:

$$\begin{aligned} A = 0.1/2 + 0.1/1 + 0.1/7 + 0.1/6 + 0.1/9 \\ + 0.3/1 + 0.3/7 + 0.3/6 + 0.3/9 \\ + 0.5/7 + 0.5/6 + 0.5/9 \\ + 0.9/6 + 0.9/9 \\ + 1/9 \end{aligned}$$

o

$$\begin{aligned} A = 0.1\ & (1/2 + 1/1 + 1/7 + 1/6 + 1/9) \\ + 0.3\ & (1/1 + 1/7 + 1/6 + 1/9) \\ + 0.5\ & (1/7 + 1/6 + 1/9) \\ + 0.9\ & (1/6 + 1/9) \\ + 1\ & (1/9) \end{aligned} \qquad 2.3.31.$$

la cual es la forma (2.3.29), con el conjunto de niveles dado por (2.3.27.) queda

$$\begin{aligned} A_{0.1} &= 2 + 1 + 7 + 6 + 9 \\ A_{0.3} &= 1 + 7 + 6 + 9 \\ A_{0.5} &= 7 + 6 + 9 \\ A_{0.9} &= 6 + 9 \\ A_1 &= 9 \end{aligned} \qquad 2.3.32.$$

Zadeh ha hecho algunas extensiones a los conceptos de conjuntos difusos ordinarios; por ejemplo tenemos los conjuntos difusos de nivel-m y los conjuntos difusos tipo-n. Un conjunto difuso de nivel-m se considera como un universo de discusión al conjunto de conjuntos difusos de nivel-(m-1), sobreentendiéndose que los conjuntos difusos de nivel-1 son conjuntos ordinarios. Para los conjuntos difusos tipo-n, los valores de las funciones de pertenencia son conjuntos difusos de tipo-(n-1) del intervalo [0,1] (en lugar de ser puntos de [0,1]). También los conjuntos difusos tipo-1 son equivalentes a los conjuntos difusos ordinarios.

En seguida, para explicitar un poco más estos dos conceptos anteriores, presento dos ejemplos en donde trato conjuntos difusos de nivel-1 y tipo-1 respectivamente.

Ejemplo 2.3.8. Para U=(a,b,c,d), si se tienen los conjuntos difusos de nivel-1 (ordinarios):

Amarillo = (0.1/a, 0.8/b)
Verde = (0.3/a, 0.2/b, 0.9/d)
Negro = (0.9/a, 0.3/b, 0.7/c, 0.4/d)
Rojo = (0.6/a, 0.4/c)

entonces para V = (Amarillo, Verde, Negro, Rojo):

Color_claro = (0.6/Amarillo, 0.7/Verde, 0.3/Rojo)

Color_obscuro = (1/Negro, 0.8/Rojo)

son conjuntos difusos de nivel-2 (pues los elementos de su universo son conjuntos difusos de nivel-1); y para

W = (Color_claro, Color_obscuro)

Color_alegre = (0.8/Color_claro, 0.3/Color_obscuro)

es un conjunto difuso de nivel-3 (el universo de discusión se compone de conjuntos difusos de nivel-2).

Ejemplo 2.3.9. Si se tienen los siguientes conjuntos difusos tipo-1 en [0,1]:

Alto = (1/0.195, 0.9/0.190, 0.8/0.185, 0.7/0.180, 0.6/0.175)

Regular = (1/0.175, 0.8/0.165; 0.6/0.155)

Bajo = (1/0.155, 0.9/0.145, 0.8/0.135)

Normal = (1/0.85, 0.9/0.80, 0.8/0.75)

Delgado = (1/0.70, 0.9/0.65)

Robusto = (1/1.00, 0.9/0.95)

Excelente = (1.0/1.0, 0.9/0.9, 0.85/0.85, 0.7/0.7, 0.65/0.65)

Bueno = (0.9/0.80, 0.95/0.75, 1.0/0.7, 0.95/0.65, 0.90/0.60) entonces para U=(Catalina,Julieta,Gloria,Sergio,Alejandro):

Altura = (Regular/Catalina, Regular/Julieta, Alto/Gloria, Alto/Sergio, Alto/Marino, Bajo/Alejandro)

Complexión = (Normal/Catalina, Normal/Julieta, Normal/Gloria, Robusto/Sergio,Robusto/Marino, Normal/Alejandro)

Certeza = (Bueno/Catalina, Excelente/Julieta, Bueno/Gloria, Bueno/Sergio, Bueno/Marino, Bueno/Alejandro)

son cojuntos difusos tipo-2 (pues los valores de las funciones de pertenencia son conjuntos difusos tipo-1).

Si ahora se definen los conjuntos difusos tipo-2 en [0,1]:

Tamaño = (Alto/0.195, Regular/0.175, Bajo/0.155)

Fuerza = (Robusto/1.0, Normal/0.85, Delgado/0.70)

Precisión = (Excelente/1.0, Bueno/0.7)

entonces

Característica =(Precisión/Catalina,Precisión/Julieta, Tamaño/Gloria, Fuerza/Sergio, Tamaño/Marino, Tamaño/Alejandro)

es un conjunto difuso tipo-3 (pues los valores de las funciones de pertenencia son conjuntos difusos tipo-2).

2.3.3 Operaciones con Conjuntos Difusos

Al igual que los conjuntos usuales, los conjuntos difusos se pueden operar para formar nuevos conjuntos. En el caso usual, con las operaciones comunes de intersección, unión y complemento, el conjunto de conjuntos de U forman un álgebra booleana, es decir se cumplen las condiciones de asociatividad, conmutatividad, elementos neutros, idempotencia, absorción, distributividad, complemento y las leyes de Morgan.

Las tres operaciones mencionadas se pueden extender de varias formas a conjuntos difusos, de modo que al restringirlas a los conjuntos usuales, coincidan con las comunes. Estas extensiones resultantes satisfacen en general solamente a algunas de las condiciones enlistadas anteriormente, y para mantener la vigencia de alguna, será obligado sacrificar a otras. Dado esto, las operaciones básicas que pueden ejecutarse en conjuntos difusos son las siguientes:

1. **El complemento** de A es denotado por $\neg A$ (o algunas veces por A') y definido por

$$\neg A = \int_U \left[1 - \mu_A(u) \right] / u \qquad 2.3.33$$

la operación de complementación corresponde a la negación. Esto es, si A es una etiqueta para un conjunto difuso, entonces no A puede ser interpretado como $\neg A$. También puede denotarse como:

$$\mu_A c(u) = 1 - \mu_A(u) \quad \text{para toda u}$$

2. **La unión** de conjuntos difusos A y B es denotado por A+B (o más convencionalmente por A U B) y es definido por

$$A + B = \int_U \left[\mu_A(u) \vee \mu_B(u) \right] / u \qquad 2.3.34$$

la unión corresponde a la conectiva o (or). Esto es, si A y B son etiquetas de conjuntos difusos, entonces A o B pueden ser interpretados como A + B. Puede entenderse que $\vee$ (= Max) y entonces se puede escribir

$$\mu_{A \cup B}(u) = \mathrm{Max}(\mu_A(u), \mu_B(u)) \quad \text{para todo u}$$

3. **La intersección** de A y B es denotada por $A \cap B$ y es definida por

$$A \cap B = \int_U \left[\mu_A(u) \wedge \mu_B(u) \right] / u \qquad 2.3.35$$

la intersección corresponde a la conectiva **y (and)**; esto es

$$A \text{ y } B = A \cap B \qquad 2.3.36$$

para este caso podemos entender que $\wedge$ Min y la notación queda

$$\mu_{A \cap B}(u) = Min(\mu_A(u), \mu_B(u)) \quad \text{para toda } u$$

4. El producto de A y B es denotado por AB y es definido por

$$AB = \int_U \mu_A(u)\mu_B(u) \ /u \qquad 2.3.37$$

esto es, A^a, donde a es algún número positivo y puede ser interpretado como

$$A^a = \int_U \left[\mu_A(u) \right]^{\alpha} /u \qquad 2.3.38$$

similarmente, si a es algún número real no negativo tal que $a \ Sup_U \mu_A(u) \leq 1$, entonces

$$aA = \int_U a\mu_A(u) \ /u \qquad 2.3.39$$

Ejemplo 2.3.10. Si

$U = 1 + 2 + ... + 10,$
$A = 0.8/3 + 1/5 + 0.6/6,$
$B = 0.7/3 + 1/4 + 0.5/6,$

entonces

$\neg A = 1/1 + 1/2 + 0.2/3 + 1/4 + 0.4/6 + 1/7 + 1/8 + 1/9 + 1/10,$

$A + B = 0.8/3 + 1/4 + 1/5 + 0.6/6,$

$A \cap B = 0.7/3 + 0.5/6,$

$AB = 0.56/3 + 0.3/6,$

$A^2 = 0.64/3 + 1/5 + 0.36/6,$

$0.4A = 0.32/3 + 0.4/5 + 0.24/6$

Como un caso especial de (2.3.38), la operación de concentración es definida como

$$CON\ (A) = A^2 \qquad 2.3.40$$

la cual es el resultado de aplicar un concentrador a un conjunto difuso A, es un conjunto difuso tal que la reducción en la magnitud del grado de pertenencia de u en A es relativamente pequeño para aquellas u que tienen un alto grado de pertenencia en A, y relativamente grande para las u con baja pertenencia. Especificamente se asumirá que esta operación tiene el efecto de elevar al cuadrado la función de pertenencia de A, o sea:

$$\mu_{CON(A)}(u) = \mu_A^2(u) \text{ para toda } u$$

Ejemplo 2.3.11. Sea U=(1,2,3,4,5,6,7,8) y definimos

Chico=(.9/1, .85/2, .7/3, .5/4, .3/5, .2/6, .1/7)

entonces:

$$\begin{aligned} CON(Chico) &= (.9^2/1,\ .85^2/2,\ .7^2/3,\ .5^2/4,\ .3^2/5,\ .2^2/6,\ .1^2/7) \\ &= (.81/1,\ .72/2,\ .49/3,\ .25/4,\ .09/5,\ .04/6,\ .01/7) \end{aligned}$$

mientras que de **dilatación** es expresada por

$$DIL\ (A) = A^{0.5} \qquad 2.3.41$$

tiene el efecto opuesto a la concentración; o sea el aumento en el grado de pertenencia para los elementos que tienen un grado alto es relativamente menor que aquellos que tienen valores pequeños, especificamente se define por:

$$\mu_{DIL(A)}(u) = (\mu_A(u))^{1/2} \text{ para toda } u$$

Ejemplo 2.3.12.

$$DIL(A) = (0.7^{1/2}/a1,\ 0.1^{1/2}/a3,\ 0.8^{1/2}/a4)$$

las operaciones de concentración y dilatación son útiles en la representación de etiquetas lingüisticas.

5. **Intensificación Contrastante**, se usa para intensificar el contraste existente entre los elementos que tienen un valor de pertenencia pequeño y los que tienen un valor grande; especificamente incrementa los valores de $\mu_A(u)$ que

están sobre 0.5 y disminuyen los que están debajo de este umbral. En otras palabras:

$$\mu_{INT(A)}(u) \geq \mu_A(u) \text{ para } \mu_A(u) \geq 0.5 \qquad \text{y}$$

$$\mu_{INT(A)}(u) \leq \mu_A(u) \text{ para } \mu_A(u) \leq 0.5$$

un ejemplo concreto de este operador es

$$\mu_{INT(A)}(u) = 2\mu_A^2(u) \qquad \text{si } 0 \leq \mu_A(u) \leq 0.5$$

$$\mu_{INT(A)}(u) = 1 - 2(1-\mu_A(u))^2 \quad \text{si } 0.5 < \mu_A(u) \leq 1$$

Ejemplo 2.3.13. Tomado el universo y el conjunto difuso chico del ejemplo 2.3.11., tenemos que

INT(Chico) = (.98/1, .96/2, .882/3, .50/4, .18/5, .08/6, .02/7)

6. Si $A_1,\ldots,A_n$ son subconjuntos de U, y $w_1,\ldots,w_n$ son pesos no negativos adicionados a la unidad, entonces la combinación convexa de $A_1,\ldots,A_n$ es un conjunto difuso de A, donde la función de pertenencia es expresada por la siguiente notación:

$$\mu_A = w_1\mu_{A_1} +\ldots+ w_n\mu_{A_n}$$

en donde + denota la suma aritmética. Este concepto de combinación convexa es útil en la representación de etiquetas lingüísticas como esencialmente, típicamente, etc., las cuales modifican los pesos asociados con los componentes de un conjunto difuso.

En otras palabras es una operación n-aria que combina un conjunto de n conjuntos difusos $A_1,\ldots A_n$ en uno solo, A. Este conjunto difuso A es una combinación ponderada de $A_1,\ldots,A_n$ en el sentido de que la función de pertenencia de A está relacionada con las de $A_1,\ldots,A_n$ por la expresión:

$$\mu_A(u) = w_1(u)\mu_{A_1}(u) +\ldots+ w_n(u)\mu_{A_n}(u)$$

donde $0 \leq w_i(u) \leq 1, \quad i=1,\ldots,n \qquad$ y

$w_1(u) +\ldots+ w_n(u) = 1 \qquad$ para toda u

Ejemplo 2.3.14. Si $u=\{0.8/3,\ 1.0/5,\ 0.6/6\}$ y $v=\{0.7/3,\ 1.0/4,\ 0.5/6\}$ entonces la combinación convexa de u y v con ponderaciones constantes $w_1=0.8$ y $w_2=0.2$ es:

$$A=\{\ [0.8*0.8 + 0.2*0.7]/3,\ [0.8*1.0]/5,\ [0.2*1.0]/4,\ [0.8*0.6 + 0.2*0.5]/6\ \}$$

$$=\{\ 0.78/3,\ 0.2/4,\ 0.8/5,\ 0.58/6\ \}$$

7. Si A1,...,An son subconjuntos difusos de U1,...,Un respectivamente, el producto cartesiano de A1,...,An es denotado por A1 x...x An y es definido como un subconjunto difuso de U1 x...x Un, donde la función de pertenencia es expresada por

$$\mu_{A_1 \times \ldots \times A_n}(u_1,\ldots,u_n)=\mu_{A_1}(u_1) \wedge \ldots \wedge \mu_{A_n}(u_n) \qquad 2.3.42$$

esto lo podemos escribir como

$$A_1 \times \ldots \times A_n=\int_{U_1 \times \ldots \times U_n} \left[\mu_{A_1}(u_1) \wedge \ldots \wedge \mu_{A_n}(u_n)\right] / (u_1,\ldots,u_n) \qquad 2.3.43$$

Ejemplo 2.3.15. Si $U_1=U_2=3+5+7$, $A_1=0.5/3 + 1/5 + 0.6/7$ y $A_2= 1/3 + 0.6/5$, entonces

$$\begin{aligned} A1 \times A2 &= 0.5/(3,3) + 1/(5,3) + 0.6/(7,3) \\ &+ 0.5/(3,5) + 0.6/(5,5) + 0.6/(7,5) \end{aligned} \qquad 2.3.44$$

8. La operación de **difusificación** en general afecta en la transformación de un conjunto no difuso a un conjunto difuso o incrementa las difusificaciones de un conjunto difuso. Es decir, la intensificación tenía por efecto transformar un conjunto difuso A en otro A^* que fuera aproximadamente igual (o menos difuso) a A. Ahora bien, la difusificación tiene el efecto opuesto, ya que transforma un conjunto difuso A en otro A~ aproximadamente igual, pero más difuso. El símbolo "~" se utiliza para representar el papel de difusificador; así por ejemplo si U es el conjunto de los números reales, entonces 3~ representa al conjunto difuso de los números reales que son aproximadamente igual a 3, esto es:

3- = (0.7/2, 0.75/2.5, 0.8/2.6, 0.85/2.7, 0.9/2.8, 0.95/2.9, 1.0/3, 0.95/3.1, 0.9/3.2, 0.85/3.3, 0.8/3.4, 0.75/3.5)

Se consideran dos enfoques a través de los que se puede llevar a cabo la difusificación de un conjunto, y que particularmente es relevante para la definición de etiquetas lingüísticas. La base de estos dos enfoques es un proceso de difusificación puntual que transforma un conjunto difuso de un solo término 1/u en U en un conjunto difuso u~ que está concentrado alrededor de u, y que se escribirá u~=k(u) (Núcleo de difusificación). A menos que se diga lo contrario el grado de pertenencia de u en k(u) se asumirá igual a 1.

Ejemplo 2.3.16. Si U=(1,2,3,4) y A(1/2), y si se desea difusificar A, entonces se dice que k(2)=(0.6/1, 1.0/2, 0.8/3, 0.3/4) es el Núcleo de difusificación.

Se pueden considerar dos casos especiales para la difusificación, al caso 1 se le conoce como difusificación de soporte y que se denota como: $SF(A;K) = (\mu_1/k(u_1),...,\mu_n/k(u_n))$; mientras que el segundo caso de le conoce como difusificación de grado y se denota como:

$$GF(A;K)=(k(\mu_1)/u_1,...,k(\mu_n)/u_n).$$

Ejemplo 2.3.17. Si U=(1,2,3,4), A=(0.8/1, 0.6/2), K(1)=(1.0/1, 0.4/2) y K(2)=(1.0/2, 0.4/1, 0.4/3) entonces:

SF(A;K)=(0.8/(1.0/1, 0.4/2), 0.6/(1.0/2, 0.4/1, 0.4/3))

=(0.8/1, 0.32/2, 0.6/2, 0.24/1, 0.24/3)

=(0.8/1, 0.6/2, 0.24/3)

Ahora, si U=(1,2,3,4), A=(0.8/1, 0.6/2), K(0.8)=(1.0/0.8, 0.6/0.7, 0.6/0.9) y K(0.6)=(1.0/0.6, 0.6/0.5, 0.6/0.7), entonces:

GR(A;K)=((1.0/0.8, 0.6/0.7, 0.6/0.9)/1, (1.0/0.6, 0.6/0.5, 0.6/0.7)/2)

2.3.4 Relaciones Difusas.

Si U es el producto cartesiano de n universos de discurso $U_1,\ldots,U_n$, entonces una n-aría relación difusa R, en U es un subconjunto difuso de U. Como en (2.3.20), R puede ser expresada como la unión de estas simples constituciones difusas $\mu_R(u_1,\ldots,u_n)/(u_1,\ldots,u_n)$, es decir,

$$R = \int_{U_1 \times \ldots \times U_n} \mu_R(u1,\ldots,un)/(u1,\ldots,un), \qquad 2.3.45$$

donde μ_R es la función de pertenencia de R.

Ejemplos comunes de relaciones difusas (binarias) son: muy grande, semejante, es aplicable para, es cerrado para, etc. Por ejemplo, si $U_1=U_2=(-\infty,\infty)$, la relación es cerrada para puede ser expresada por

$$\text{es cerrada para} = \int_{U_1 \times U_2} e^{-a|u1-u2|}/(u_1,u_2), \qquad 2.3.46$$

donde a es un factor de escala.

DEFINICION 2.3.1. Sean U y V dos universos de discusión arbitrarios. Una relación binaria difusa R de U a V es un subconjunto difuso del producto cartesiano UxV, caracterizada por la función de pertenencia μ_R: UxV ---> [0,1], que asocia a cada par (u,v) con su grado de pertenencia $\mu_R(u,v)$ en R.

Esto quiere decir que R puede expresarse como:

$$\mu R=(\mu_R(u_1,v_1)/<u_1,v_1>,\mu_R(u_2,v_2)/<u_2,v_2>,\ldots,\ \mu_R(u_m,v_n)/<u_m,v_n>) \qquad 2.3.47.$$

donde u_i, i=1,2,...,n y v_j, j=1,2,...,n representan a los elementos de U y V repectivamente, y $<u_i, v_j>$ es un elemento de UxV que corresponde al par ordenado de u_i y v_j.

Ejemplo 2.3.18. Sea U=(0,1,5,9), y V=(10,20), entonces la siguiente es una relación difusa:

Mucho_menor_que = (0.8/<0,10>, 0.9/<0,20>, 0.8/<1,10>, 0.9/<1,20>, 0.5/<5,10>, 0.6/<5,20>, 0.1/<9,10>, 0.4/<9,20>)

2.3.5 Principio de Extensión.

El **principio de extensión** para conjuntos difusos es en esencia una identidad básica, la cual permite el mapeo o una relación para ser extendidos los puntos en U para subconjuntos difusos de U. Más específicamente, suponemos que f es un mapeo de U a V, y A es un subconjunto difuso de U, expresado como

$$A = \mu_1 u_1 + \ldots + \mu_n u_n \qquad 2.3.48$$

Entonces el principio de extensión afirma que

$$f(A) = f(\mu_1 u_1 + \ldots + \mu_n u_n) = \mu_1 f(u_1) + \ldots + \mu_n f(u_n). \qquad 2.3.49$$

Es decir la imagen de A bajo f puede ser deducida del conocimiento de las imágenes de $u_1, \ldots, u_n$ bajo f.

Ejemplo 2.3.19. Tenemos

$$U = 1 + 2 + \ldots + 10,$$

y f es la operación del cuadrado. Pequeño es un subconjunto difuso de U definido por

$$\text{pequeño} = 1/1 + 1/2 + 0.8/3 + 0.6/4 + 0.4/5 \qquad 2.3.50$$

usando (2.3.49) tenemos[3]

$$\text{pequeño}^2 = 1/1 + 1/4 + 0.8/9 + 0.6/16 + 0.4/25 \qquad 2.3.51$$

Si el soporte de A es continuo, es decir es

$$A = \int \mu_A(u)/u \qquad 2.3.52$$

entonces la notación del principio de extensión asume la forma siguiente:

$$f(A) = f\left(\int_U \mu_A(u)/u\right) = \int_V \mu_A(u)/f(u) \qquad 2.3.53$$

entendiendo que f(u) es un punto en V, y $\mu_A(u)$ es el grado de pertenencia de f(A), la cual es un subconjunto difuso de V.

---

3 Nota: tener en cuenta que la definición de pequeño^2 es diferente de la forma que se muestra en (2.3.38)

En algunas aplicaciones es conveniente usar la forma modificada del principio de extensión, el cual sigue la forma (2.3.53), por la descomposición de A en componentes de los conjuntos-niveles antes que ser difuso. (por la forma (2.3.28) podemos escribir)

$$A = \int_0^1 \alpha A_\alpha \qquad 2.3.54$$

donde A_α es un conjunto nivel-α de A, la declaración del principio de extensión asume la forma:

$$f(A) = f\left(\int_0^1 \alpha A_\alpha\right) = \int_0^1 \alpha f(A_\alpha) \qquad 2.3.55$$

cuando el soporte de A es continuo, y

$$f(A) = f\left(\sum_\alpha \alpha A_\alpha\right) = \sum_\alpha \alpha f(A_\alpha) \qquad 2.3.56$$

cuando cualquier soporte de A es un conjunto contable o los conjuntos nivel de A forman una colección contable.

Comentario 2.3.2. Escrito en la forma (2.3.53), el principio de extensión extiende el dominio de f puntos en U para subconjuntos difusos de U. Por contraste (2.3.55) extiende el dominio de f subconjuntos no difusos de U para subconjuntos difusos de U.

En muchas de las aplicaciones del principio de extensión, encontramos el siguiente problema: cuando tenemos una n-aría función f, la cual es un mapeo de un producto Cartesiano $U_1 \times \ldots \times U_n$ para un espacio V, y un conjunto difuso (relación) A en $U_1 \times \ldots \times U_n$, el cual es caracterizado por una función de pertenencia $\mu_A(u_1,\ldots,u_n)$ con u_i, $i=1,\ldots,n$, denotando un punto genérico en U_i. Una aplicación directa del principio de extensión (2.3.53) produce la notación siguiente:

$$f(A) = f\left(\int_{U_1 \times \ldots \times U_n} \mu_A(u_1,\ldots,u_n)/(u_1,\ldots,u_n)\right) \qquad 2.3.57$$

$$= \int \mu_A(u_1,\ldots,u_n)/f(u_1,\ldots,u_n).$$

Sin embargo, en muchas instancias A no es las proyecciones $A_1,\ldots,A_n$ en $U_1,\ldots,U_n$ respectivamente. En tales casos asumimos que la función de pertenencia A es expresada por,

$$\mu_{A}(u_1,\ldots,u_n) = \mu_{A_1}(u_1) \wedge \mu_{A_2}(u_2) \wedge\ldots\wedge \mu_{A_n}(u_n) \qquad 2.3.58$$

en donde μ_{A_i}, $i=1,\ldots,n$ es la función pertenencia de A_i. Tomando en cuenta (2.3.45), esto es equivalente a decir que A es el producto Cartesiano de proyecciones,

$$A = A_1 \times\ldots\times A_n$$

es decir, esto implica que A es un conjunto grande cuyas proyecciones en $U_1,\ldots,U_n$ son $A_1,\ldots,A_n$ respectivamente.

Ejemplo 2.3.20. Suponemos que como en el Ejemplo 2.3.19.

$$U_1=U_2= 1 + 2 + 3 +\ldots+ 10$$

y

$$A_1 = 2\sim = \text{aproximadamente } 2 = 1/2 + 0.6/1 + 0.8/3 \qquad 2.3.59$$

$$A_2 = 6\sim = \text{aproximadamente } 6 = 1/6 + 0.8/5 + 0.7/7 \qquad 2.3.60$$

y

$$f(u_1,u_2) = u_1 \times u_2 = \text{producto aritmético de } u_1 \text{ y } u_2$$

usando (2.3.58) y aplicando el principio de extensión como es expresado por (2.3.57), entonces para este caso tenemos:

$$\begin{aligned} 2\sim \times 6\sim &= (1/2 + 0.6/1 + 0.8/3) \times (1/6 + 0.8/5 + 0.7/7) \\ &= 1/12 + 0.8/10 + 0.7/14 + 0.6/6 + 0.6/5 + 0.6/7 \\ &+ 0.8/18 + 0.8/15 + 0.7/21 \\ &= 0.6/5 + 0.6/6 + 0.6/7 + 0.8/10 + 1/12 + 0.7/14 \\ &+ 0.8/15 + 0.8/18 + 0.7/21 \end{aligned} \qquad 2.3.61$$

Esto es, el producto aritmético de los números difusos aproximadamente 2 y aproximadamente 6, es un número difuso dado por (2.3.61).

Más generalmente, * es una operación binaria definida en U x V con valores en W. Esto es, si $u \in U$ y $v \in V$, entonces

$$w = u * v \qquad\qquad w \in W$$

Ahora, suponemos que A y B son subconjuntos difusos de U y V respectivamente, con

$$A = \mu_1 u_1 + \ldots + \mu_n u_n,$$

$$B = v_1 v_1 + \ldots + v_m v_m, \qquad 2.3.62$$

usando el principio de extensión bajo la proposición (2.3.58), la operación * puede ser extendida para subconjuntos difusos de U y V, por la relación definida

$$A * B = \left(\sum_i \mu_i u_i \right) * \left(\sum_j v_j v_j \right)$$

$$= \sum_{i,j} (\mu_i \wedge v_j)(u_i \wedge v_j) \qquad 2.3.63$$

esto lo podemos comprobar mediante el caso cuando A=2~, B=6~ y *=X, como en el ejemplo 2.3.20., la aplicación de (2.3.63) produce la expresión para 2 X 6.

2.3.6 Conjuntos Difusos con Funciones Caracteristicas.

Los conjuntos difusos con funciones de pertenencia son motivados por la unión de cerradura, la cual existe entre el concepto de una lingüística de verdad, con valores tales como verdadero, completamente verdadero, muy verdadero, más o menos verdadero, etc. Por otro lado, los conjuntos difusos en los cuales el grado de pertenencia es especificado en términos lingüísticos tales como bajo, mediano, alto, muy bajo, no bajo y no alto, etc.

Bajo esto suponemos que A es un subconjunto difuso de un universo de discusión U, y los valores de la función de pertenencia μ_A de A son dados en el intervalo [0,1] para los subconjuntos difusos. Para diferenciar tales conjuntos difusos los llamaremos de **Tipo 2**, los conjuntos difusos donde las funciones de pertenencia son mapeados de U para [0,1] son clasificados de **Tipo 1**.

DEFINICION 2.3.2. Un conjunto difuso es de Tipo n n=2,3,..., si los rangos de la función de pertenencia esta sobre los conjuntos difusos de Tipo n-1. Los rangos de una funcion de pertenencia de un conjunto difuso de Tipo 1 estan en el intervalo [0,1].

Para definir las operaciones de complementación, unión, intersección, etc., para conjuntos de Tipo 2, se emplea el principio de extensión.

La expresión de la funcion de pertenencia para la intersección de A y B, donde son subconjuntos difusos de Tipo 1 de U:

$$\mu_{A \cap B}(u) = \mu_A(u) \wedge \mu_B(u) \qquad u \in U$$

si $\mu_A(u)$ y $\mu_B(u)$ son intervalos en [0,1] antes que puntos en [0,1] fijos para u.

$$\mu_A(u) = [a_1, a_2]$$

$$\mu_B(u) = [b_1, b_2]$$

donde a_1, a_2, b_1 y b_2 dependen de u, entonces la aplicación del principio de extensión (2.3.55) para la función $\wedge$ (Min).

$$[a_1, a_2] \wedge [b_1, b_2] = [a_1 \wedge b_1 \, , \, a_2 \wedge b_2] \qquad 2.3.65$$

Esto es, si A y B tienen intervalos valuados en funciones de pertenencia como se muestra en la Figura 2.10, entonces su intersección es una curva en intervalos valuados cuyos valores para cada u esta dado por (2.3.65). Considerar el caso donde, para cada u, μ_A y μ_B son subconjuntos difusos en el intervalo [0,1]. Podemos asumir que estos subconjuntos son convexos, es decir tienen intervalos como conjuntos nivel. En otras palabras, podemos asumir que para cada α en [0,1], los conjuntos nivel-α de μ_A y μ_B son funciones de pertenencia de intervalos valuados Figura 2.11.

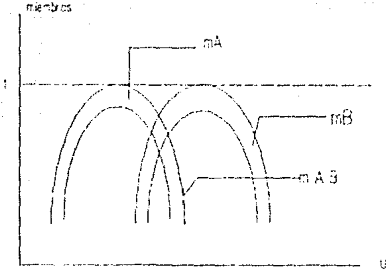


Figura 2.10 Intersección de conjuntos con funciones de pertenencia, intervalos-valuados

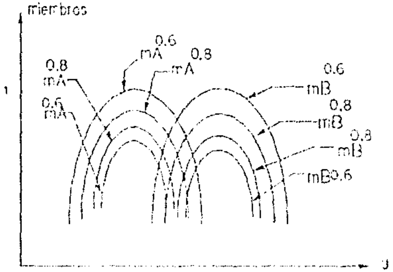


Figura 2.11 Conjuntos-nivel de funciones de pertenencia difuso μ_A y μ_B.

Aplicando la forma del conjunto nivel del principio de extensión (2.3.55) para los conjuntos nivel-α de μ_A y μ_B nos conduce a la siguiente definición de la intersección de conjuntos difusos de Tipo 2.

DEFINICION. 2.3.3. A y B son subconjuntos difusos de Tipo 2 de U tal que para cada $u \in U$, $\mu_A(u)$ y $\mu_B(u)$ son subconjuntos difusos convexos de Tipo 1 de [0,1]. Implican que para cada α en [0,1], los conjuntos de nivel-α de las funciones de pertenencia difusas μ_A y μ_B son funciones de pertenencia $\mu_A{}^{\alpha}$ y $\mu_B{}^{\alpha}$ en intervalos valuados.

Los conjuntos nivel-α de la función de pertenencia difusa de la intersección de A y B denotado por $\mu^{\alpha}_{A\cap B}$, con los conjuntos nivel-α μ_{A}^{α} y μ_{B}^{β} definidos por cada u mediante

$$\mu_A^{\alpha} = \{ v|v_A(v) \geq \alpha \}, \qquad 2.3.66$$

$$\mu_B^{\alpha} = \{ v|v_B(u) \geq \alpha \}, \qquad 2.3.67$$

donde $v_A(v)$ denota el grado de pertenencia de un punto v, $v \in [0,1]$, en el conjunto difuso $\mu_A(u)$, e igualmente para μ_B. Entonces para cada u

$$\mu^{\alpha}_{A\cap B} = \mu_A^{\alpha} \wedge \mu_B^{\alpha} \qquad 2.3.68$$

en otras palabras el conjunto nivel-α de la función de pertenencia difusa de la intersección de A y B es el mínimo (en el sentido de 2.3.65) de los conjuntos nivel-α de las funciones de pertenencia difusas de A y B. Por medio de la identidad de resolución de (2.3.28.) podemos expresar $\mu_{A\cap B}$ como

$$\mu_{A\cap B} = \int_0^1 \alpha(\mu_A^{\alpha} \wedge \mu_B^{\alpha}) \qquad 2.3.69$$

para el caso donde μ_A y μ_B tiene soportes finitos que son μ_A y μ_B de la forma

$$\mu_A = \alpha_1 v_1 + \ldots + \alpha_n v_n, \quad v_i \in [0,1],\ i=1,\ldots,n \qquad 2.3.70$$

y

$$\mu_B = \beta_1 w_1 + \ldots + \beta_m w_m, \quad w_j \in [0,1],\ j=1,\ldots,m \qquad 2.3.71$$

donde α_i y β_j son el grado de pertenencia de v_i y w_j en μ_A y μ_B, respectivamente, la expresión para $\mu_{A\cap B}$ puede ser dada empleando el principio de extensión en la forma (2.3.63), es decir empleando esta forma para la operación de $\wedge$(= Min), obtenemos

$$\begin{aligned}\mu_{A\cap B} &= \mu_A \wedge \mu_B \\ &= (\alpha_1 v_1 + \ldots + \alpha_n v_n) \wedge (\beta_1 w_1 + \ldots + \beta_m w_m) \qquad 2.3.72 \\ &= \sum_{i,j} (\alpha_i \wedge \beta_j)(v_i \wedge w_j)\end{aligned}$$

la expresión deseada para $\mu_{A\cap B}$.

Ejemplo 2.3.21. Como una ilustracion de (2.3.69) suponemos que como un punto **u** el grado de pertenencia de **u** en A y B son etiquetados como **alto** y **medio** respectivamente, con **alto** y **medio** como subconjuntos difusos de V = 0 + 0.1 + 0.2 +...+ 1, definidos por medio de las expresiones

$$\textbf{alto} = 0.8/0.8 + 0.8/0.9 + 1/1 \qquad 2.3.73$$

$$\textbf{medio} = 0.6/0.4 + 1/0.5 + 0.6/0.6 \qquad 2.3.74$$

los conjuntos nivel de alto y medio son expresados por

$$\text{alto}_{0.6} = 0.8 + 0.9 + 1,$$

$$\text{alto}_{0.8} = 0.8 + 0.9 + 1,$$

$$\text{alto}_{1} = 1,$$

$$\text{medio}_{0.6} = 0.4 + 0.5 + 0.6$$

$$\text{medio}_{1} = 0.5$$

y consecuentemente los conjuntos nivel-α de la interseccion estan dados por

$$\begin{aligned}\mu^{0.6}_{A\cap B}(u) &= \text{alto}_{0.6} \wedge \text{medio}_{0.6}\\ &= (0.8 + 0.9 + 1) \wedge (0.4 + 0.5 + 0.6)\\ &= 0.4 + 0.5 + 0.6\end{aligned} \qquad 2.3.75$$

$$\begin{aligned}\mu^{0.8}_{A\cap B}(u) &= \text{alto}_{0.8} \wedge \text{medio}_{0.8}\\ &= (0.8 + 0.9 + 1) \wedge 0.5\\ &= 0.5\end{aligned} \qquad 2.3.76$$

y

$$\begin{aligned}\mu^{1}_{A\cap B}(u) &= \text{alto}_{1} \wedge \text{medio}_{1}\\ &= 1 \wedge 0.5\\ &= 0.5\end{aligned} \qquad 2.3.77$$

combinando (2.3.75), (2.3.76) y (2.3.77), el conjunto difuso representa el grado de pertenencia de u en la intersección de A y B:

$$\mu_{A \cap B}(u) = 0.6/(0.4 + 0.5 + 0.6) + 1/0.5 \qquad 2.3.78$$

$$= \text{medio}$$

lo cual equivale a la declaración

$$\text{alto} \wedge \text{medio} = \text{medio} \qquad 2.3.79$$

Este mismo resultado puede ser obtenido usando (2.3.72), esto es,

$$\text{alto} \wedge \text{medio} = (0.8/0.8 + 0.8/0.9 + 1/1) \wedge$$

$$(0.6/0.4 + 1/0.5 + 0.6/0.6)$$

$$= (0.6/0.4 + 1/0.5 + 0.6/0.6)$$

$$= \text{medio} \qquad 2.3.80$$

Observación 2.3.1. El resultado dado en el Ejemplo 2.3.21. puede ser visto como una conclusión general en donde (2.3.75) respecto a una extensión de desigualdad $\leq$ de números reales para subconjuntos difusos de la línea real. Específicamente, en el caso de números reales a,b, tenemos la equivalencia

$$a \leq b \Longleftrightarrow a \wedge b = a \qquad 2.3.81$$

usando esto como la base $\leq$ para la extensión de intervalos, tenemos que de 2.3.65 y 2.3.81 encontramos:

$$[a_1\ ,\ a_2] \leq [b_1\ ,\ b_2] \Longleftrightarrow a_1 \leq b_1 \ \text{ y } \ a_2 \leq b_2 \qquad 2.3.82$$

esto nos lleva a la siguiente definición.

DEFINICION 2.3.4. Sean A y B dos subconjuntos difusos convexos de la línea real; con A_α y B_α conjuntos en el nivel-α de A y B respectivamente. Entonces una extensión de desigualdad $\leq$ para subconjuntos difusos convexos de la línea real es expresada por

$$A \le B \Longleftrightarrow A \wedge B = A \qquad 2.3.83$$

$$\Longleftrightarrow A_\alpha \wedge B_\alpha = A_\alpha$$

$$\text{para toda } \alpha \text{ en } [0,1] \qquad 2.3.84$$

donde $A_\alpha \wedge B_\alpha$ es definida por ecuación 2.3.65.

En el caso del Ejemplo 2.3.21., es fácil verificar por inspección que:

$$\text{medio}_\alpha \le \text{alto}_\alpha \quad \text{para toda } \alpha \qquad 2.3.85$$

por medio de (2.3.84), podemos concluir que:

$$\text{medio} \wedge \text{alto} = \text{medio} \qquad 2.3.86$$

lo cual concuerda con (2.3.79).

CAPITULO 3

LOGICA DIFUSA

3.1 CONCEPTO DE UNA VARIABLE LINGüISTICA.

Hasta ahora se ha establecido que una variable lingüística difiere de una variable numérica en que sus valores no son números, son palabras o declaraciones en un lenguaje natural o artificial. Las palabras en general son menos precisas que los números, pero el concepto de variable lingüística tiene el propósito de proveer un medio de aproximación del fenómeno, el cual suele ser demasiado complejo para una descripción en términos cuantitativos convencionales.

Los conjuntos difusos representan las restricciones asociadas con los valores de una variable lingüística, que pueden ser vistos como la suma de varias subclases de elementos en un universo de discusión. Por ejemplo, el adjetivo hermoso es una suma de caracteristicas complejas de la apariencia de un individuo; esto puede ser visto como una etiqueta para un conjunto difuso el cual representa un restricción impuesta por una variable difusa llamada hermoso. Desde este punto de vista, entonces, los términos muy hermoso, no hermoso, extremadamente hermoso, completamente hermoso, etc., son nombres de conjuntos difusos que resultan de la operación del conjunto difuso llamado hermoso con los llamados modificadores muy, no, extremadamente, completamente, etc. Estos son conjuntos difusos, junto con el conjunto difuso etiquetado hermoso juega el papel de valores de la variable lingüística apariencia.

Lo importante del concepto de variable lingüística es su estimación de variable de alto orden más que una variable difusa. En el sentido que una variable lingüística toma variable difusas como sus valores. Por ejemplo, los valores de una variable linguistica llamada Edad pueden ser: joven, no joven, viejo, muy viejo, no joven y no viejo, completamente viejo, etc., cada uno de los cuales es el nombre de una variable difusa. Si X es el nombre de tal variable difusa, la restricción impuesta por X puede ser interpretada como el significado de X. De este modo, si la restricción impuesta por la variable difusa llamada viejo es un subconjunto difuso de U = [0 , 100], definida por

$$R(\text{viejo}) = \int_{50}^{100} \left[1 + \left(\frac{u-50}{5} \right)^{-2} \right]^{-1} / u, \qquad u \in U \qquad 3.1.1$$

entonces el conjunto difuso representado por R(viejo) puede ser tomado para el significado de viejo (Figura 3.1).

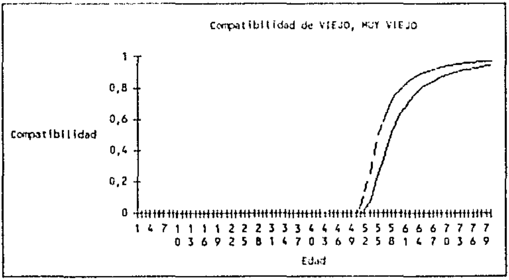


Figura 3.1 Función de compatibilidad de viejo y muy viejo.

En general el concepto de variable lingüística está conformada por una parte estructural y otra de comportamiento que asocia dos reglas:

Regla Sintáctica: la cual puede tener la forma de una gramática para generar los nombres de los valores de la variable.

Regla Semántica: la cual define el procedimiento algoritmico para calcular el significado de cada valor.

Estas dos reglas constituyen la parte esencial de una variable lingüística estructurada.

DEFINICION 3.1.1. Una variable lingüistica es caracterizada por una quintupla (æ, T(æ), U, G, M).

En donde æ es el nombre de la variable. T(æ) (o simplemente T) denota el **conjunto-término** de æ, que es, el conjunto de nombres de **valores lingüísticos** de æ, con cada valor existe una variable difusa denotada genericamente por X y el rango sobre un universo de discusión U, el cual es asociado con la **variable base** u. G es la **regla sintáctica** (usualmente tiene la forma de una gramática) para generar los nombres X, de valores de æ. y M es una **regla semántica** para asociar con cada X su **significado**, M(X) es un subconjunto difuso de U.

En particular X es el nombre generado por G y es llamado un **término**. Un término consiste de una palabra o palabras las que funcionan como una unidad (es decir, siempre ocurren al mismo tiempo) es llamado **término átomico**. Un término que contiene uno o más términos átomicos es un **término compuesto**. Una concatenación de componentes de un término compuesto es un subtérmino. Si $X_1, X_2, \ldots$ son términos en T, entonces T puede ser expresado como la unión

$$T = X_1 + X_2 + \ldots \qquad 3.1.2$$

En donde T es generada por una gramática G, T puede ser escrita como T(G).

El significado, M(X) de un término X es definido por la restricción, R(X) en la variable base, u es impuesta por la variable difusa llamada X. Esto es

$$M(X) = R(X), \qquad 3.1.3$$

entendiendo que R(X) -y por lo tanto M(X)- puede ser vista como un subconjunto difuso de U tomando el nombre de X, la conexión entre æ, el valor linguístico X y la variable base u.

Ejemplo 3.4.1. Considerar una variable lingüistica llamada Edad, esto es æ=Edad, con U=[0,100]. Un valor linguistico de Edad puede ser llamado viejo, con viejo siendo un término átomico. Otro valor puede ser llamado muy viejo, el cual es un término compuesto, que contiene viejo como un componente átomico y tener a muy y viejo como subtérminos. El valor de Edad llamado más o menos joven es un término compuesto el cual contiene joven como término átomico y más o menos es un subtérmino.

El conjunto-término asociado con Edad puede ser expresado como

T(Edad) = viejo + muy viejo + no viejo + más o menos joven + completamente joven + no muy viejo y no muy joven +...

3.1.4

en el que cada término es el nombre de una variable difusa en el universo de discusión $U=[0,100]$. La restricción impuesta por un término, puede ser R(viejo) que constituye el significado de viejo. Esto es, si R(viejo) es definida por (3.1.1), entonces la definición del valor lingüístico viejo está dado por

$$M(\text{viejo}) = \int_{50}^{100} \left[1 + \left(\frac{u-50}{5} \right)^{-2} \right]^{-1} /u, \qquad 3.1.5$$

o más simple

$$\text{viejo} = \int_{50}^{100} \left[1 + \left(\frac{u-50}{5} \right)^{-2} \right]^{-1} /u, \qquad 3.1.6$$

similarmente, el significado de un valor lingüístico tal como muy viejo puede ser expresado como (ver Figura 3.1)

$$M(\text{muy viejo}) = \text{muy viejo} = \int_{50}^{100} \left[1 + \left(\frac{u-50}{5} \right)^{-2} \right]^{-2} /u,$$

3.1.7

la ecuación de asignación en el caso de una variable lingüística asume la forma

$$X = \text{término en } T(æ) = \text{nombre generado por } G \qquad 3.1.8$$

lo cual implica que la definición asignada a X es expresada por

$$M(X) = R(\text{término en } T(æ)), \qquad 3.1.9$$

en otras palabras, el significado de X es dado por la implicación de la regla semántica M para los valores asignados a X por (3.1.8). Además definido por (3.1.3), M(X) es identica a la restricción impuesta por X.

Esto quiere decir, que la ecuación de asignación puede usualmente ser escrita como

$$æ = \text{nombre en } T(æ) \qquad 3.1.10$$

Por ejemplo, si æ = Edad, y viejo es un término en T(æ) podemos escribir

$$\text{Edad} = \text{viejo} \qquad 3.1.11$$

Es importante notar que el símbolo de igualdad en (3.1.10) no representa una relación simétrica como en el caso de una igualdad aritmética. Esto es, la forma definida (3.1.11) no puede ser escrita como

$$\text{viejo} = \text{Edad}$$

Para ilustrar el concepto de una variable lingüística se puede considerar un ejemplo en el cual T(æ) contiene justamente pocos términos y las reglas sintácticas y semánticas son triviales.

Ejemplo 3.1.2. Considerar una variable lingüística llamada Número, la cual es asociada con el conjunto de términos finito

$$T(\text{Número}) = \text{pocos} + \text{varios} + \text{muchos} \qquad 3.1.12$$

en donde cada término representa una restricción en los valores de u en el universo de discusión.

$$U = 1 + 2 + 3 + \ldots + 10 \qquad 3.1.13$$

Estas restricciones son asumidas para ser subconjuntos difusos de U, los cuales son definidos como :

$$\text{Pocos} = 0.4/1 + 0.8/2 + 1/3 + 0.4/4, \qquad 3.1.14$$

$$\text{varios} = 0.5/3 + 0.8/4 + 1/5 + 1/6 + 0.8/7 + 0.5/8, \qquad 3.1.15$$

$$\text{muchos} = 0.4/6 + 0.6/7 + 0.8/8 + 0.9/9 + 1/10 \qquad 3.1.16$$

Esto es

$$R(\text{Pocos}) = M(\text{Pocos}) = 0.4/1 + 0.8/2 + 1/3 + 0.4/4, \qquad 3.1.17$$

y lo mismo para los otros términos en T. La implicación de (3.1.17) en donde pocos es el nombre de una variable difusa, que es un valor de la variable lingüística Número. Entonces el valor pocos asignado a la variable lingüística Número es

$$\text{Número} = \text{pocos} \qquad 3.1.18$$

entendiendo que la asignación de Número es una variable difusa llamada pocos.

Ejemplo 3.1.3. Una variable lingüística compuesta llamada **(æ,Y)** es asociada con la variable base **(u,v)** sobre el rango del universo de discusión **U x V**, donde

$$U \times V = (1 + 2 + 3 + 4) \times (1 + 2 + 3 + 4) \qquad 3.1.19$$

$$= (1,1) + (1,2) + (1,3) + (1,4)$$

$$\begin{matrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{matrix}$$

$$+ (4,1) + (4,2) + (4,3) + (4,4), \qquad 3.1.20$$

con la condición que

$$i \times j = (i,j), \qquad i,j=1,2,3,4 \qquad 3.1.21$$

Además asumimos que el conjunto de terminos de **(æ,Y)** comprende justo dos términos:

$$T = \text{aproximadamente igual} + \text{más o menos igual} \qquad 3.1.22$$

donde aproximadamente igual y más o menos igual son nombres de relaciones difusas binarias definidas por la relación de matrices

$$\text{aproximadamente igual} = \begin{bmatrix} 1 & 0.6 & 0.4 & 0.2 \\ 0.6 & 1 & 0.6 & 0.4 \\ 0.4 & 0.6 & 1 & 0.6 \\ 0.2 & 0.4 & 0.6 & 1 \end{bmatrix} \qquad 3.1.23$$

y

$$\text{más o menos igual} = \begin{bmatrix} 1 & 0.8 & 0.6 & 0.4 \\ 0.8 & 1 & 0.8 & 0.6 \\ 0.6 & 0.8 & 1 & 0.8 \\ 0.4 & 0.6 & 0.8 & 1 \end{bmatrix} \qquad 3.1.24$$

En estas matrices de relación, el elemento (i,j) representa la compatibilidad de la pareja (i,j), con la restricción. Por ejemplo, la entrada (2,3) en aproximadamente igual - es 0.6-, es la compatibilidad de la pareja ordenada (2,3) con la restricción binaria llamada aproximadamente igual.

Para un valor asignado aproximadamente igual para (æ,Y) es escrito como

$$(æ,Y) = \text{aproximadamente igual} \qquad 3.1.25$$

ésta es una asignación para (æ,Y), la cual es una restricción binaria en los valores de (u,v) en el universo de discusión.

Comentario 3.1.2. En términos de la analogía de la maleta; una variable lingüística Definición 3.1.1., puede ser semejante a una maleta dura en la que puede ponerse maletas suaves, como se ilustra es la Figura 3.2. Una maleta suave corresponde a una variable difusa la cual es asignada como un valor lingüístico para æ, con X jugando el papel del título-nombre de la maleta suave.

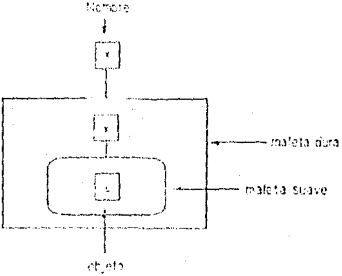


Figura 3.2 Analogía de la maleta para una variable lingüística.

3.1.1 Variables Lingüísticas Estructuradas.

En los ejemplos anteriores, el conjunto-término contiene solamente un número pequeño de términos, de ésta manera es práctica la lista de elementos de T(æ). En un caso más general el número de elementos de T(æ) puede ser infinito, y necesariamente el uso de un algoritmo mejor dicho un procedimiento de una tabla de busqueda, para generar los elementos de T(æ) para calcular sus significados.

Una variable lingüística æ puede ser estructurada si su conjunto-término T(æ), y la función M, asocia un significado a cada término en el conjunto-término, siendo caracterizado algoritmicamanete. En este sentido las reglas sintácticas y semánticas asociadas con una variable lingüística estructurada puede ser vista como procedimientos algoritmicos, para generar los elementos de T(æ) y calcular el significado de cada término en T(æ) respectivamente. De aquí podemos asumir que las variables lingüísticas las entenderemos como estructuradas.

Ejemplo 3.1.4. Para ilustrar el papel que juegan las reglas sintácticas y semánticas en el caso de una variable lingüística estructurada, consideramos una variable llamada **Edad** donde los términos son ejemplificados por: viejo, muy viejo, mucho muy viejo, mucho mucho muy viejo, etc. Esto es, el conjunto de términos de **Edad** puede ser escrito como:

$$T(\text{Edad}) = \text{viejo} + \text{muy viejo} + \text{mucho muy viejo} + \ldots \qquad 3.1.26$$

es claro que por inspección cada término en T(Edad) es de la forma viejo o muy, muy...muy viejo. Para deducir esta regla en una forma más general, procedemos como sigue.

Tenemos xy que denota la concatenación de cadena de caracteres x y y, por ejemplo x=muy, y=viejo, xy=muy viejo. Si A y B son conjuntos de cadenas, por ejemplo,

$$A = x_1 + x_2 + \ldots \quad , \qquad 3.1.27$$

$$B = y_1 + y_2 + \ldots \quad , \qquad 3.1.28$$

donde x_i y y_i son cadenas de caracteres, entonces la concatenación de A y B es denotada por AB y es definida como el conjunto de cadenas.

$$AB = (x_1 + x_2 + \ldots)(y_1 + y_2 + \ldots)$$
$$= \Sigma\, x_i y_i \qquad 3.1.29$$

Ahora, si tenemos A = muy y B = viejo + muy viejo, entonces

muy(viejo + muy viejo) = muy viejo + muy muy viejo

3.1.30

Usando ésta notación, la expresión dada por T(Edad) o simplemente T, puede ser tomada para ser la solución de la ecuación

$$T = \text{viejo} + \text{muy } T \;, \qquad 3.1.31$$

en palabras decimos que cada término en T es de la forma viejo o muy, seguida por algún término en T.

La ecuación (3.1.31) puede ser resuelta por iteración, usando la ecuación de recursión

$$T^{i+1} = \text{viejo} + \text{muy } T^{i} \;, \qquad i=0,1,2,\ldots \qquad 3.1.32$$

con el valor inicial de T^{i} siendo el conjunto de entrada θ. Esto es

$T^{0} = \theta$

T1 = viejo

T2 = viejo + muy viejo 3.1.33

T3 = viejo + muy viejo + muy muy viejo

y la solución de (3.1.31) es dada por

$T = T^{\infty}$ = viejo + muy viejo + muy muy viejo + muy muy muy viejo +... ,

3.1.34

En consideración a la regla sintáctica expresada por (3.1.31) y su solucion (3.1.34), equivalentemente la regla sintáctica puede ser caracterizada por un sistema de producción

T ---> viejo 3.1.35

T ---> muy T 3.1.36

para el cual (3.1.31) juega el papel de una representación algebráica.

En este caso, un término en T puede ser generado a través de un procedimiento de derivación éstandar envolviendo una aplicación sucesiva de las reglas reescritas (3.1.35) y (3.1.36), comenzando con el símbolo T. Esto es, si T es reescrito como **muy** T y entonces T en **muy** T es reescrita como viejo, obtenemos el término muy viejo. En una forma similar el término muy muy muy viejo puede ser obtenido de T por cadena de derivación.

T --->muy T --->muy muy T --->muy muy muy

T ---> muy muy muy muy viejo 3.1.37

Volviendo a la regla semántica para **Edad** notamos que para calcular el significado de un término tal como muy ... muy viejo necesitamos conocer el significado de viejo y el significado de muy. El término viejo juega el papel de un término **primario**, es decir, un término donde el significado debe ser especificado como un dato inicial, para proporcionar una base y calcular el significado de los términos compuestos en T. Como para el término muy, éste actua como una etiqueta lingüística, siendo una modificación del significado de este operando. Si -como aproximación muy simple- asumimos que muy actua como un concentrador, entonces

$$\begin{aligned} \text{muy viejo} &= \mathbf{CON}(\text{viejo}) \\ &= \text{viejo}^2 \end{aligned} \qquad 3.1.38$$

Consecuentemente, la regla semántica para **Edad** puede ser expresada como

$$M(\text{muy},\ldots,\text{muy viejo}) = \text{viejo}^{2n}, \qquad 3.1.39$$

donde n es el número de ocurrencias de muy en el término muy...muy viejo y M(muy...muy viejo) es el significado de muy...muy viejo. Además, si el término primario viejo es definido como

$$\text{viejo} = \int_{50}^{100} \left[1 + \left(\frac{u-50}{5} \right)^{-2} \right]^{-1} /u, \qquad 3.1.40$$

entonces

$$M(\text{muy} \ldots \text{muy viejo}) = \int_{50}^{100} \left[1 + \left(\frac{u-50}{5} \right)^{-2} \right]^{-2n} / u \, ,$$

$$n = 1,2,\ldots \qquad 3.1.41.$$

Esta ecuación proporciona la regla semántica explicita para el cálculo de la definición de los términos compuestos generados por (3.1.31), desde el conocimiento del significado del término primario viejo y la etiqueta muy.

3.1.2 Las Etiquetas Lingüisticas: Operadores.

En el campo de la semántica difusa cuantitativa, al significado de un termino "x" se le representa como un conjunto difuso M(x) del universo de discusión. Desde este punto de vista , uno de los problemas básicos en semántica es que se desea calcular el significado de un término compuesto $x=x_1 \ldots x_n$, partiendo del conocimiento del significado de sus componentes atómicos x_i.

La idea básica sugerida por Zadeh, es que una etiqueta linguística tal como "muy", "mas o menos", "ligeramente", etc., puede considerarse como un operador que actua sobre un conjunto difuso asociado al significado de su operando. Por ejemplo, en el caso del término compuesto "muy alto", el operador "muy" actua en el conjunto difuso asociado al significado del operando "alto". Una representación aproximada para una etiqueta linguística se puede lograr en términos de combinaciones o composiciones de las operaciones básicas. Se debe tomar en cuenta que las representaciones se proponen para ilustrar el enfoque, más que para proporcionar una definición exacta de la etiquetas linguísticas.

Zadeh considera, que las etiquetas linguísticas pueden clasificarse en dos categorias que informalmente se definen como sigue:

tipo I: las que pueden representarse como operadores que actuan en un conjunto difuso: "muy", "mas o menos", "mucho", "ligeramente", "esencialmente", "altamente", "bastante", etc., y

tipo II: las que requieren una descripcion de como actuan en los componentes del conjunto difuso (operando): "esencialmente", "estrictamente", "prácticamente", etc..

En otra palabras, las etiquetas lingüísticas pueden ser caracterizadas como operadores más que construcciones complicadas sobre las operaciones primitivas de conjuntos difusos.

Ejemplos de etiquetas de tipo I.

Considérese A, un conjunto difuso en U, que representa el significado de un término como x=viejo. Ahora, sea x* = muy x =muy viejo, y A* el conjunto difuso que representa el significado de x*. La idea es considerar a la etiqueta lingüística "muy", como un operador que transforma al conjunto difuso A (significado de x) en otro conjunto difuso A* (significado de x*).

De acuerdo a éste punto de vista y sabiendo que el lenguaje natural es muy rico y complejo, tomamos el operador "muy" el cual podemos concretizar con un significado de que aun cuando no tenga validez universal sea sólo una aproximación.

Ahora bien, asumimos que si el significado de un término x es un conjunto difuso A, entonces el significado de muy X es:

A=Chico

Chico = (0.9/1, 0.85/2, 0.7/3, 0.5/4, 0.3/5, 0.2/6, 0.1/7)

$$A^{*} = CON(A) = A^{2}$$

se puede ver en la Figura 3.3, que la concentración del conjunto difuso Chico representa al conjunto difuso muy Chico.

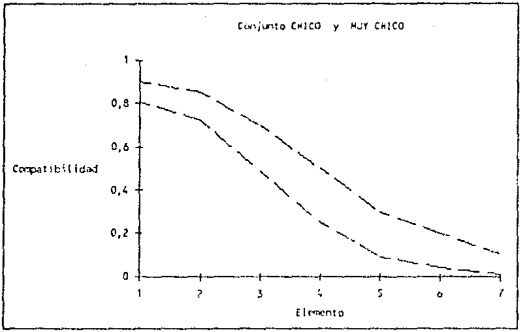


Figura 3.3 Representación de los conjuntos difusos Chico y muy Chico

mas y menos

Se pueden definir etiquetas lingüísticas artificiales, por ejemplo: mas, menos, que son instanceas de lo que puede llamarse acentuador y desacentuador respectivamente, cuya función es proporcionar ligeras variantes de la concentración y la dilatación. Así:

$$\text{mas } x = x^{1.25} \quad \text{y} \quad \text{menos } x = x^{0.75}$$

los exponentes se eligen de modo que se dé la igualdad aproximada: mas mas x = menos muy x

y que además se puedan usar para definir etiquetas lingüísticas cuyo significado difiere ligeramente de otras, por ejemplo:

altamente chico = mas "que" muy chico

altamente chico = mas(CON(chico))

si gráficamos este conjunto podemos observar la Figura 3.4, que muestra la relación entre la etiqueta linguistica muy y la etiqueta altamente.

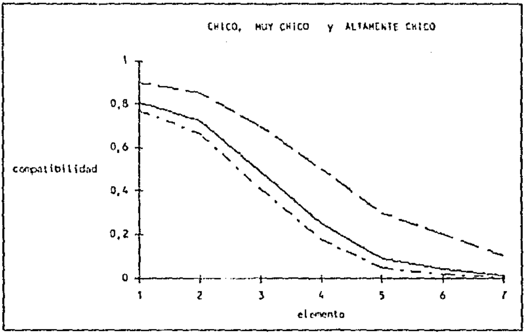


Figura 3.4 Representación de los conjuntos chico , muy chico y altamente chico.

mas o menos

Otra etiqueta lingüística interesante es mas o menos que en sus usos más comunes como mas o menos inteligente, mas o menos rectangular, etc., juega el papel de difusificador.

Suponiendo nuevamente al conjunto difuso chico ya usado antes, y

k(1) = (1/1, 0.9/2, 0.8/3, 0.7/4, 0.6/5),

k(2) = (1/2, 0.85/3, 0.70/4, 0.55/5, 0.40/6),

k(3) = (1/3, 0.8/4, 0.6/5, 0.4/6, 0.2/7)

k(4) = (1/4, 0.75/5, 0.5/6, 0.25/7, 0.01/8), entonces

mas o menos chico = SF(chico;k)=

(.90/1,.85/2,.723/3,.63/4,.54/5,.34/6,.140/7,.005/8)

Gráficando éste conjunto podemos observar el efecto del difusificador k en el conjunto difuso chico y de ésta forma podemos ver el conjunto mas o menos chico Figura 3.5.

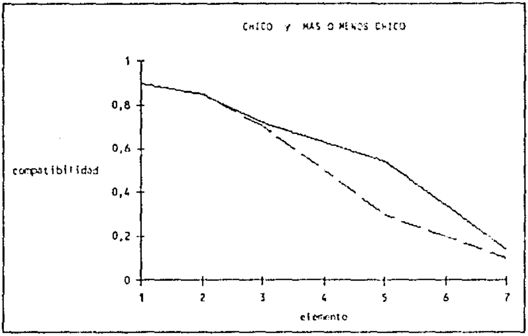


Figura 3.5 Representación del conjunto chico y mas o menos chico

ligeramente

Su efecto es dependiente de la definición de proximidad u ordenamiento en el dominio de su operando. Existen casos, sin embargo, en los que su significado puede definirse en términos de etiquetas lingüísticas tipo I, bajo la suposición de que el dominio del operando es un conjunto ordenado linealmente.

Así:

$$\text{ligeramente } x = \text{norm } (x \cap \neg(x)^2)$$
$$= \text{norm } (x \cap \neg\text{con}(x))$$

ejemplo:

ligeramente chico = norm(chico y no muy chico)

la Figura 3.6, muestra los elementos del universo de discusión para chico y ligeramente chico.

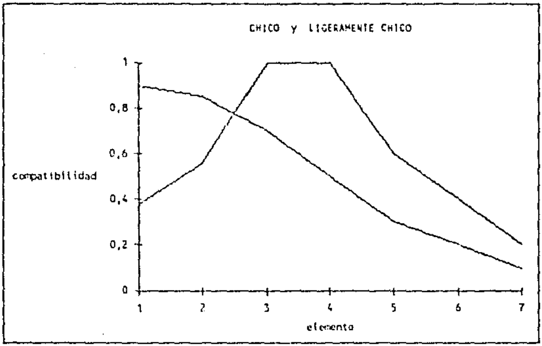


Figura 3.6 Representación de chico y ligeramente chico.

clase de

Es una etiqueta lingüística que tiene el efecto de reducir el grado de pertenencia de los elementos que están en el "centro" (grados de pertenencia grandes) de una clase x e incrementa el de aquellos que están es su periferia (grados de pertenencia pequeños). Se puede tener la aproximación:

$$\text{clase de } x = \text{norm}(\neg(x)^4 \cap x^{1/2}) = \text{norm}(\neg\text{con}^2(x) \cap \text{dil}(x))$$

$$= \text{norm(no muy muy x y mas o menos x)}$$

donde $\neg\text{con}^2(x)$ sirve para reducir el grado de pertenencia de aquellos puntos que están en el centro de la clase, mientras que dil(x) aumenta el grado de pertenencia de los puntos que están en la periferia; como se puede ver la Figura 3.7, y comparandola con la Figura 3.6 que es muy parecida, sin embargo la etiqueta lingüística clase de, encierra más fuertemente a los elementos del conjunto difuso chico que la etiqueta ligeramente; es decir los elementos 1,2,6 y 7 tienen valores mayores de pertenencia en la clase chico que en ligeramente chico.

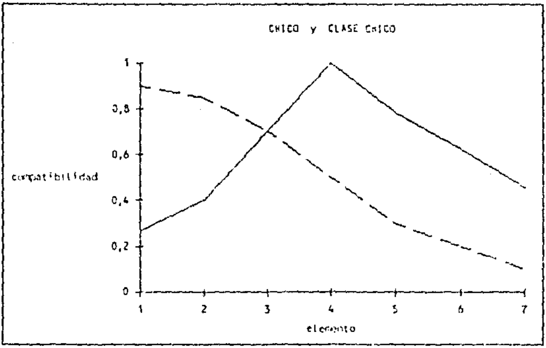


Figura 3.7 Representación del conjunto chico y clase chico

regular

Es una etiqueta que tiene el efecto de reducir el grado de pertenencia de aquellos elementos que tienen tanto un alto grado de pertenencia al conjunto como los de aquellos que lo tienen pequeño, y solo aumenta el de aquellos elementos que tienen un grado de pertenencia cercano al 0.5. Se puede tener la aproximación:

$$\text{regular } x = \text{norm}(\text{int}(x) \cap -(\text{int}(\text{con}(x))))$$

$$= \text{norm}(\text{int}(x) \cap -(\text{int}(\text{muy } x)))$$

Ejemplo de etiquetas tipo II.

Su caracterización envuelve una descripción de la forma en que afectan a los componentes del operando, y por lo tanto es más compleja que las etiquetas de tipo I. En general, la definición de una etiqueta de este tipo debe formularse como un algoritmo difuso que envuelve etiquetas de tipo I.

Ejemplo: etiqueta esencialmente, si se tiene:

decente = (0.4/bueno, 0.3/honesto, 0.2/educado, 0.1/atractivo), como la magnitud de las ponderaciones es una medida de la importancia del atributo asociado, intuitivamente la etiqueta esencialmente tiene el efecto de aumentar las ponderaciones de los atributos importantes y disminuir los que relativamente no lo son.

Se normalizan las ponderaciones:

$$\frac{0.4}{0.4} = 1; \quad \frac{0.3}{0.4} = 0.75; \quad \frac{0.2}{0.4} = 0.5; \quad \frac{0.1}{0.4} = 0.25$$

elevando al cuadrado los pesos normalizados, sumandolos y dividiendo cada peso normalizado por dicha suma:

$$1^2 + 0.75^2 + 0.5^2 + 0.25^2 = 1.87$$

$$\frac{1^2}{1.87} = 0.53; \quad \frac{0.75^2}{1.87} = 0.30; \quad \frac{0.50^2}{1.87} = 0.13; \quad \frac{0.25^2}{1.87} = 0.04$$

entonces:

esencialmente decente = (0.53/bueno, 0.30/honesto, 0.13/educado, 0.04/atractivo).

3.1.3 Variables Linguísticas Booleanas.

La variable lingüística Booleana típicamente envuelve un número finito de términos primarios, un número finito de etiquetas, las conectivas y y o, y la negación no. Por ejemplo, el conjunto de términos de una variable linguística Booleana Edad puede ser

T(Edad) = joven + viejo + no joven + no viejo + muy joven + muy muy joven + no muy joven y no muy viejo + completamente joven + más o menos viejo + extremadamente viejo + ...

3.1.42

Más formalmente, una variable lingüística Boolena puede ser definida como sigue:

Definición 3.1.2. Una variable lingüística Booleana, es una variable linguística donde los términos X son expresiones Booleanas en variables de la forma Xp, hXp, X o hX, en el que h es una etiqueta lingüística, Xp es un término primario y hX es el nombre de un conjunto difuso, resultando de la actuación con h en X.

Como ilustración, en el caso de la variable lingüística **Edad**, donde el conjunto de términos está definido por (3.1.42), el término no muy joven y no muy viejo, con h=muy, Xp=joven y Xp=viejo. Similarmente, en el caso del término muy muy joven , h=muy muy y Xp=joven.

Ejemplo 3.1.5. Edad es una variable lingüística Booleana con el conjunto de términos

$$T(\text{Edad}) = \text{joven} + \text{no joven} + \text{viejo} + \text{no viejo} + \text{muy joven} + \text{no joven y no viejo} + \text{joven o viejo} + \text{joven o (no muy joven y no muy viejo)} + \ldots, \qquad 3.1.43$$

si identificamos y con intersección, o con unión, no con complementación y muy con concentración (ver 3.1.38.), la definición de un valor típico de **Edad** puede ser escrito por inspección. Por ejemplo,

$$M(\text{no joven}) = \neg\text{joven},$$

$$M(\text{no muy joven}) = \neg(\text{joven}^2), \qquad 3.1.44$$

$$M(\text{no muy joven y no muy viejo}) = \neg(\text{joven}^2) \cap \neg(\text{viejo}^2)$$

$$M(\text{joven o viejo}) = \text{joven} \cup \text{viejo}$$

estas ecuaciones expresan el significado de un término compuesto como una función de los significados de sus términos primarios, que lo constituyen. Esto quiere decir, si joven y viejo son definidos como

$$\text{joven} = \int_0^{25} 1/u + \int_{25}^{100} \left[1 + \left(\frac{u-25}{5} \right)^2 \right]^{-1} /u, \qquad 3.1.45$$

$$\text{viejo} = \int_{50}^{100} \left[1 + \left(\frac{u-50}{5} \right)^{-2} \right]^{-1} /u, \qquad 3.1.46$$

entonces

$$H(\text{joven o viejo}) = \int_{0}^{25} 1/u + \int_{25}^{50} \left[1 + \left(\frac{u-25}{5} \right)^{2} \right]^{-1} /u$$

$$+ \int_{50}^{100} \left[1 + \left(\frac{u-25}{5} \right)^{2} \right]^{-1} \vee \left[1 + \left(\frac{u-50}{5} \right)^{-2} \right]^{-1} /u$$

3.1.47

su representación gráfica (ver Figura 3.8).

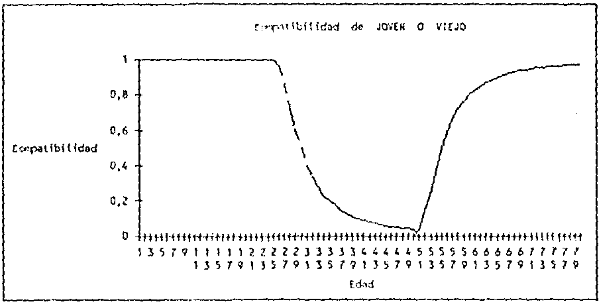


Figura 3.8 Función de compatibilidad para joven o viejo.

Ahora, lo que constituye una apropiada representación para una etiqueta particular, por ejemplo más o menos o completamente o esencialmente, no resulta tan simple. Para ejemplificar este punto con la etiqueta más o menos, usamos la notación (2.3.41).

$$H(\text{más o menos } X) = DIL(X) = X^{0.5} , \qquad 3.1.48$$

Por ejemplo, si X=viejo, y viejo es definido por (3.1.46), entonces

$$\text{más o menos viejo} = \int_{50}^{100} \left[1 + \left(\frac{u-50}{5} \right)^{-2} \right]^{-0.5} /u \qquad 3.1.49$$

Muchas veces más o menos actua como difusificador en el sentido de la notación (2.3.48); ahora nuevamente tomando el significado del término reciente especificado como:

$$\text{reciente} = 1/1974 + 0.8/1973 + 0.7/1972 \qquad 3.1.50$$

y más o menos reciente es definido como el resultado de activar reciente con difusificador SF, es decir,

$$\text{más o menos reciente} = SF(\text{reciente};K) \qquad 3.1.51$$

donde la semilla de SF es definida por

$$K(1974) = 1/1974 + 0.9/1973 ,$$

$$K(1973) = 1/1973 + 0.9/1972 , \qquad 3.1.52$$

$$K(1972) = 1/1972 + 0.8/1971.$$

Sustituyendo los valores de K en (2.3.48), obtenemos la definición de más o menos reciente, es decir

$$\text{más o menos reciente} = 1/1974 + 0.9/1973 + 0.72/1972 + 0.56/1971 \qquad 3.1.53$$

Por otro lado, si asumimos la etiqueta más o menos para ser un dilatador, entonces podemos tener

$$\begin{aligned}\text{más o menos reciente} &= (1/1974 + 0.8/1973 + 0.7/1972)^{0.5} \\ &= 1/1974 + 0.9/1973 + 0.84/1972\end{aligned} \qquad 3.1.54$$

se puede observar que 3.1.53 difiere de 3.1.54 principalmente en que el término **0.56/1971** está ausente. Esto es, si éste término es de importancia en la definicion de más o menos reciente, entonces la aproximación por un dilatador para más o menos no puede ser bueno.

En el ejemplo 3.1.5. tenemos deducida la regla semántica por inspección, tomando las ventajas de la evaluación de expresiones Booleanas. Una técnica más general, usando el método [15,16] el cual es una adaptación de la aproximación empleada por Knuth en [17] para definir la semántica de lenguajes de libre-contexto.

Ejemplo 3.1.6. El conjunto de términos del ejemplo 3.1.5., es generado por una gramática de libre contexto $G=(V_T,V_N,T,P)$, en la cual los no terminales (categorías sintácticas) son denotadas por T, A, B, C, D, es decir,

$$V_N = T + A + B + C + D + E \; , \qquad 3.1.55$$

mientras que el conjunto de terminales (componentes de términos en T) son expresados por

$$V_T = \text{joven} + \text{viejo} + \text{muy} + \text{no} + \text{y} + \text{o} + (+) \qquad 3.1.56$$

y el sistema de producción P, es dado por

$$\begin{array}{lll} T \dashrightarrow A & C \dashrightarrow D & \\ T \dashrightarrow T \text{ o } A & C \dashrightarrow E & \\ A \dashrightarrow B & D \dashrightarrow \text{muy } D & 3.1.57 \\ A \dashrightarrow A \text{ y } B & E \dashrightarrow \text{muy } E & \\ B \dashrightarrow C & D \dashrightarrow \text{joven} & \\ B \dashrightarrow \text{no } C & E \dashrightarrow \text{viejo} & \\ C \dashrightarrow (T) & & \end{array}$$

El sistema de producción P puede sólo ser representado en una forma algebraica, como el conjunto de ecuaciones

$$\begin{aligned} T &= A + T \text{ o } A \, , \\ A &= B + A \text{ y } B \, , \\ B &= C + \text{no } C \, , \\ C &= (T) + D + E \, , & 3.1.58 \\ D &= \text{muy } D + \text{joven} \, , \\ E &= \text{muy } E + \text{viejo}. \end{aligned}$$

La solución de éste conjunto de ecuaciones para T genera el conjunto de términos T expresado por (3.1.43). Similarmente, las soluciones para A, B, C, D y E conjuntos de términos producidos, constituyen las categorías sintácticas denotadas por las mismas respectivamente.

La solución de (3.1.58) puede ser obtenida iterativamente, como en (3.1.32) usando la ecuación recursiva.

$$(T,A,B,C,D,E)^{i+1} = f((T,A,B,C,D,E)^{i}) \ , \ i=0,1,2,... \qquad 3.1.59$$

con

$$(T,A,D,C,D,E)^{\circ} = (\theta,...,\theta)$$

donde (T,A,B,C,D,E) son una séxtupla, cuyos componentes son los no terminales en (3.1.58); f es el mapeo definido por el sistema de ecuaciones (3.1.58); θ es el conjunto de entrada; y $(T,A,B,C,D,E)^i$ es la i-ésima iteración de (T,A,B,C,D,E). La solución de (3.1.58) es el punto fijo de f, dado por $(T,A,B,C,D,E)^{\infty}$. Sin embargo, esto es verdadero para toda i que

$$(T,A,B,C,D,E)^{i} \subset (T,A,B,C,D,E) \ , \qquad 3.1.60$$

lo que significa que cada componente en la sextupla del lado izquierdo de (3.1.60.), es un subconjunto del componente correspondiente a la derecha de misma notación. Entonces ésta notación, genera más y más términos en cada categoría sintáctica T,A,B,C,D,E, como las i iteraciones en (3.1.59).

En una fase más convencional, un término en T puede ser no muy joven y no muy viejo generado por G a través de una sucesión de sustituciones (derivaciones), envolviendo las producciones en P, con cada derivación encadenada empezando con T y terminando en un término generado por G. Por ejemplo, la derivación encadenada para los términos no muy joven y no muy viejo es:

T --> A --> A y B --> B y B --> no C y B --> no D y B --> no muy D y B --> no muy joven y B --> no muy joven y no C --> no muy joven y no E --> no muy joven y no muy E --> no muy joven y no muy viejo.

3.1.61

Esta derivación encadenada puede ser deducida de un árbol de sintaxis (análisis) mostrado en la Figura 3.9, la cual exhibe la frase estructurada del término no muy joven y no muy viejo, en términos de las categorías sintácticas T, A, B, C, D, E. En efecto, el procedimiento para generar los términos en **T**, usando la gramática **G** constituye la regla sintáctica para la variable **Edad**.

La regla semántica para **Edad** es inducida por la regla sintáctica descrita, en el sentido que el significado de un término en T es determinado en parte por el árbol de sintaxis. Especificamente en (3.1.57) es asociado con una relación entre los conjuntos difusos etiquetados por los símbolos terminal y no terminal correspondientes.

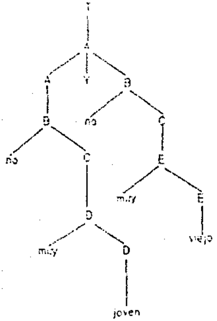


Figura 3.9 Arbol sintáctico para no muy joven y no muy viejo

Para mostrar el resultado del sistema de producción y ecuaciones asociadas usamos los índices I y D para diferenciar los símbolos entre los lados izquierdo y derecho respectivamente (+ = unión).

T --> A	$\Longrightarrow T_I = A_D$,	3.1.62
T --> T o A	$\Longrightarrow T_I = T_D + A_D$,	3.1.63
A --> B	$\Longrightarrow A_I = B_D$,	3.1.64
A --> A y B	$\Longrightarrow A_I = A_D \cap B_D$,	3.1.65
B --> C	$\Longrightarrow B_I = C_D$,	3.1.66
B --> no C	$\Longrightarrow B_I = \neg C_D$,	3.1.67
C --> (T)	$\Longrightarrow C_I = T_D$,	3.1.68

$$C \dashrightarrow D \quad \Longrightarrow C_I = D_D, \qquad 3.1.69$$

$$C \dashrightarrow E \quad \Longrightarrow C_I = E_D, \qquad 3.1.70$$

$$D \dashrightarrow \text{muy } D \quad \Longrightarrow D_I = (D_D)^2, \qquad 3.1.71$$

$$E \dashrightarrow \text{muy } E \quad \Longrightarrow E_I = (E_D)^2, \qquad 3.1.72$$

$$D \dashrightarrow \text{joven} \quad \Longrightarrow D_I = \text{joven}, \qquad 3.1.73$$

$$E \dashrightarrow \text{viejo} \quad \Longrightarrow E_I = \text{viejo}. \qquad 3.1.74$$

Este sistema dual es empleado para calcular el significado de un término compuesto en T.

1. El término en cuestión, no muy joven y no muy viejo es analizado por medio de un algoritmo apropiado para G [18], generando un árbol de sintaxis como se muestra la Figura 3.9, las hojas de éste árbol de sintaxis son (a) términos primarios donde el significado es especificado a priori; (b) modificadores de nombres (es decir, etiquetas, conectivas, negación, etc.); y (c) marcas tales como parentesis los cuales sirven como ayuda para análizar.

2. Empezando de lo profundo (abajo), a los términos primarios se les asigna su significado y usando la ecuación (3.1.62) el significado de no terminales es conectado a las hojas calculadas. Entonces, los subárboles que tienen éstos no terminales como sus raices son borrados. Este proceso es repetido hasta que el significado del término asociado con la raíz del árbol de sintaxis es calculado.

Aplicando éste procedimiento para el árbol de sistaxis mostrado en la Figura 3.10, primero asignamos a joven y viejo los significados expresados por (3.1.45) y (3.1.46), entonces usando (3.1.73) y (3.1.74), encontramos

$$D_7 = \text{joven} \qquad 3.1.75$$

y

$$E_{11} = \text{viejo} \qquad 3.1.76$$

usando (3.1.71) y (3.1.72) obtenemos lo siguiente:

$$D_6 = D_7^2 = \text{joven}^2 \qquad 3.1.77$$

y

$$E_{10} = E_{11}^2 = \text{viejo}^2 \qquad 3.1.78$$

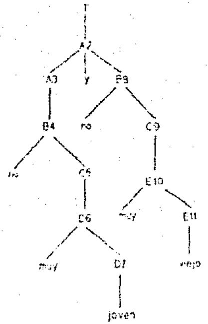


Figura 3.10 Cálculo de la definición de no muy joven y no muy viejo.

continuando de ésta manera obtenemos

$$C5 = D6 = joven^2 \,, \qquad 3.1.79$$

$$C9 = E10 = viejo^2 \,, \qquad 3.1.80$$

$$B4 = \neg C5 = \neg(joven^2) \,, \qquad 3.1.81$$

$$B8 = \neg C9 = \neg(viejo^2) \,, \qquad 3.1.82$$

$$A3 = B4 = \neg(joven^2) \,, \qquad 3.1.83$$

$$A2 = A3 \cap B8 = \neg(joven^2) \cap \neg(viejo^2) \,, \qquad 3.1.84$$

y de aqui

$$\text{no muy joven y no muy viejo} = \neg(joven^2) \cap \neg(viejo^2)$$

que son los que obtuvimos por inspección en (3.1.44).

La idea básica del procedimiento descrito es para relacionar los significados de un término compuesto, para que de éste se constituyan los terminos primarios por medio de un sistema de ecuaciones, el cual sea determinado por la gramática que generen los términos en T.

3.2 VARIABLES LINGÜISTICAS DE VERDAD Y LOGICA DIFUSA.

En discusiones diarias frecuentemente caracterizamos el grado de verdad de una declaración por medio de expresiones tales como muy verdadero, completamente verdadero, más o menos verdadero, esencialmente verdadero, falso, completamente falso, etc. La simulitud entre estas expresiones y los valores de una variable lingüística indican que en situaciones en las cuales lo verdadero o falso de una afirmación *no* puede ser definida, esto puede ser apropiado para tratar verdadero como una variable linguistica, para la cual verdadero o falso son simplemente dos términos primarios de este conjunto de términos, antes de ser un par de puntos extremos en el universo de valores de verdad. Tal variable y sus valores son llamados una **variable lingüística de verdad** y sus **valores lingüísticos de verdad** respectivamente.

Tratando a verdadero como una variable linguistica conduce a una **lógica lingüística difusa**, o simplemente **lógica difusa**, la cual es completamente diferente de la convencional dos-valores o igual a la lógica de n-valores. Esta lógica difusa proporciona una base, que puede ser llamado **razonamiento aproximado**, que es un modo de razonamiento en el cual los valores de verdad y las reglas de inferencia son difusas antes que precisas. En muchos casos, el razonamiento aproximado es parecido al razonamiento humano, que suele ser mal definido o situaciones no confiables.

El término **proposición** puede ser empleado para denotar declaraciones de la forma "**u es A**", donde **u** es un nombre de un objeto y **A** es el nombre de un posible subconjunto difuso de un universo de discusión U, por ejemplo "Juan es joven", "X es pequeño", "la manazana es roja", etc. Si **A** es interpretada como un predicado difuso[1], entonces la declaración "**u es A**" puede ser traducida como "**u es propiedad de A**". Equivalentemente, "**u es A**" puede ser interpretada como una ecuación de asiganción, en la cual un conjunto difuso llamado A es asignado como un valor a la variable lingüística, que denota un atributo de u, por ejemplo:

---

1 Más pecisamente, un predicado difuso puede ser visto como el equivalente de la función de pertenencia de un conjunto difuso. Para simplificar esta terminología, A y μA puede ser referenciada como un predicado difuso.

Juan es joven ---> **Edad**(Juan) = joven

X es pequeño ---> **Magnitud**(X) = pequeño

la manzana es roja ---> **Color**(manzana) = roja

Una proposición tal como "**u es A**" puede ser asumida para ser asociada con dos subconjuntos difusos:

i) la definición de **A**, **M**(**A**) que es un subconjunto difuso de **U** llamado **A**; y

ii) el valor de verdad de "**u es A**", el cual es definido para ser un posible subconjunto difuso de un universo **V** de valores de verdad. En el caso de dos valores lógicos, **V=T + F** (**T** = verdadero, **F**= falso). Puede asumirse que **V** = **[0 , 1]**.

Un valor de verdad que es un punto en [0,1], (**V**(**A**) = 0.8) puede ser referenciado como un valor de verdad numérico. Los valores de verdad numéricos juegan el papel de valores de verdad de la variable base para la variable lingüística **verdad.** Los valores lingüísticos de verdad pueden ser referenciados como **valores de verdad lingüísticos.** Más específicamente, asumimos que verdad es el nombre de la variable lingüística Booleana en donde el término primario es verdadero, con falso no definido como la negación de verdadero, éste es su imagen espejo con respecto a el punto 0.5 en [0,1]. El conjunto de términos de verdad es asumido como sigue:

T(**verdad**) = verdadero + no verdadero + muy verdadero + mas o menos verdadero + muy muy verdadero + esencialmente verdadero + muy (no verdadero) + no muy verdadero +...+ falso + no falso + muy falso +...+...+ no muy verdadero y no muy falso +... , 3.2.1

en el cual los términos son nombres de los valores de verdad. El significado de un término primario verdadero es asumido para ser un subconjunto difuso del intervalo V=[0,1] caracterizado por una función pertenencia. Más preciso, verdadero puede ser observado como el nombre de una variable difusa cuya restricción es el conjunto difuso representado en la Figura 3.11.

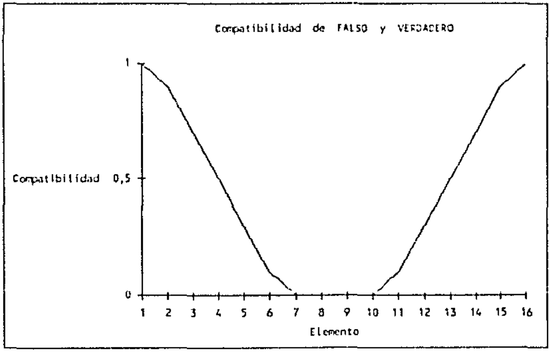


Figura 3.11. Funciones de compatibilidad de valores de verdad lingüísticos verdadero y falso.

Una posible aproximación para la función de pertenencia de verdadero está dada por la expresión

$$\mu_{verdadero}(v) = 0 \qquad \text{para} \quad 0 \le v \le a$$

$$= 2\left[\frac{v-a}{1-a}\right]^2 \qquad \text{para} \quad a \le v \le \frac{a+1}{2}$$

$$= 1 - 2\left[\frac{v-1}{1-a}\right]^2 \qquad \text{para} \quad \frac{a+1}{2} \le v \le 1$$

3.2.2.

en el que $v = \frac{(1+a)}{2}$ está como un punto que atravieza.(Notando que el soporte de verdadero es el intervalo [0,1]). En cuanto falso le corresponde

$$\mu_{falso}(v) = \mu_{verdadero}(1-v) \qquad 0 \le v \le 1.$$

Algunas veces es simple asumir que verdadero es un subconjunto del universo finito de valores de verdad

$$V = 0 + 0.1 + 0.2 + \ldots + 0.9 + 1 \qquad 3.2.3$$

y podemos definir a verdadero como

$$\text{verdadero} = 0.5/0.7 + 0.7/0.8 + 0.9/0.9 + 1/1$$

donde el par 0.5/0.7, por ejemplo, forman la compatibilidad del valor de verdad 0.7 con verdadero 0.5.

En general podemos afirmar que

v(u es: valor lingüístico de una variable Booleana lingüística æ)

= valor lingüístico de una variable lingüística Booleana de verdad τ

3.2.4

como

v(Juan es alto y moreno y bien parecido) = no muy verdadero y no muy falso

donde alto y moreno y bien parecido es un valor lingüístico de una variable llamada æ = Apariencia, y no muy verdadero y no muy falso es una variable lingüística de verdad τ. En forma abreviada tenemos:

$$v(X) = T$$

con X es un valor lingüístico de æ y T es de τ.

3.2.1 Conectivas Lógicas en Lógica Difusa.

Para construir una base para la lógica difusa es necesario entender el significado de operaciones lógicas tal como negación, disyunción, conjunción e implicación para operandos, los cuales tienen valores lingüísticos de verdad antes que numéricos.

Tenemos que A es un subconjunto difuso de un universo de discurso U y $u \in U$, entonces tenemos dos declaraciones:

(a) El grado de pertenencia de u en el conjunto difuso A es $\mu_A(u)$.

(b) El valor de verdad del predicado difuso A es $\mu_A(u)$.

3.2.5.

son equivalentes. Esto es, cuando hacemos la pregunta "¿ Cuál es el grado de pertenencia de u en $A \cap B$ dado el grado de pertenencia difusos de u en A y B?".

Para la respuesta a ésta pregunta usamos el principio de extensión. Especificamente, si $v(A)$ es un punto en $V=[0,1]$ representa el valor de verdad de la proposición "u es A", (o simplemente A), donde u es un elemento de un universo de discusión U, entonces el valor de verdad de no A (o $\neg A$) está dado por:

$$v(\text{no } A) = 1 - v(A) \qquad 3.2.6$$

ahora suponemos que $v(A)$ no es un punto en $[0,1]$, pero si un subconjunto difuso de $[0,1]$ expresado como

$$v(A) = \mu_1/v_1 + \ldots + \mu_n/v_n \qquad 3.2.7$$

donde los v_i son puntos en $[0,1]$ y los μ_i son el grado de pertenencia en $v(A)$. Aplicando el principio de extensión (2.3.59) para (3.2.6) obtenemos la expresión para $v(\text{no } A)$ como un subconjunto difuso de $[0,1]$, es decir,

$$v(\text{no } A) = \mu_1/(1-v_1) + \ldots + \mu_n/(1-v_n) \qquad 3.2.8$$

en particular, si el valor de verdad de A es verdadero, esto es,

$$v(A) = \text{verdadero} \qquad 3.2.9$$

entonces el valor de verdad falso puede ser definido como:

$$\text{falso} = v(\text{no } A) \qquad 3.2.10$$

Por ejemplo, si

$$\text{verdadero} = 0.5/0.7 + 0.7/0.8 + 0.9/0.9 + 1/1 \qquad 3.2.11$$

entonces el valor de verdad de no A está dado por,

$$\text{falso} = v(\text{no } A) = 0.5/0.3 + 0.7/0.2 + 0.9/0.1 + 1/0$$

Comentario 3.2.1. puede notarse que si,

$$\text{verdadero} = \mu_1/v_1 + \ldots + \mu_n/v_n \qquad 3.2.12$$

entonces por (2.3.33)

$$\text{no verdadero} = (1-\mu_1)/v_1 + \ldots + (1-\mu_n)/v_n \qquad 3.2.13.$$

Por contradicción, si

$$v(A) = \text{verdadero} = \mu_1/v_1 + \ldots + \mu_n v_n \ , \qquad 3.2.14$$

entonces

$$\text{falso} = v(\text{no } A) = \mu_1/(1-v_1) + \ldots + \mu_n/(1-v_n) \qquad 2.5.15$$

La misma aplicación para etiquetas. Por ejemplo, la definición de muy (3.1.38)

$$\text{muy verdadero} = \mu_1^2/v_1 + \ldots + \mu_n^2/v_n \qquad 3.2.16$$

por otro lado, el valor de verdad de muy A es expresado por

$$v(\text{muy } A) = \mu_1/v_1^2 + \ldots + \mu_n/v_n^2 \qquad 3.2.17$$

Con conectivas binarias tenemos que v(A) y v(B) son valores de verdad lingüísticos de proposiciones A y B respectivamente. Para simplificar la notación podemos adoptar la convención para escribir -v(A) y v(B) son puntos en [0,1]-

$$v(A) \wedge v(B) \quad \text{para} \quad v(A \text{ y } B)\ , \qquad 3.2.18$$

$$v(A) \vee v(B) \quad \text{para} \quad v(A \text{ o } B)\ , \qquad 3.2.19$$

$$v(A) \Longrightarrow v(B) \quad \text{para} \quad v(A \Longrightarrow B)\ , \qquad 3.2.20$$

$$\neg v(A) \quad \text{para} \quad v(\text{no } A)\ , \qquad 3.2.21$$

entendiendo que $\wedge$, $\vee$ y $\neg$ se reducen a **Min**(conjunción), **Max**(disyunción) y -(negación) operaciones, donde v(A) y v(B) son puntos en [0,1].

Ahora, si v(A) y v(B) son valores de verdad lingüísticos como:

$$v(A) = \alpha_1/v_1 + \ldots + \alpha_n/v_n \ , \qquad 3.2.22.$$

$$v(B) = \beta_1/w_1 + \ldots + \beta_m/w_m \qquad 3.2.23.$$

donde los v_i y w_j son puntos en [0,1] y los α_i y β_j son su repectivo grado de pertenencia en A y B, entonces aplicando el principio de extensión para v(A y B), obtenemos

$$v(A \text{ y } B) = v(A) \wedge v(B)$$
$$= (\alpha_1/v_1 + \ldots + \alpha_n/v_n) \wedge (\beta_1/w_1 + \ldots + \beta_m/w_m)$$
$$= \sum_{i,j} (\alpha_i \wedge \beta_j)/(v_i \wedge w_j)$$

3.2.24

los valores de verdad de A y B son un subconjunto difuso de [0,1] cuyo soporte comprende los puntos $v_i \wedge w_j$, $i=1,\ldots,n$, $j=1,\ldots,m$, con el grado de pertenencia $(\alpha_i \wedge \beta_j)$ respectivos. Hay que tener en cuenta que ésta expresión equivale a la función de pertenencia de la intersección de conjuntos difusos teniendo funciones de pertenencia difusas (2.3.82).

Ejemplo 3.2.2. Suponemos que

$$v(A) = \text{verdadero}$$
$$= 0.5/0.7 + 0.7/0.8 + 0.9/0.9 + 1/1$$

3.2.25

y

$$v(B) = \text{no verdadero}$$
$$= 1/0 + 1/0.1 + 1/0.2 + 1/0.3 + 1/0.4 + 1/0.5 + 1/0.6 + 0.5/0.7 + 0.3/0.8 + 0.1/0.9$$

3.2.26

$$v(A \text{ y } B) = \text{verdadero} \wedge \text{no verdadero}$$
$$= 1/(0 + 0.1 + 0.2 + 0.3 + 0.4 + 0.5 + 0.6) + 0.5/0.7 + 0.3/0.8 + 0.1/0.9$$
$$= \text{no verdadero}$$

3.2.27

En una forma similar, para el valor de verdad de A o B, y obtenemos

$$v(A \text{ o } B) = v(A) \vee v(B)$$
$$= (\alpha_1/v_1 + \ldots + \alpha_n/v_n) \vee (\beta_1/w_1 + \ldots + \beta_m/w_m)$$
$$= \sum_{i,j} (\alpha_i \wedge \beta_j)/(v_i \vee w_j)$$

3.2.28

El valor de verdad de A ==> B depende en la manera en la cual ==> es definida para valores de verdad numéricos. Esto es, si definimos

$$v(A \Longrightarrow B) = \neg v(A) \vee v(A) \wedge v(B) \qquad 3.2.29$$

para el caso donde v(A) y v(B) son conjuntos en [0,1], aplicamos el principio de extensión y tenemos que

$$v(A \Longrightarrow B) = [(\alpha_1/v_1 + \ldots + \alpha_n/v_n) \Longrightarrow (\beta_1/w_1 + \ldots + \beta_m/w_m)]$$

$$= \sum_{i,j} (\alpha_i \wedge \beta_j)/(1-v_i) \vee (v_i \wedge w_j) \qquad 3.2.30$$

para el caso donde v(A) y v(B) son subconjuntos difusos de [0,1].

Comentario 3.2.3. Es importante tener claro la diferencia entre y en verdadero y no verdadero, con $\wedge$ en verdadero $\wedge$ no verdadero. En la forma que concierne con el significado de términos verdadero y no verdadero, e y es definido por la relación

$$M(\text{verdadero y no verdadero}) = M(\text{verdadero}) \cap M(\text{no vedadero}) \qquad 3.2.31$$

donde M es la función de mapeo de un término en (Definición 3.1.1.). Por contraste, en el caso de verdadero $\wedge$ no verdadero son afectados con el valor de verdad de verdadero $\wedge$ no verdadero, el cual es derivado de la equivalencia (3.2.19.).

$$v(A \text{ y } B) = v(A) \wedge v(B) \qquad 3.2.32.$$

Esto quiere decir que , en (3.2.31) $\cap$ es la operación de intersección de conjuntos difusos, ya que en (3.2.32) $\wedge$ es de conjunción. Para ilustrar la diferencia tenemos que dado V= 0 + 0.1 +...+ 1, y P y Q son subconjuntos difusos de V definidos por

$$P = 0.5/0.3 + 0.8/0.7 + 0.6/1 \qquad 3.2.33$$
$$Q = 0.1/0.3 + 0.6/0.7 + 1/1 \qquad 3.2.34$$

entonces

$$P \cap Q = 0.1/0.3 + 0.6/0.7 + 0.6/1 \qquad 3.2.35$$

ya que

$$P \wedge Q = 0.5/0.3 + 0.8/0.7 + 0.6/1 \qquad 3.2.36$$

CAPITULO 4

CARITAS DE CHERNOFF.

4.1 GENERALIDADES.

En la década de los cincuentas se mostró gran interés en la representación gráfica de datos teniendo como fundamentos aspectos analíticos y conceptuales para la existencia de este campo. Sin embargo no es hasta la década de los setentas cuando empieza la atención a las gráficas como una disciplina intelectual con grandes aplicaciones. Es posible que las técnicas de graficación fueron usadas comunmente por científicos e ingenieros antes que el trabajo desarrollado por Playfair enfocado a aplicaciones estadisticas.

Ahora bien, es claro que con el desarrollo tecnológico de las computadoras y su considerable impacto se despierta un gran interés por las gráficas. La computadora tiene la capacidad de producir una gran cantidad de resultados, pero a su vez presenta el problema de resumir grandes cantidades de datos, entonces surge ahora la necesidad de buscar los métodos gráficos que permitan resumir y presentar la información.

Dado este problema, ahora el compromiso de los investigadores en el área de computación es buscar modelos que no caigan en regresiones tediosas, en contradicción de datos y modelos inapropiados; una representación gráfica debe cumplir con ciertos requisitos para ser bien aceptada, es decir, su presentación debe tener un eficaz interés creativo y ser atractiva para la atención del usuario, y de esta manera mantener una relación visual que sea claramente comprendida y fácil de recordar, así mismo los datos resumidos se puedan visualizar con una sola mirada ya que en una imagen completa y bien balanceada se pueden encontrar relaciones escondidas para analizar e invetigar.

La representacion gráfica mediante caritas fué propuesta por Herman Chernoff en 1971 como un nuevo modelo gráfico para representar datos multivariados fácil de examinar. La gráfica es un icono de cara cuyos rasgos faciales: ojos, nariz, boca etc.; son modificados en forma y tamaño para representar vectores estadísticos (Figura. 4.1).

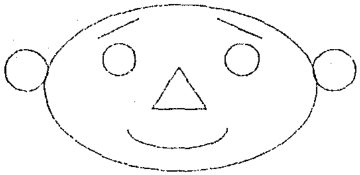


Figura 4.1 Carita de Chernoff

Estas caritas representan un punto en un espacio de k-dimensiones (k es un escalar el cual define el número de puntos) y cambia de posición de acuerdo a los cambios de las variables.

La representación gráfica de los datos multivariados tienen como propósito mejorar la percepción del analista, para detectar y comprender un fenómeno importante. Sirviendo a su vez, como un medio mnemotécnico para recordar mejor las conclusiones y de esta forma sea mayor la comunicación de éstas a otros. El uso de una carita ha proporcionado una riqueza mayor dentro de las técnicas gráficas, debido a que para lograr sus objetivos se basa en la observación del dibujo, enseñando a reconocer y responder a los cambios en las expresiones faciales.

Cuando el analista observa una carita examina los datos en términos de grupos (clusters) y no como una distribución uniforme. La agrupación es precisamente el método que proporciona a las gráficas el poder analítico.

En el programa de dibujo de caritas los datos multivariados que tienen propiedades similares pueden sólo ser representadas por caritas similares. Los grupos (clusters) similares son llamados simplemente familias.

4.2 TECNICAS PARA EL DESPLEGADO DE DATOS MULTIVARIADOS.

Existen varios métodos para representar una gráfica de datos multivariados, $x=(x_1,x_2,...,x_k)$ donde k es la dimensión del vector descrito. Encontramos el método de perfiles lineales tal vez uno de los más simples de la representación gráfica de puntos multivariados. Este representa un punto por medio de una linea poligonal la cual conecta los distintos topes correspondientes a los valores de las variables, que son ordenadas a lo largo de una linea horizontal o vertical. Es decir, una variación consiste en usar una línea poligonal conectando los puntos de las coordenadas (i, xi). Ver la Figura 4.2.

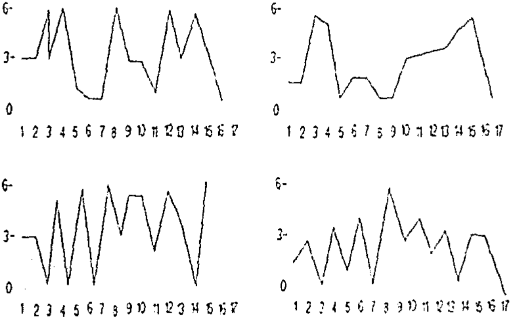


Figura 4.2 Representación de Perfiles Lineales.

También tenemos el método de los perfiles circulares que es una variación de los perfiles lineales, en donde la linea poligonal conecta puntos localizados en rayas igualmente espaciadas y la distancia del centro representa el valor para cada una de las variables como se muestra en la Figura 4.3.

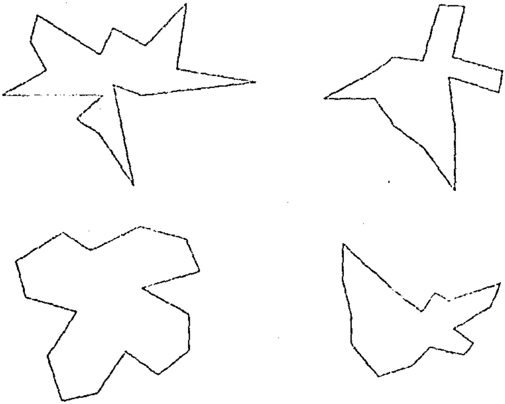


Figura 4.3 Representación de Perfiles Circulares.

Otro método es el de las **Estrellas** donde el polígono conecta los puntos de las coordenadas polares formando círculos, es decir, suponemos que cada unidad de datos consiste de $p\geq 2$ observaciones no negativas; en dos dimensiones podemos construir círculos de radio fijo con p rayas expandidas que salen del centro de éste. La longitud de las rayas representa los valores de las variables, los puntos finales de las rayas pueden conectarse con líneas rectas para formar la estrella, la cual representa una observación multivariada que forma un grupo de acuerdo a similitudes.

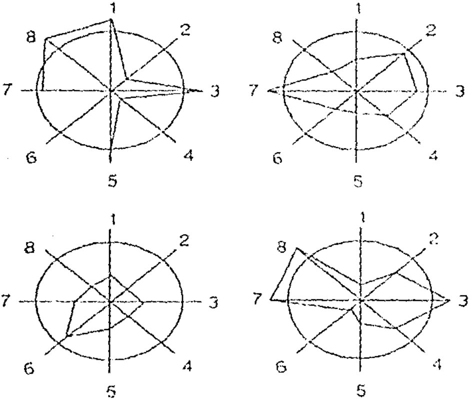


Figura 4.3 Método de las Estrellas.

Andrews introduce la representación de Series de Fourier, este método es mejor conocido como curvas de seno. Este consiste de un vector de p-dimensiones tal que [x1,x2,...,xp] medidas son representadas por las series de Fourier finitas:

$$(x_1, x_2, \ldots, x_k) \dashrightarrow \frac{1}{\sqrt{2}} x_1 + x_2 \sin t + x_3 \cos t + x_4 \sin 2t +$$

$$x_5 \cos 2t + \ldots$$

las medidas son coeficientes para formar una gráfica periódica en el rango $-\pi \leq t \leq \pi$, las representaciones de las observaciones multivariadas por medio de las series pueden ser visualmente agrupadas, teniendo en cuenta que en una gráfica las curvas pueden ser limitadas a 5 o 6.

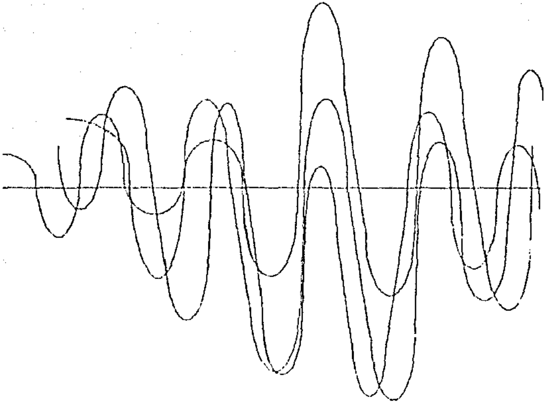


Figura 4.5 Método del Andrews curvas de seno.

Casi al mismo tiempo aparecen las figuras llamadas **metroglyphs** un código particular escrito por Herbert T. Davis para Laboratorios de Sandia Alburquerque Nuevo México.

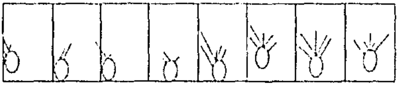


Figura 4.6 Método de Metroglyphs.

Finalmente Chernoff sigirió la representación de p observaciones, desarrollando un programa por computadora el cual dibuja la caricatura de una carita determinada por 18 paramétros tales como la longitud de la nariz, curvatura de la boca, tamaño de los ojos, etc. Designanado $k \leq 18$ variables para cada carita, entonces la cara resultante representa x.

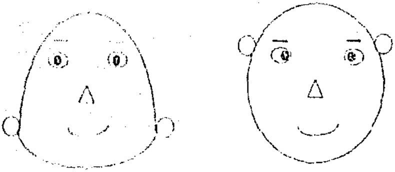


Figura 4.7 Método de Caritas de Chernoff.

De los métodos mencionados los más usuales han sido curvas del seno, las figuras metroglyphs y las caritas de Chernoff.

4.3 VANTAJAS Y DESVENTAJAS SOBRE EL USO DE CARITAS.

Cada una de las tecnicas usadas para desplegar datos multivariados tienen sus ventajas y desventajas. El método de las caritas de Chernoff es de gran utilidad para analizar, dado que su representación presenta un gesto, el cual es llamado psicológicamente expresión de ideas u objetos; las series de Fourier y los perfiles son menos faverables en respuestas emocionales, sin embargo éstas representan algunas ventajas de cálculo. Ahora bien, cuando es presentada una carita podemos decir que contamos con un lenguaje común para nuestras discusiones. Nosotros hablamos de nariz grande, de oreja levantada y esto no resulta tan confuso; dado que podemos asociar por ejemplo, una gran sonrisa a un buen negocio, a un paciente sano y esto resulta facilmente entendido en la observación y así mismo dar una traducción.

El hecho de que una sonrisa puede ser usada para representar una variable de "éxito/fracaso", los ojos puedan representar una posición política, la frente puede representar la inteligencia; pero también esto trae como consecuencia la interpretación de caras subjetivas y precisamente la subjetividad es la distinción de la metodología de las caras con otras técnicas. El tipo de representación metroglyphs parece objetivo pero no establece; hay que tener en cuenta que las caritas no son propuestas como un método para tomar decisiones finales por el contrario es un medio para el estudio de los datos, y si bien es cierto lo subjetivo de las caritas atrae la atención de los usuarios generando conclusiones y esto es bueno, ya que si la tecnica aparenta objetividad ocasiona que el usuario la trate finalmente como tal y esto puede ser una desventaja.

Otra ventaja es la posibilidad de concentrar un subconjunto de variables de datos para hacer las gráficas, la concentración es siempre deseable asociando las variables con los rasgos faciales. Ciertamente uno de los más serios problemas con las caritas de Chernoff, es que su construcción depende de los rasgos faciales y esto puede ocasionar la distorsión en la representación de datos causando impresiones erróneas. Si todos los rasgos faciales son independientes esto no garantiza igualdad y en realidad no es muy probable que la carita pueda ser vista como la unión de las diferentes variables independientes. La dependencia de los rasgos faciales es un serio problema por ejemplo, la posición de las pupilas depende de otros rasgos faciales para garantizar que estas están en los ojos, la estructura de la boca depende de la cara grande y ancha, las orejas al mismo nivel, la nariz larga y la forma de cara son de gran influencia para la boca; así como también los ojos grandes dependen de la nariz larga y la cara grande. Estas dependencias deben ocurrir en órden para asegurar las posiciones correctas de los rasgos faciales.

En el programa original de Chernoff las caras son normalizadas en constantes como cara grande y ancha. La normalización reduce las dependencias pero no las elimina. La restricción de los rangos de los rasgos faciales reduce un poco las dependencias.

En general, se puede decir que cada aplicación que contempla un método gráfico proporciona un conjunto de atributos, que se asocian consecuentemente a las aplicaciones. Con las caritas de Chernoff como representación gráfica tenemos:

Ilustración y comunicación.- Las caritas sirven como un medio de comunicación de la información de forma ilustrativa a un auditorio; considerando que ésta misma información ha sido previamente análizada y entendida. Si un importante aspecto es la comunicación las caritas llegan a tener un fuerte impacto.

Cálculo.- En el tratamiento de datos multivariados las caritas son muy útiles debido a que la representación se lleva a cabo sobre los rasgos que recuerdan psicológicamente objetos e ideas. Cuando una gran cantidad de información es comunicada es importante presentar algo que sea recordado. La precisión en los datos suele ser una de las controversias; pero, las caritas no son propuestas como un método para decisiones finales, se les considera como un medio para agrupar datos que de forma compacta se representan en el icono de cara.

Análisis.- Si los datos no han sido comprendidos, con frecuencia utilizamos representaciones que permitan el análisis para el entendimiento de la existencia de relaciones y regularidades. Cuando los datos son complejos para ser análizados las caritas ofrecen una nueva técnica para comprensión de forma auto explicatoria. La multidimensionalidad de los datos requiere de cierto número de indicadores para las variables, resolviéndolo en las caras con los rasgos faciales.

4.4 ANALISIS MULTIVARIADO.

La investigación científica es un proceso de aprendizaje iterativo; los objetivos se refieren a la explicación de un fenómeno físico o social, que debe ser especificado y probado por la acumulación y análisis de datos. Un análisis de datos acumulados por medio de la experimentación u observación puede originar una explicación modificada del fenómeno. Durante todo este proceso de aprendizaje las variables frecuentemente son adicionadas o borradas del estudio; y esto es lo complejo de muchos fenómenos ya que se requiere de un investigador para coleccionar (recoger) observaciones de muchas variables diferentes, ahora bien, los datos incluyen medidas simultáneas en todas la variables y la necesidad de entender las relaciones entre éstas ocasiona lo que es llamado **análisis multivariado**.

Los datos multivariados consisten de la observación de varias variables diferentes para un número de individuos u objetos; y datos de este tipo se presentan en todas las ramas de la ciencia, desde psicología hasta biología; por tal razón los métodos de análisis de datos multivariados constituyen un incremento importante en el área de la estadística. Entonces formalmente podemos definir el análisis multivariado como una parte de la estadística la cual se ocupa de la relaciones entre conjuntos de variables dependientes e individuos que conducen a otras.

Un ejemplo de análisis multivariado puede ser los resultados de exámen; cuando los estudiantes hacen varios exámenes se obtienen por cada estudiante un conjunto de ellos, las variables son las diferentes materias y los individuos son los estudiantes. El análisis de datos de éste tipo es usualmente simple, los promedios son calculados para cada variable y para cada individuo, pero no sólo se pueden sacar estos resultados sino también calcular el promedio por materia y el promedio de todos los estudiantes e incluso hacer comparaciones de promedios por materia.

La característica distintiva del análisis multivariado es la consideración de un conjunto de n objetos, en cada uno de los cuales son observados los valores de las p variables. El conjunto de objetos puede ser completo o sólo una muestra de un conjunto grande, las variables pueden ser continuas o discontinuas y lo mismo puede ser un subconjunto de un conjunto grande.

El hecho de denotar el número de variables por p, y el número de individuos u objetos por n, nos conduce a decir que tenemos un total de (nxp) medidas; usualmente podemos referirnos a

$$x_{rj} = \text{r-esima observación en la j-esima variable}$$

$$(r=1,\ldots,n \; ; \; j=1,\ldots,p)$$

esto es una matriz con elementos x_{rj}, que puede ser llamada la matriz de datos y que puede ser denotada por X es decir,

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{23} & \cdots & x_{2p} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ x_{n1} & & \cdots & x_{np} \end{bmatrix}$$

esto es muy práctico dado que la matriz de datos puede tener n renglones y p columnas. El valor de n usualmente es mucho más grande que el valor de p, y la forma obvia para registrar los datos es por medio de un vector renglón para cada individuo, generalmente la matriz de datos puede verse como un vector de n renglones o como un vector de p columnas

$$X = \begin{bmatrix} x_1^T \\ x_2^T \\ \cdot \\ \cdot \\ \cdot \\ x_n^T \end{bmatrix} = \begin{bmatrix} y_1, & y_2, & \cdots & ,y_p \end{bmatrix}$$

donde x_i^T denota la transpuesta de x_i.

En general podemos decir que las medidas (comunmente llamados datos) frecuentemente son colocados y desplegados en distintas formas, sin embargo lo más común es usar arreglos para estos casos. Los datos mutivariados aparecen siempre en todas las investigaciones buscando entender un fenómeno físico o social, seleccionando un número $p \geq 1$ de variables o caracteres para un registro. Los valores de estas variables son todas registradas por cada item (campo) distinto individual o ensayo experimental. Entonces, dada ésta situación se puede usar la notación de matrices donde x_{rj} indica el valor particular de la r-esima variable, que es observada en el j-esimo item o ensayo, consecuentemente n medidas en p variables pueden ser desplegadas como la notación matricial anterior donde los renglones son las variables y las columnas son los items.

Por ejemplo, tenemos la selección de 4 ingresos (recibos) de un libro universitario, con la finalidad de investigar la naturaleza de venta del libro. Cada recibo prevee entre una y otra cosa el número de libros vendidos en su totalidad y el total entre cada venta. Entonces, se puede observar el número correspondiente en los recibos como 4 medidas en dos variables:

variable 1 (venta dollar):	42	52	48	58
variable 2 (núm. de libros):	4	5	4	3

usando la notación mencionada tenemos:

$x_{11} = 42 \qquad x_{12} = 52 \qquad x_{13} = 48 \qquad x_{14} = 58$
$x_{21} = 4 \qquad x_{22} = 5 \qquad x_{23} = 4 \qquad x_{24} = 3$

y entonces el arreglo de datos X es:

$$X = \begin{bmatrix} 42 & 52 & 48 & 58 \\ 4 & 5 & 4 & 3 \end{bmatrix}$$

con dos renglones y cuatro columnas.

Considerando los datos en forma de arreglos facilita la exposición y permite los cálculos numéricos para ser ejecutados en orden y de manera eficiente. La eficiencia se logra, primero con la descripción de cálculos numéricos como operaciones en arreglos y segundo su implementación en computadoras, dado que en la actualidad existen muchos lenguajes y paquetes estadísticos para calcular operaciones en arreglos. El estudio de métodos multivariados es grandemente facilitado por el uso del álgebra de matrices. El álgebra de matrices nos proporciona un modelo estadístico conciso, además las relaciones formales expresadas en términos de matrices son facilmente programadas en computadora; además los métodos multivariados son basados en el modelo probabilístico conocido como la distribución normal multivariada.

Los objetivos de las investigaciones científicas para los cuales los métodos multivariados dan más naturalidad son:

- Reducción de datos o simplificación estructural.
- Ordenamiento y agrupacion .
- Investigación de la dependencia entre variables.
- Construcción de hipótesis y prueba.

La elección del método más apropiado depende del tipo de datos, el tipo de problema y el ordenamiento de objetivos los cuales son orientados para el análisis. Un tema en el análisis multivariado es la simplificación, que en otras palabras es sumar una gran cantidad de datos por medio de pocos parámetros relativamente. Muchas de las técnicas multivariadas son exploratorias ya que buscan la generación de hipótesis antes que hacer una prueba. Una distinción entre muchas técnicas es que algunas son de análisis, preocupadas primeramente por la relación entre variables, mientras que otras se preocupan por las relaciones entre individuos. Otro tema tratado en el análisis multivariado es la clasificación de un problema individual direccionado para cuestionarse si los individuos forman clusters (grupos), o si estos son más o menos dispersos al azar sobre el dominio de variación. Las técnicas para formar grupos de individuos son usualmente llamadas análisis de Clusters.

4.5 ANALISIS DE CLUSTERS.

El objetivo básico del análisis de clusters (grupos) es, encontrar los "grupos naturales" de algún conjunto de individuos (objetos, o puntos, o unidades, o cualquier cosa). Este conjunto de individuos puede formar una población completa o ser una muestra de alguna población grande. Más formalmente, el análisis de clusters es un intento para distribuir (colocar) un conjunto de individuos en un conjunto de exclusividad mutua completo, tales grupos contienen individuos similares entre si. Los grupos forman una partición, que puede ser subdividida en conjuntos pequeños o agrupaciones grandes; esto forma una estructura frecuentemente llamada un árbol jerárquico y es expresado en forma de diagrama.

Para aceptar un análisis de cluster necesitamos medir la semejanza (o no semejanza) de cada par de individuos. Las semejanzas algunas veces son observadas directamente, mientras que en otros casos son derivadas de la matriz de datos mediante alguna operación apropiada.

Aunque el análisis de cluster fué particularmente popular en los sesentas, en años muy recientes se ha presentado con una variedad de objetivos tales como:

(a) Exploración de datos.
(b) Reducción de datos.
(c) Generación de hipótesis, y
(d) Predicción basada en grupos.

El primer objetivo es explicado por si mismo; frecuentemente un análisis de cluster comienza con la idea clara de un conjunto de individuos estructurado, ahora si los cluster (grupos) son compactos entonces pueden reducir la información de n individuos a información de un número pequeño de grupos. Con el análisis se pueden generar algunas ideas (o hipótesis) respecto a la estructura de la población.

En particular la agrupación, el análisis y la agrupación de variables son tres tópicos que se consideran en el análisis de cluster; el término de agrupación es usualmente aplicado a los métodos de grupos, el análisis es usuado cuando una simple población es homogenea, así como la agrupación no natural y en la división de una población en subgrupos, el tercer tópico la agrupación de variables antes que individuos tiene como objetivo ver si se pueden encontrar subconjuntos de variables las cuales son altamente correlacionadas entre ellas o el cálculo de algún promedio de estos o bien para representar el subconjunto sin serias pérdidas de información.

4.6 APLICACIONES DE LAS CARITAS DE CHERNOFF.

El uso de caritas ha atraído considerablemente la atención de una gran cantidad de experimentos, que han sido aplicaciones que van desde los cráteres de la luna hasta la política Soviética en Sub-Saharam Africa. Chernoff público su primer artículo de representación gráfica de datos multivariados usando caritas en la Journal of the American Statistical Association en 1971. Siguiendo la publicación de éste artículo un número de personas trabajarón en varios campos usando éste programa. Algunos de estos trabajos son:

Lawrence A. Bruckner del Laboratorio Cientifico Los Alamos de la Universidad de California, trabajo bajo el apoyo de USERDA (United States Energy Reseach and Development Administration), estudiando diez grupos de aceites de la costa (Bruckner 1976).

En 1973 Chernoff aplicó su metodología para un experimento geológico, en el cual uso las caritas para representar el contenido mineral de cincuenta y tres muestras de una montaña en Colorado.

Otro trabajo fué el de Bud Good, en Bud Godde Sports Computer para presentar los resultados del juego de football.

Newton y Johnson desarrollaron junto con las caritas programas en linea que ejecutarán análisis estadístico con polígonos, gráficas de barras, flechas y diagramas. El trabajo fué usado en investigaciones del corazón, estudio de enfermedades respiratorias y comparaciones antropológicas.

Johns Hopkins inició un número de proyectos usando caritas de Chernoff en investigaciones de comunicaciones icónicas. Su primer trabajo fué en el Hospital de Servicios de Salud Pública de los Estados Unidos para desarrollar un método de proyecciones psiquiátricas.

El más reciente, el Dr. David L. Huff en la Universidad de Texas planeó usar las caritas para desarrollar los indicadores regionales urbanos que midan las condiciones de vida.

4.7 EJEMPLO DE CARITAS DE CHERNOFF.

Para dibujar una carita (Figura 4.1) se requieren de 20 números, los cuales se describen en la Figura 4.8, los cuales son rasgos que son controlados por los parámetros $x_1, x_2, \ldots, x_n$; por ejemplo x_6 controla la longuitud de la nariz y x_{20} el ancho de la nariz. En particular para este ejemplo las variables las dividimos en cinco grupos (Figura 4.9).

VARIABLES RANGO	RASGOS CONTROLADOS
X1	ANCHO CARA
X2	NIVEL OREJA
X3	ALTURA CARA
X4	CARA ELIPTICA GRANDE
X5	CARA ELIPTICA CHICA
X6	LONGITUD NARIZ
X7	BOCA CENTRADA (NIVEL)
X8	CURVATURA BOCA
X9	LONGITUD BOCA
X10	LONGITUD OJOS
X11	SEPARACION DE OJOS
X12	INCLINACION OJOS
X13	FORMA OJOS
X14	LONG. MEDIA DE OJOS
X15	POSICION DE PUPILAS
X16	ALTURA DE CEJAS
X17	INCLINACION DE CEJAS
X18	LONGITUD DE CEJAS
X19	DIAMETRO DE OREJAS
X20	ANCHO NARIZ

Figura 4.8 Parámetros de los 20 vectores estadísticos.

NOMBRE DE VARIABLES	NUM. RASGO	DESCRIPCION DE RASGO
CONTROL PERSONAL		
ESTUDIANTES ACADEMICOS	X2	NIVEL OREJAS
DEPARTAMENTOS DE TECNICAS	X17	INCLINACION CEJA
TECNICAS ECONOMICAS SOVIETICAS	X18	LONGITUD CEJA
CONT. DE PERSONAL MILITAR MIL FIJO		
MILES DE PERSONAS SOVIETICAS	X1	ANCHO CARA
MILES DE PERSONAS EN LA UNION SOVIETICA	X4	CARA ELIPTICA GRANDE
PUERTOS VISITADOS POR LA U. SOVIETICA	X5	CARA ELIPTICA CHICA
CONTACTOS ECONOMICOS		
CONVENIOS ECONOMICOS	X19	DIAMETRO OREJA
IMPORTACIONES	X11	SEPARACION DE OJOS
EXPORTACIONES	X12	INCLINACION OJOS
AYUDA ECONOMICA EXTENDIDA	X15	POSICION PUPILAS
AYUDA ECONOMICA ANUAL	X14	LONGITUD MEDIA DE OJOS
CONTACTOS DE AYUDA MILITAR		
CONVENIOS DE AYUDA MILITAR	X7	NIVEL BOCA
EXTENSIONES DE AYUDA MILITAR	X8	CURVATURA DE BOCA
CONCESIONES MILITARES	X9	LONGITUD BOCA
CARACTERISTICAS		
PRODUCTO NACIONAL BRUTO	X3	ALTURA CARA
TAMAÑO DE PRESUPUESTO	X6	LONGITUD NARIZ
TAMAÑO DE FUERZAS ARMADAS	X20	ANCHO NARIZ
RASGOS FIJOS		
FIJO	X10	NIVEL OJOS
FIJO	X13	FORMA OJOS
FIJO	X16	ALTURA CEJAS

Figura 4.9 Asociación de variables.

Se presentan algunas caritas representando la política exterior Soviética en algunos países de Sub-Saharam Africa, durante el periodo de 1964 a 1975. Como primer instancia se presentan tres grupos (grupos basados en rasgos faciales similares) los cuales representan hechos de tal politica. El grupo uno es contruido usando los valores máximos de los hechos politicos, el grupo dos representa el valor promedio de estos hechos y el grupo tres muestra los valores mínimos Figura 4.10.

Figura 4.10 Representacion de Grupos de la política exterior Soviética

La Figura 4.11 muestra un modelo Soviético presentando a Guinea, Somalia y República de Mali. En cada uno de los tres conjuntos de caras, la longitud de las cejas representa (número de técnicas económicas Soviéticas en el país), un cambio en el ángulo de las cejas (número de técnicas en el Departamento de entrenamiento en la URSS), un cambio en la posición de las pupilas representa (la ayuda económica extendida), otro indicador es la relación del ángulo de la ceja y lo amplio de la frente (técnicas y el departamento de personal militar para la Unión Soviética), la curvatura de la boca controla (la ayuda militar extendida).

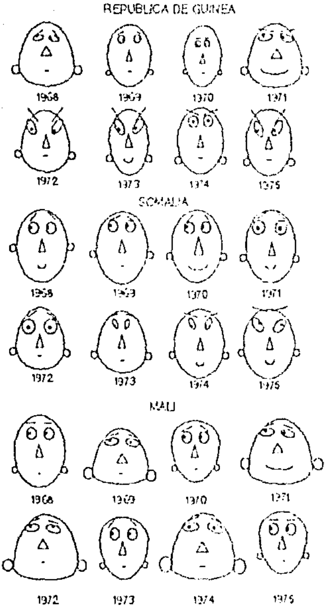


Figura 4.11 Caras de la Política exterior Sovietica de 1972 a 1975

En la Figura 4.12, las personas de República del Congo en 1975, presentan una variación en el diámetro de la oreja indicando (número de contratos económicos), lo alto y lo bajo de la boca (incremento en contratos de ayuda militar). También en éste grupo se puede observar que tanto en República del Congo y Nigeria el número de estudiantes en la Unión Soviética es elevado (altura de la oreja) y aun más tienen contratos de ayuda económica (diámetro de la oreja), así como también varios contratos de ayuda militar (nivel de la boca), personal militar Soviético (ancho de la cara), incremento de visitas al puerto (lo amplio de la barba) para 1969, 1970 y 1975.

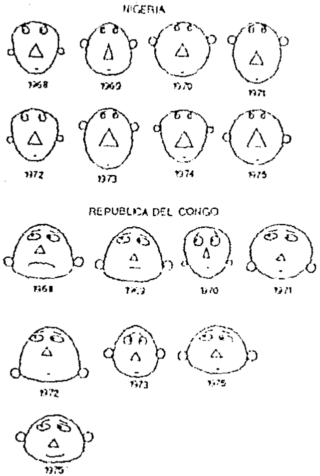


Figura 4.12 Cambio de caras de 1968 a 1975.

Las caritas de Chernoff son un modelo de icono, que como diagnóstico o herramienta de evaluación presenta un aceptable acercamiento para la representación de datos multivariados. Este modelo presenta ventaja en el análisis de datos multivariados, para percibir relaciones complejas o para detectar cambios y similitudes, indicadores que pueden formar modelos (patrones) o grupos (clusters), que no siempre son visibles con una simple correlación basada en la teoría lineal de dos dimensiones. Como puede notarse el icono carita nos representa de forma gráfica datos tomados de un archivo o base de datos, previamente construidos; es decir, nos permite analizar y sacar conclusiones de una forma menos complicada y porque no, hasta menos tediosa que cuando tenemos que leer varios documentos llenos de números (datos), que resultan menos atractivos para cualquier lector.

A lo largo de éste capitulo he hablado de la idea inicial de las caritas de Chernoff y sus gran aceptación como un método gráfico para representar datos multivariados por medio de rasgos faciales, lo cual da como resultado una agrupación de datos similares, que se pueden observar y sacar las relaciones existentes entre ellos.

Ahora, lo que se pretende en éste trabajo es tomar la carita como un método icónico para representar una consulta a una base de datos, esto es, tendremos un editor de caritas, que con sus rasgos faciales formaran una sintaxis (consulta); el tipo de sintaxis estará dada en términos de variables lingüísticas, las cuales se basan en la teoría de conjuntos difusos, que en lugar de números son palabras en lenguaje natural.

" VIVIR EN PROVECHO
PARA
MORIR SATISFECHO "

Porque:

" Se nace al fulgor del rayo, pero
cuando se fructifica en la vida,
el resplandor prevalece aún
después de la muerte "

Dr. Cristóbal del Río González.

CAPITULO 5

CONSULTA DIFUSA.

5.1 GENERALIDADES.

En la actualidad la mayor parte de los sistemas de base de datos usan lenguajes de consulta artificial . Ejemplos de tales lenguajes son ALPHA[19], SEQUEL[20], QUEL[21], Query-by-Example[22], los cuales son usados para accesar sistemas de base de datos relacional[23]. La característica básica de tales lenguajes es su sintaxis y semántica, que son precisamente formuladas y definidas. Estos sistemas de base de datos son el tipo de sistemas de símbolos inteligentes[24], donde las capacidades funcionales y las reglas fundamentales de decisión son influenciadas por los dos valores lógicos. Más específicamente, las bases de datos de tales sistemas contienen (o supuestamente contienen) el dato operacional preciso de una actividad (por ejemplo, los enteros en el rango de [1,100] son usados para representar la edad de una entidad, valores numéricos son usados para denotar el salario de un individuo, etc.). Secundariamente, las solicitudes sometidas para procesamiento son formuladas en base a atributos y valores precisos de la base de datos. Finalmente, los datos recuperados del sistema son asumidos para ser una respuesta exacta de la solicitud requerida por los usuarios.

Ciertamente existen dificultades pragmáticas y una cantidad enorme de inteligencia invocada en la construcción e implementación de tales sistemas precisos. Pero, hay que notar que su capacidad funcional, es limitada a situaciones o tareas caracterizadas por los dos valores lógicos. Tales capacidades son ciertamente sólo un subconjunto de los que los humanos poseen.

Una parte de nuestra inteligencia no es incorporada dentro de los sistemas de recuperación que existen actualmente. Esto es, se almacena información aproximada o imprecisa en la base de datos natural (por ejemplo Newton fué un gran inventor, el sol esta muy lejos de la tierra, la tierra es casi esferica, etc.). Podemos hacer preguntas aproximadas acerca de nuestros datos almacenados (¿Juan es un hombre elegante?, ¿Maria es una mujer hermosa?, ¿Estados Unidos de America es un país viejo?, etc.). Finalmente podemos hacer razonamientos imprecisos basados sobre nuestros factores o premisas aproximadas almacenadas.

Los términos tales como muy, grande, casi, viejo, varios, etc., son llamados términos difusos[25] (punto 3.3 capitulo 3). Tales términos constituyen una gran clase de datos usados por los humanos para almacenar o cumunicar información difusa del mundo real. Esto puede verse en el capitulo 2 con el uso de conjuntos difusos[26], que son la clase en donde su transición no es tan fuerte de pertenencia a no pertnencia. Por ejemplo, la clase de mujeres hermosas, hombres viejos, negocios grandes, datos reelevantes, algoritmos eficientes, etc. Incidentalmente, los términos tales como viejo, joven, muy viejo, más o menos joven , no muy joven, etc., pueden ser observados como valores lingüisticos del atributo edad comparada a los números 20,30, etc., los cuales son valores numéricos de este atributo. Una caracteristica básica de un valor lingüístico tal como viejo es que es menos preciso y menos informativo que un valor numérico tal como 20. Esto también es basado en lo que es llamado el principio de incompatibilidad[27], " un principio en el cual sostiene que la alta precisión es incompatible con la alta complejidad".

Los aspectos formales para ésto son basados en la teoría de conjuntos difusos y variables lingüísticas las cuales se discutieron en el capítulo 2 y 3. La teoría de conjuntos difusos representa un intento para construir un cuadro conceptual para un tratamiento sistemático de difusiones tanto cuantitativo como cualitativo.

Los iconos basados en **caritas de Chernoff** en combinación con la teoría de conjuntos difusos y variable lingüística son una nueva modalidad en el intento del manejo de consultas difusas en una base de datos. En el presente capitulo trataremos consultas difusas con un lenguaje conocido como es SEQUEL y con los iconos caritas de Chernoff, y al mismo tiempo hacer una comparación entre ambas formas. Entendiendo entonces que una **Consulta Difusa** es aquella que usa imprecisiones o predicados difusos (por ejemplo Edad="muy joven", Salario="más o menos alto", Año de contrato="reciente", Salario > 20,000, etc.). Ahora bien, como una base para el procesamiento consulta difusa es introducido un sistema de recuperación difusa, que también se basa en teoría de conjuntos difusos y variables lingüisticas.

Un modelo de este tipo de sistema tiene como primer paso el procesamiento de consultas difusas que consiste del significado asignado a los **Términos Difusos** (valores lingüisticos), de un conjunto de términos, usados para la formulación de la consulta. El significado de un término difuso es definido como un conjunto difuso en un Universo de discusión, el cual contiene los valores numéricos de un dominio de una relación en el sistema de base de datos.

El sistema de recuperación difusa se desarrolla con técnicas de un modelo de alto nivel las cuales pueden ser usadas en un sistema de base de datos y con la posibilidad de implementar tales técnicas es un medio ambiente de estudio real.

5.2 MODELO DE DATOS.

Por lo general las técnicas para el procesamiento de consultas difusas son aplicables más a sistemas de base de datos. Sin embargo, al tratar el problema en sistemas relacionales se asume que el modelo de datos del sistema consiste de una colección de n relaciones normalizadas.

Ejemplo 5.2.1. La Figura 5.1, representa algunas entradas de la relación

EMP(MNO, NOMBRE, EDAD, DNO, NOJEFE, SALARIO, AÑOCONT)

donde los renglones de la tabla corresponden a las tuplas de la relación EMP. Cada tupla especifica los datos de un empleado en la compañía. Los elementos de una tupla EMP respectivamente consiste del número del empleado, nombre, edad, número de departamento, número de su jefe, salario y año de contratación. Los nombres MNO, NOMBRE, EDAD, etc., son el dominio de la relación EMP. La Figura 5.2, representa algunas entradas de la relación

DEPT(DNO, DNOMBRE)

donde una tupla de la relación DEPT contiene el número de departamento y el nombre del departamento. La relación DEPT es definida solamente en los dominios de **DNO** y **DNOMBRE**.

Comentario 5.2.1.- En general un modelo de datos no puede ser limitado para describir la información precisa de las entradas en una empresa. La Figura 5.3, ilustra otra posibilidad para la relación EMP en la cual algunos valores lingüísticos (términos difusos) tienen que ser usados para construir los datos de algunas entradas en la relación. Esto es, por ejemplo, los términos Joven y Muy alto representan respectivamente los valores de los dominios **EDAD** y **SALARIO** de la tupla x_1. Sin embargo, en este trabajo restringimos el modelo de datos para situaciones precisas. En otras palabras, se usará términos difusos para formular consultas contra un modelo de datos, el cual contiene información precisa de las entradas.

	ENO	NOMBRE	EDAD	DNO	JEFE	SALARIO	AÑOCONT
t1	101	ZECUA	25	15	104	40,000	1972
t2	102	REYES	30	15	105	15,000	1974
t3	103	CHAPA	30	15	106	20,000	1974
t4	104	LEON	55	15	NINGUNO	20,000	1960
t5	105	RUIZ	25	19	106	25,000	1975

Figura 5.1 Relación EMP.

DNO	DNOMBRE
15	ADMINISTRACION
19	CONTABILIDAD
25	INGENIERIA

Figura 5.2 Relación DEPT.

	MNO	NOMBRE	EDAD	DNO	JEFE	SALARIO	AÑOCONT
x1	101	ZECUA	JOVEN	15	104	MUY ALTO	1972
x2	102	REYES	30	15	105	15,000	1974
x3	103	CHAPA	30	15	106	20,000	1974
x4	104	LEON	VIEJO	15	NONE	20,000	1960
x5	105	RUIZ	5	13	106	ALTO	MUY RECIEN

Figura 5.3 Relación EMP con términos difusos en algunos de los dominios.

5.3 CONSULTA DIFUSA EN LENGUAJE SEQUEL.

Tomando las relaciones EMP y DEPT y usando la sintaxis del lenguaje SEQUEL. Una consulta no difusa en este lenguaje es como sigue:

Q1- Lista los nombres de empleados que tengan 20 años y que ganen más de 20,000.

```
SELECT NOMBRE
FROM EMP
WHERE EDAD=20 AND SALARIO > 20,000;
```

la consulta está dada por un bloque, el cual consiste de las cláusulas SELECT, FROM Y WHERE. Los parámetros de SELECT regresan los valores del dominio (NOMBRE) de la relación (EMP) especificada en la cláusula FROM. Estos valores son los campos de las tuplas las cuales pueden satisfacer la condicion calificada (EDAD=20 y SALARIO > 20,000) de la cláusula WHERE. También, en SEQUEL una consulta puede consistir de uno o más bloques (anidados) y varios parámetros pueden ser usados en cada cláusula. Una consulta difusa es áquella en la que la cláusula WHERE contiene uno o más términos difusos.

Q2- Lista los nombres de empleados que son jovenes y ganan más de 15,000.

```
SELECT NOMBRE
FROM EMP
WHERE SALARIO > 15,000 AND EDAD="JOVEN";
```

La cláusula WHERE de ésta consulta consiste del predicado preciso SALARIO > 15,000 y el predicado difuso EDAD="JOVEN". Se puede notar que el valor lingüistico JOVEN juega el mismo papel que el valor numérico, puede ser 20 en la formulación del predicado asociado. La diferencia principal entre los dos predicados EDAD="JOVEN" y EDAD=20 es que la primera es menos informativa y menos precisa que la segunda. Escencialmente ésta es característica de los predicados difusos los que hacen toda la consulta imprecisa o difusa. Incidentalmente podemos considerar que la imprecisión de una consulta difusa relaciona a las difusiones de la consulta semántica. En otras palabras la respuesta recuperada de una consulta difusa en una base de datos consiste de un conjunto difuso de datos.

Q3- Lista los nombres de empleados que ganen muy alto y no muy viejo.

```
SELECT NOMBRE
FROM EMP
WHERE SALARIO="MUY ALTO" AND EDAD="NO MUY VIEJO";
```

Q4- Lista los empleados que hayan sido recientemente contratados y ganen 25,000.

```
SELECT NOMBRE
FROM EMP
WHERE AÑOCONT="RECIENTE" AND SALARIO=25,000;
```

Q5- Localizar el salario promedio de empleados que son más o menos viejos, ganen más de 15,000 y que esten en el departamento 15.

```
SELECT AVE(SALARIO)
FROM EMP
WHERE EDAD="MAS O MENOS VIEJO" AND SALARIO > 15,000 AND
        DNO=15;
```

Los ejemplos mencionados muestran la aplicación de valores lingüisticos (términos difusos) para la construcción de las llamadas consultas difusas.

5.4 PROCESAMIENTO DE UNA CONSULTA DIFUSA.

Las técnicas para el procesamiento de consulta difusa son basadas en algunos resultados teóricos. Convencionalmente un componente importante de cualquier técnica de procesamiento de una consulta difusa consiste de una operación de igualdad. Bajo ésta operación los elementos de información de una consulta calificada (o simplemente una consulta) son comparadas con los elementos de un registro de la base de datos. Esta comparación puede determinar el registro que puede o no satisfacer la consulta correspondiente. Desafortunadamente éste esquema no puede ser usado directamente en orden para procesar una consulta difusa. La dificultad es que los elementos de información de un registro de base de datos son precisos mientras que una consulta difusa es precisamente imprecisa y difusa.

La aproximación para tratar éste problema en el procesamiento de consulta difusa es basado en el esquema de recuperación asociativa[28,29]. Este sistema de recuperación es apliamente usado en el área de sistemas de recuperacion documental; en éste esquema un término en la consulta es reemplazado por los términos asociados los cuales son definidos sintácticamente o semánticamente. Siguiendo este esquema y considerando la definición de un término difuso (definido en el capitulo 2) como los términos asociados semánticamente a un término difuso que son direccionados a una aproximación. Esto indica que primero los términos difusos en una consulta pueden ser reemplazados por sus significados asociados; después la operación de igualdad puede ser aplicada para comparar los datos de los conjuntos difusos (los significados de términos difusos) con los correspondientes elementos de los registros de la base de datos.

Finalmente se busca una aproximación conveniente y sistemática para introducir el procesamiento de consulta difusa en conjunción con las actividades de la recuperación de un sistema de recuperación difusa.

Sistema de Recuperación Difusa.

Se considera un sistema de recuperación difusa (o simplemente un sistema de recuperación) por una cuadrupla (X,T,Q,τ) donde X es una colección de registros de datos (la siguiente X puede ser asumida para ser una n-aria relación normalizada), T es el conjunto de términos; Q es un conjunto consulta; $\tau : Q \quad X \quad X \dashrightarrow [0,1]$ (llamada la función de igualdad) asignada para cada par (q,x) donde $q \in Q$ y $x \in X$, un número $\tau(q,x)$ en el intervalo $[0,1]$ llamado el índice de igualdad para la consulta q y el registro x.

Registro de datos X.

Los registros X consisten de las tuplas de una n-aria relación normalizada. El término X puede ser usado para denotar el nombre de la relación así como la colección de tuplas $(x_1, x_2, \ldots, x_m)$. Las etiquetas $A_1, A_2, \ldots, A_n$ denota los nombres de dominios de la relación X. Además, la notación $x_i[A_j]$ indican el valor del dominio A_j en la tupla x_i de la relación X. Esto es, tomando como referencia a la Figura 5.1, X="EMP", A_1="ENO", A_2="NOMBRE", A_3="EDAD". Entonces, se puede notar que $x_1[A1]=101$, $x_2[A6]=15{,}000$, etc.

Conjunto de Términos T.

Informalmente el conjunto de términos T es la colección de los valores lingüísticos permitidos (términos difusos) de las variables lingüísticas asociadas con algunos de los dominios de la relación X. Formalmente T es definido como

$$T = \Sigma\ T_i$$

donde Σ denota la unión y T_i es el conjunto de términos de la variable lingüística representada por el nombre del dominio A_i (ver definición 3.1.1. de variable lingüística capítulo 3).

Ejemplo 5.4.1. Basado en las discusiones del Capítulo 2, entonces el conjunto de términos de la relación EMP ilustrada en la Figura 5.1, puede ser considerada como:

T_3 = (viejo, muy viejo, no viejo, más o menos viejo, completamente joven,...)

T_6 = (alto, bajo, muy alto, más o menos alto,...)

T_7 = (reciente, más o menos reciente, muy reciente,...)

donde éstos corresponden a los atributos EDAD, SALARIO y AÑOCONT respectivamente. De esta forma el conjunto de términos es $T = T_3 + T_6 + T_7$.

Conjunto Consulta Q

Una consulta q del conjunto de consulta Q es definido recursivamente como:

$$q = p + \sim q + (q \text{ o } q) + (q \text{ y } q)$$

donde + denota "o", ~ denota la negación, y p es definido como:

$$p = A_i = 't'$$

donde $t \in T_i$ y A_i es un nombre del dominio de la relación X.

Ejemplo 5.4.2. Basado en el Ejemplo 5.4.1., una consulta q puede ser:

EDAD = "VIEJO" y (AÑOCONT = "MAS O MENOS RECIENTE" o SALARIO = "NO MUY ALTO")

Función de Igualdad τ

La función de igualdad τ es formulada en base a los siguientes pasos.

P1- Suponer una consulta q, que es el predicado difuso A_i="t", donde $t \in T_i$. Asimismo, asumimos que el significado del término t es un subconjunto difuso del universo de discusión U_i, definido como

$$M(t) = \Sigma\, \mu_i(u)/u$$

donde $u \in U_i$. Se puede notar que U_i es el universo de discusión de la variable lingüística asociada con el dominio A_i. Los indices de la igualdad $\tau(q,x)$ donde $x \in X$, son definidos por

$$\tau(q,x) = u_i(u)$$

en donde $u=x[A_i]$, es el valor del dominio A_i en la tupla x.

P2- Si p y q son consultas en Q, y x es una tupla de X, el conjunto

$$\tau(p \text{ y } q,\ x) = \min(\tau(p,x)\ ,\ \tau(q,x))$$

$$\tau(p \text{ o } q,\ x) = \max(\tau(p,x)\ ,\ \tau(q,x))$$

y

$$\tau(\neg q\ ,\ x) = 1 - \tau(q,x)$$

P3- Si q es cualquier consulta en Q, entonces $\tau(q,x)$ es obtenida por la aplicación de cálculos definidos en los pasos P1-P2.

Ejemplo 5.4.3. Suponemos las consultas q_1, q_2, q_3 y q_4 son respectivamente los términos EDAD="VIEJO", EDAD="JOVEN", SALARIO="ALTO" y AÑOCONT="RECIENTE". Los significados de VIEJO, ALTO, RECIENTE y ALTO son subjetivamente definidos como:

$$\mu_{joven}(u) = \begin{cases} 1 & \text{Para } u < 25 \\ \left[1 + \left[\dfrac{u-25}{5} \right]^{2} \right]^{-1} & \text{Para } u \geq 25 \end{cases}$$

$$\mu_{viejo}(u) = \begin{cases} 0 & \text{Para } u \leq 50 \\ \left[1 + \left[\dfrac{u-50}{5} \right]^{-2} \right]^{-1} & \text{Para } u \geq 50 \end{cases}$$

$$\mu_{alto}(u) = \begin{cases} \left[1 + \left[\dfrac{u-15000}{5000} \right]^{-2} \right]^{-1} & \text{Para } u \geq 15{,}000 \\ 0 & \text{otro caso} \end{cases}$$

$$M(\text{RECIENTE}) = 1.0/1976 + 0.8/1975 + 0.6/1974 + 0.4/1973$$

Usando las reglas de igualdad mencionadas, la información contenida de la relación EMP en la Figura

5.1, y los significados de los términos difusos, los indices de igualdad son tabulados en la Figura 5.4.

CONSULTAS / REGISTROS	q1 EDAD VIEJO	q2 EDAD JOVEN	q3 SALARIO ALTO	q4 AÑOCONT RECIENTE	q5 q2 AND [illegible]
x1	0.0	1.0	1.0	0.0	1.0
x2	0.0	0.5	0.0	0.6	0.5
x3	0.0	0.5	0.5	0.6	0.5
x4	0.5	0.0	0.5	0.0	0.0
x5	0.0	1.0	0.8	0.8	0.8

Figura 5.4 Los indices de igualdad $\tau(qi,xi)$ para las consultas y las tuplas de la relación EMP.

Hay que notar que en la Figura 5.4, la consulta q_5, es EDAD="JOVEN" y (SALARIO = "ALTO" o AÑOCONT = "RECIENTE")

Comentario 5.4.2. El indice de igualdad $\tau(q,x)$ puede ser considerado como el grado de satisfacción de la consulta q por el registro x. Esto es, comparado para los casos convencionales donde $\tau(q,x)$ es 1 o 0 dependiendo sea o no satisfecha la consulta q por x respectivamente.

Sistema de Respuesta.

Dada una consulta q, el sistema de respuesta f(q) es un subconjunto difuso de X donde la función de pertenencia es definada por

$$\mu_{f(q)}(X) = \tau(q,x) \qquad \text{para } x \in X$$

en otras palabras,

$$f(q) = \Sigma\ \tau(q,x)/x$$

Ejemplo 5.4.4. Tenemos las consultas q_2, q_4 y q_5, definidas en el Ejemplo 5.4.3. Usando los indices de igualdad de la Figura 5.4, entonces, el sistema de respuestas consiste respectivamente de

$f(q_2)$=f(EDAD="JOVEN") = $1.0/x_1 + 0.5/x_2 + 0.5/x_3 + 1.0/x_5$

$f(q_4)$=f(AÑOCONT="RECIENTE") = $0.6/x_2 + 0.6/x_3 + 0.8/x_5$

$f(q_5)$=f(EDAD = "JOVEN" y (SALARIO = "ALTO" o AÑOCONT = "RECIENTE"))
= $1.0/x_1 + 0.5/x_2 + 0.5/x_3 + 0.8/x_5$

Comentario 5.4.3. $\mu f(q)$ el grado de pertenencia del registro x en la respuesta recuperada **f(q)**, puede ser vista como el grado de pertenencia del contenido de información de la tupla x para la consulta q. Si $\mu_{f(q)}(x_i) > \mu_{f(q)}(x_j)$ entonces se puede decir que el registro x_i es más pertinente que x_j para la consulta q. En otras palabras la consulta q es satisfecha más con x_i que con x_j. Tal que, los valores pertinentes pueden ser usados para definir un orden en los registros en **f(q)**.

Estrategias de Recuperación.

En la práctica los registros **X** son parte de los datos almacenados en una base de datos. Los resultados establecen alguna relación entre la respuesta recuperada **f(q)** del sistema (X, T, Q, τ) y la lista invertida de la relación X.

Se asume que la estructura almacenada de los registro X (en la base de datos) es una organización de archivo invertida. Tal que L(u) denota una lista invertida para el valor **u** en un dominio de la relación X. Esto es, por ejemplo, si u es un valor del dominio EDAD, entonces L(u) es un subconjunto de registros X, donde el valor de EDAD para cada miembro de L(u) es u. Se puede notar que la lista L(u) contiene las direcciones almacenadas o identificadores de los registros incluidos. Para simplificar, la etiqueta x_i puede ser usada para denotar un registro (tupla), así como las direcciones almacenadas.

Proposición 5.4.1. Tenemos que $t \in T_i$, es un término difuso donde su significado es definidos por

$$M(t) = \Sigma\ \mu_t(u)/u$$

donde μ_t es la función de pertenencia del significado M(t). Si tenemos que la consulta **q** es el predicado difuso A_i="t", entonces, la respuesta recuperada f(q) del sistema (X,T,Q,τ) puede consistir de

$$f(q) = \sum_{\mu_t(u)} \mu_t(u) \cdot L(u)$$

donde Σ denota la unión de conjuntos difusos y $\mu_t(u) \cdot L(u)$ denotan un conjunto difuso en X, con dos funciones características definidas por

$$\mu_{\mu t(u) \cdot L(u)}(X) = \begin{cases} \mu_t(u) & \text{para } x \in L(u) \\ 0 & \text{en otro caso} \end{cases}$$

Ejemplo 5.4.5. Considerando la relación EMP de la Figura 5.1 Las listas invertidas correspondientes al dominio EDAD consiste de

$$L(EDAD = 25) = (x_1 \; , \; x_5)$$
$$L(EDAD = 30) = (x_2 \; , \; x_3)$$
$$L(EDAD = 55) = (x_4)$$

Ahora t es el término difuso JOVEN, donde su significado está dado en el Ejemplo 5.4.3. De acuerdo con la proposición 5.4.1., la respuesta recuperada f(q) para la consulta **q** definida por EDAD="JOVEN", consiste de

$$f(q) = \mu_t(25) \cdot L(EDAD=25) + \mu_t(30) \cdot L(EDAD=30) + \mu_t(55) \cdot L(EDAD=55)$$

$$= 1.0(x_1, x_5) + 0.5(x_2, x_3) + 0.0(x_4)$$

$$= 1.0/x_1 + 0.5/x_2 + 0.5/x_3 + 1.0/x_5$$

Otro resultado que es de gran ayuda para las estrategias de recuperación está dado mediante la siguiente proposición.

Proposición 5.4.2. Para cualquier consulta p y q en Q del sistema (X, T, Q, τ)

$$f(p \text{ y } q) = f(p) \cap f(q)$$

y

$$f(p \text{ o } q) = f(p) \cup f(q)$$

donde $\cap$ y $\cup$ son la intersección y unión de conjuntos difusos.

5.5 TECNICAS CONCEPTUALES PARA EL PROCESAMIENTO DE UNA CONSULTA DIFUSA.

El sistema de recuperación difusa (X, T, Q, τ) es un modelo de alto nivel para las técnicas que pueden ser usadas en un sistema de base de datos para el procesamiento de consultas difusas. Podemos considerar el procesamiento de las consultas con la siguiente estructura

```
SELECT Nombre-dominio 1, Nombre-dominio 2,...
FROM X
WHERE q;
```

donde los parámetros de SELECT son algunos nombres de dominios de la relación X, y q es una consulta en el sistema (X, T, Q, τ).

Ejemplo 5.5.6. Una consulta con la sintaxis anterior puede ser considerada como (ver Figura 5.1)

```
SELECT NOMBRE,DNO
FROM EMP
WHERE EDAD="JOVEN" AND (SALARIO="ALTO" OR AÑOCONT =
          "RECIENTE");
```

De acuerdo al procesamiento de una consulta difusa tenemos los siguientes puntos.

Punto 1. Estructura de Almacenamiento.- Asumimos que la estructura de almacenamiento de una relación X en el sistema de base de datos es una organización de archivo invertida. Del mismo modo un directorio es mantenido por cada relación X. El directorio es hecho de listas direccionadas invertidas, asociadas con los valores en el dominio de la relación X. Sin embargo, de acuerdo con la Proposición 5.4.1. el directorio contiene información adicional relacionada a los valores lingüísticos (términos difusos). Más especificamente, se puede asumir que los términos del conjunto de términos T suma las compatibilidades (significados) de tales términos con los valores numéricos en los dominios de la relación X, que son como parte también del directorio.

Por ejemplo la Figura 5.5, representa un ejemplo de tales directorios; los términos difusos tales como VIEJO, JOVEN, RECIENTE y ALTO son parte del directorio de datos. La compatibilidad de un término difuso con un valor numérico en un dominio es denotado por un número asignado para la ramificación, que conecta el término anterior con el posterior. Esto es, por ejemplo los valores 0.5 y 1.0 respectivamente representan las compatibilidades del término

JOVEN con los valores numéricos correspondientes a 30 y 25 de EDAD (ocurridos en la relación EMP). Además, un término x_i es una lista invertida usada para indicar las direcciones almacenadas de la correspondiente tupla x_i en la base de datos.

Punto 2. Diccionario.- El conjunto de términos T del sistema (X, T, Q, r) suman los correspondientes significados de estos términos residentes en la base de datos. Esto es, T juega el papel de diccionario. El significado de cualquier término es almacenado en una forma tabulada o rutinas semánticas, las cuales automaticamente generan el significado del término. Tales rutinas pueden especificamente ser aplicadas para asignar significado a términos compuestos tales como muy joven, más o menos reciente, no muy rico, etc. Estas rutinas son construidas en base a las técnicas desarrolladas por L. A. Zadeh[27,30] (mencionadas en el capítulo 3 lógica difusa). Incidentalmente el administrador de base de datos[23] puede ser responsable para mantener el significado de términos difusos en el sistema.

Puede notarse que el significado de un término difuso en un directorio, como lo menciona el punto 1, es un subconjunto definido por el diccionario T. Más especificamente, el directorio contiene las compatibilidades de un término difuso con solamente aquellos valores (de un dominio) los cuales existen inmediatamente en la base de datos. Es decir, por ejemplo, el significado del término joven como es definido en el Ejemplo 5.4.3. relaciona a muchos valores numéricos, el mismo término en el directorio mostrado en la Figura 5.5, es asociado con los valores 25 y 30 (existentes en la base de datos). El significado de un término en el diccionario T puede ser usado para construir las compatibilidades correspondientes aplicables a un directorio.

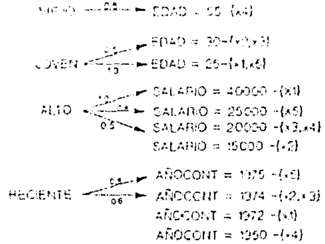


Figura 5.5 Una parte del directorio del archivo almacenado EMP.

Punto 3. Interprete.- Tomando en cuenta el modelo del interprete SEQUEL[31], con algunas simplificaciones se puede asumir que el sistema contiene un interprete que consiste de tres grandes partes; un **Parser** (analizador), un **Evaluador Difuso** (o simplemente un evaluador) y un **Scanner** (Explorador). La Figura 5.6, muestra la relación entre estos componentes. El Evaluador puede recuperar un subconjunto difuso de una tupla de identificadores (la tupla de direcciones almacenadas) de la base de datos. Esto es denotado por **f(q)** en la Figura 5.6. Las estrategias de recuperación del Evaluador son basadas en las Proposiciones 5.4.1. y 5.4.2. El Evaluador de respuesta, **f(q)** puede ser considerado como los identificadores de las tuplas, los cuales satisfacen algunos grados de la consulta califica q, junto con el grado de satisfacción de cada tupla.

La calificación q y el nombre de la relación X de una consulta son suministradas al Evaluador por medio del Parser. Similarmente, el Parser puede someter los parametros del SELECT a el Scanner.

El Scanner puede recuperar los elementos de datos asociados (expresados por los parametros del SELECT) de las tuplas identificadas por f(q) de la base de datos. La salida del Scanner es denotada por R en la Figura 5.6, es un subconjunto difuso de una proyección[32] en la relación X. Desde luego los dominios de ésta proyección son especificados en el SELECT. Ya que R es un conjunto difuso, el grado de pertenencia en R puede ser usado para definir un ordenamiento en los elementos de R.

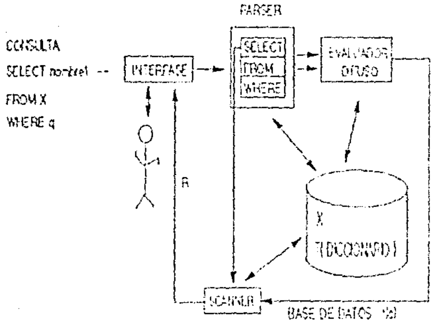


Figura 5.6 Componentes funcionales de un sistema base de datos para el procesamiento de una consulta difusa.

Ejemplo 5.5.2. Considerando la consulta del ejemplo 5.5.1. Siguiendo la proposición 5.4.2., la salida del Evaluador consiste de

$$f(q)=f(\text{EDAD} = \text{"JOVEN"}) \cap (f(\text{SALARIO} = \text{"ALTO"}) \cup f(\text{AÑOCONT} = \text{"RECIENTE"}))$$

los significados de los términos difusos en estos predicados son definidos en el ejemplo 5.4.3. De acuerdo con el punto 1, la proposición 5.4.1., y la Figura 5.1, cada subcomponente de respuesta **f(q)** puede ser obtenida usando las listas invertidas de la relación EMP mostrada en la Figura 5.5. Esto es, el Evaluador recupera las listas invertidas correspondientes, pudiendo ejecutar las operaciones siguientes:

$$\begin{aligned} f(\text{EDAD} = \text{"JOVEN"}) &= 0.5L(\text{EDAD} = 30) + 1.0L(\text{EDAD} = 25) \\ &= 0.5(x_2, x_3) + 1.0(x_1, x_5) \\ &= 1.0/x_1 + 0.5/x_2 + 0.5/x_3 + 1.0/x_5 \end{aligned}$$

$$\begin{aligned} f(\text{SALARIO}=\text{"ALTO"}) &= 1.0L(\text{SALARIO}=40{,}000) + \\ &\quad 0.8L(\text{SALARIO}=25{,}000) + \\ &\quad 0.5L(\text{SALARIO}=20{,}000) \\ &= 1.0/x_1 + 0.5/x_3 + 0.5/x_4 + 0.8/x_5 \end{aligned}$$

$$f(\text{AÑOCONT} = \text{"RECIENTE"}) = 0.8L(\text{AÑOCONT} = 1975) + 0.6L(\text{AÑOCONT} = 1974)$$

$$= 0.6/x_2 + 0.6/x_3 + 0.8/x_5$$

Usando la salida del Evaluador mencionada obtenemos:

$$f(q) = 1.0/x_1 + 0.5/x_2 + 0.5/x_3 + 0.8/x_5$$

los x_i en $f(q)$ representa los idenficadores de tuplas de la relación EMP. Usando estos identificadores y los nombres de los campos NOMBRE y DNO, el Scanner puede sujetarse a los datos mostrados en la Figura 5.7, para la interface. Cada valor de la columna del índice de igualdad en esta figura representa el grado de satisfacción de la consulta del renglón correspondiente de la tabla. La interface del sistema puede entonces presentar las entradas ordenadas (basado en los índices de igualdad) de la tabla Figura 5.7, como la respuesta recuperada del sistema.

NOMBRE	DNO	COMB. IND
ZECUA	15	1.0
RUIZ	19	0.8
REYES	15	0.5
CHAPA	15	0.5

Figura 5.7 La salida del Scanner del Ejemplo 5.5.2

5.6 CARITAS DE CHERNOFF COMO CONSULTAS DIFUSAS.

Se menciono en el capitulo 4 que la idea original de las caritas propuestas por el Profesor Chernoff es representar de una forma gráfica datos multivariados de tal manera que cada carita encapsule un grupo de datos similares. Esto es, dada una base de datos y mediante alguna técnica de clasificación o análisis de cluster(grupos) obtener los datos o grupos de datos con ciertas similitudes. Una vez obtenida ésta información representarla mediante el icono de carita, proporcionando ilustración y comunicación de esta información, tomando en cuenta que cuando una gran cantidad de información es presentada es importante presentarla mediante algo que sea recordado y las caritas mediante rasgos faciales son idóneas para recordarse, ya que como un medio para agrupar datos de forma compacta nos permite un análisis y así mismo entender las relaciones entre los datos.

Ahora bien, las caritas de Chernoff en éste trabajo se proponen como un modelo de icono para el procesamiento de consultas difusas, el cual nos ofrece una semántica muy rica. Dichas caritas asocian a operadores que permiten la clasificación de datos y facilite la elaboración de consultas difusas a una base de datos. Es decir, el icono de carita se presenta como una sintaxis de la consulta difusa (basada en conjunto difusos capítulo 2) y con la cual obtener la información de la base de datos que satisfaga tal consulta.

La propuesta de este modelo icónico como sintaxis de una consulta de tipo difuso se fundamenta de la siguiente forma:

Conjunto Difusos ····> Lógica Difusa

↓

Colsulta Difusa

↓

Caritas de Chernoff

Tomando como base la teoría de conjuntos difusos (capítulo 2), que nos conduce a la lógica difusa (capítulo 3), en donde las variables lingüísticas toman valores que son palabras y no números, que comunmente son usadas en el discurso humano, tales valores o términos como muy, grande, joven, viejo, varios, más o menos, etc., son llamados términos difusos los cuales constituyen una gran clase de datos usados por el hombre para almacenar o comunicar información difusa al mundo real.

Considerando esto y llevandolo al manejo de información en un modelo de base de datos, el cual contendrá información precisa nos conduce a lo que es llamado consulta difusa, en donde se usan términos no precisos para formar una sintaxis. La sintaxis se puede presentar mediante un lenguaje de tipo textual (por ejemplo SEQUEL punto 5.3), pero ahora se intenta manejar de forma gráfica por medio de caritas de Chernoff, que con cada rasgo facial representará una variable lingüística para que finalmente encapsule una consulta bajo términos difusos.

La sintaxis representada por un icono de carita estará manejada por tres partes que son básicas para cualquier interprete; un Parser (analizador), un Evaluador difuso (evaluador) y un Scanner (explorador). En donde por medio de programación automática se obtendrán datos ordenados que cumplan con la consulta.

Este procedimiento se llevará a cabo bajo el esquema presentado en la Figura 5.8, que muestra los componentes funcionales de un sistema de base de datos para el procesamiento de una consulta difusa, teniendo como entrada un un icono de carita.

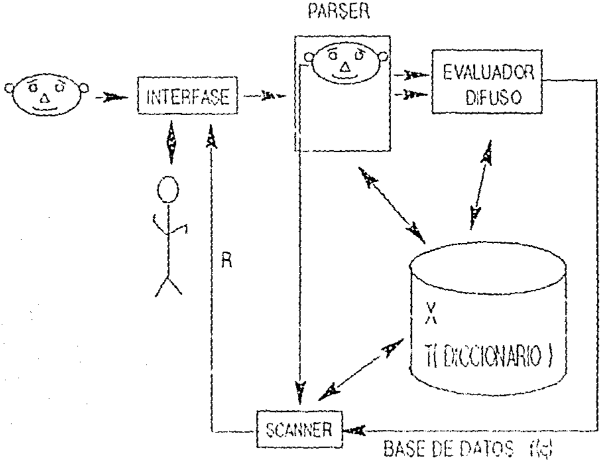


Figura 5.8 Componentes funcionales de un sistema base de datos para el procesamiento de una consulta difusa.

En la sección 5.3 de este capitulo se presento la sintaxis de consultas difusas con el lenguaje SEQUEL, en este lenguaje la sintaxis está dada por medio de un bloque que consiste de las cláusulas SELECT, FROM y WHERE. Ahora, la idea es que mediante seis rasgos faciales básicos cara, boca, nariz, cejas, orejas y ojos se elabore la consulta difusa; cada uno de los rasgos puede representar una variable la cual va presentar variaciones de acuerdo a la forma de los rasgos, a continuación se presenta las distintas formas que pueden tomar los rasgos:

RASGOS	FORMA
Cara	Ancha Redonda Ovalada
Boca	Sonriente seria triste
Nariz	Ancha Larga Recta
Ceja	Inclinada hacia arriba forma horizontal Inclinada hacia abajo
Oreja	A nivel de Ojos A nivel de Nariz A nivel de Boca Redonda Ovalada
Ojos	Redondos Ovalados forma de línea

Las distintas formas que puede presentar cada rasgo van a representar los valores altos, medios o bajos que puede tomar cada variable asignada.

La consulta difusa que se basa en predicados o enunciados y no números, contiene variables que toman valores lingüísticos o términos difusos (capitulo 2), por ejemplo si una variable llamada sueldo asume la siguiente forma:

sueldo = x

donde x es un valor del conjunto de términos difusos sueldo, dado por

T(sueldo)= alto + muy alto + altamente alto + bajo + muy bajo + altamente bajo + más o menos alto +...

y si, ahora decimos

$$x = \text{alto}$$

la asignación resultante es

$$\text{sueldo} = \text{alto}$$

alto es un término difuso del conjunto de términos sueldo, el cual tendrá un significado dado por

$$M\ (\text{alto}) = \begin{cases} \left[1 + \left[\dfrac{u-15000}{5000} \right]^{-2} \right]^{-1} & \text{Para } u \geq 15{,}000 \\ 0 & \text{otro caso} \end{cases}$$

esta ecuación de compatibilidad es obtenida mediante análisis de semántica como fué comentado en el capítulo 3. La variable sueldo puede asumir cualquier valor del conjunto de términos, esto quiere decir que puede ser

$$\text{sueldo} = \text{muy alto}$$
$$\text{sueldo} = \text{altamente alto}$$

cuyos significados también estarán dados por una función de compatibilidad en base a etiquetas lingüísticas (ver sección 3.1.2), entonces muy alto será definida por la operación de concentración (CON) y la función de compatibilidad resultante es

$$M(\text{muy alto}) = CON(M(\text{alto}))$$

y la forma que asume la función de compatibilidad es

$$M(\text{muy alto}) = \begin{cases} \left[1 + \left[\dfrac{u-15000}{5000} \right]^{-2} \right]^{-2} & \text{Para } u \geq 15{,}000 \\ 0 & \text{otro caso} \end{cases}$$

ahora para el término altamente alto usamos la etiqueta mas definida como

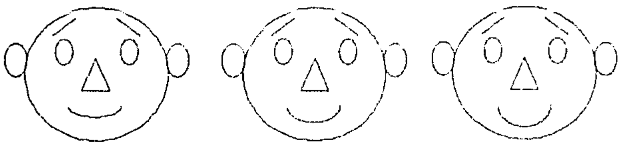
$$\text{mas } x = x^{1.25}$$

y la operación de concentración (CON) y la forma que asume la función es la siguiente:

$$M(\text{altamente alto}) = \text{mas}(CON(\text{alto}))$$

Estas funciones de compatibilidad para cada término nos darán el significado de cada caso de la variable sueldo; es decir con esto se podrá obtener el grado de pertenencia de las variables lingüísticas alto, muy alto o altamente alto; en otras palabras se analizará que tanto pertenece un valor a un subconjunto por ejemplo, si tenemos un valor de 15,000 podremos observar en que subconjunto tiene mayor grado de pertenencia si a alto, muy alto o altamente alto, dado que substituyendo 15,000 a cada una de las funciones de compatibilidad respectivamente obtendremos un valor en el intervalo de [0,1] y en la medida que el valor este más cercano a 1 podremos decir que el elemento pertenence más al conjunto de la variable lingüística.

A los rasgos faciales se les asignará una variable por ejemplo sueldo lo asignaremos a boca, que asumirá un valor lingüístico, ya que tomando en cuenta lo que se dijo en el capitulo 4 que una sonrisa puede representar una variable de éxito, podemos decir que una gran sonrisa puede representar un sueldo alto, ahora para muy alto y altamente alto la sonrisa debe mostrar mayor profundidad pues se esta hablando de mayor exito, esto quiere decir que la curvatura de la boca va a variar, por lo tanto tendremos entonces un rango de variación de la curvatura de la boca, que va depender de la etiqueta lingüística que se trate.



alto　　　muy alto　　　altamente alto

Los rasgos faciales asocian un tipo de dato que va a variar de acuerdo a valores difusos, la sintaxis o consulta agrupará varios datos de éste tipo, que formarán la cuestión.

Debe considerarse también que ésta sintaxis va contener conectivas entre las variables (rasgos) por ejemplo edad = joven y sueldo = alto; la conectiva y es muy distinta de las conectivas de los términos difusos, es decir la inclusión de un rasgo u otro va tener conectivas y/o; para definir estas conectivas usaremos color, si los rasgos que tratamos son del mismo color estaremos indicando que existe una conectiva y la cual indica que se tienen que cumplir las dos variables para que sea válida la respuesta y si el color es distinto se tratará de la conectiva o que quiere decir que con cualquiera que se cumpla será válida; para el caso de asignación de base de datos usaremos patrones o texturas en la carita o sea podremos tener carita blanca, carita rellena o algúna otra textura que diferencíe, teniendo finalmente un modelo gráfico de caritas que además de manejar la variación de sus rasgos faciales tendrá colores y texturas para agrupar más información.

Ejemplo Tomado el ejemplo 5.4.3. para consultas en lenguaje SEQUEL tenemos lo siguiente:

1. Tendremos una base de datos (Figura 5.1) la que contendrá la información precisa, de la cual se va obtener la información para las variables lingüísticas; para esto como primera instancia tendremos que hacer una asignación de rasgos faciales con las variables de la base de datos que se trate.

2. La asignación de rasgos faciales se llevará a cabo tomando de la estructura de base de datos un campo para cada rasgo facial, para este caso tendremos:

RASGO	VARIABLE ASOCIADA
Curvatura de la boca	Salario
Posición de la ceja	Edad
Nivel Orejas	Año de contratación

   los rasgos que no son asignados con una variable quedarán fijos, ésto quiere decir que los rasgos tomarán la forma de la representación de valores medios, esto es:

   La Cara será redonda, los ojos redondos, naríz ancha y la forma de las orejas será redonda.

3. Se contara con la base de datos de compatibilidades como la (Figura 5.4), es decir cada uno de los registros de las variables de la base de datos (Figura 5.1) maestra tendrá asociado su respetivo grado de pertenencia para la etiqueta lingüítica que se trate, estos valores se obtendrán mediante las funciones de pertenencia (compatibilidad).

4. Entonces las consultas con caritas se presentarán de la siguiente manera:

C1.

Con esta carita estaremos indicando lista los empleados que tengan un salario alto y que sean jovenes o bien que su contratación sea reciente.

Podemos darnos cuenta que la boca expresa sonrisa estaremos hablando de salario alto, las cejas tienen una inclinación hacia arriba se trata de joven ya que si estuviera hacia abajo indicaría que se trata de viejo, las orejas estan a nivel de la boca por lo tanto de trata de un valor reciente.

En lenguaje SEQUEL la sintaxis tomaría la siguiente forma:

```
SELECT Nombre
FROM Emp
WHERE   Salario="ALTO"   AND   (Edad="JOVEN"   OR
AñoCont="RECIENTE')
```

Ahora internamente lo que realizará el interprete en base a conjuntos difusos[1] sería:

Alto = 1.0/x1 + 0.5/x3 + 0.5/x4 + 0.8/x5

Joven = 1.0/x1 + 0.5/x2 + 0.5/x3 + 1.0/x5

1 Para los valores de las etiquetas linguísticas ver el Apendice A.

Reciente = 0.6/x2 + 0.6/x3 + 0.8/x5

Joven o Reciente= 1.0/x1 + 0.6/x2 + 0.6/x3 + 1.0/x5

Alto y (Joven o Reciente)= 1.0/x1 + 0.5/x3 + 0.8/x5

los resultados obtenidos son:

ZECUA
CHAPA
RUIZ

C2.

Con está carita indicaremos lista todos los empleados cuyo salario sea muy alto y sean jovenes o recien contratados.

Podemos notar en la variación de la curvatura de la boca que es más profunda que en el caso de la primera carita, los rasgos para edad y contratación son los mismos por lo tanto representa a joven y reciente respectivamente.

En SEQUEL la consulta sería:

```
SELECT Nombre
FROM Emp
WHERE  Salario="MUY  ALTO"  AND  (Edad="JOVEN"  OR
AñoCont="RECIENTE')
```

Con el interprete de conjuntos difusos:

Muy Alto = 0.9/x1 + 0.25/x3 + 0.25/x4 + 0.64/x5

Muy Alto y (Joven o Reciente)=0.9/x1+ 0.25/x3 + 0.64/x5 los resultados obtenidos son:

ZECUA
CHAPA
RUIZ

C3.

Lista los empleados cuyo salario sea altamente alto y sean viejos o año de contratación reciente.

En SEQUEL la sintaxis es:

```
SELECT Nombre
FROM Emp
WHERE Salario="ALTAMENTE ALTO" AND (Edad="VIEJO" OR
AñoCont="RECIENTE')
```

Altamente Alto = $0.9/x_1 + 0.18/x_3 + 0.18/x_4 + 0.57/x_5$

Viejo = $0.5/x_4$

Reciente = $0.6/x_2 + 0.6/x_3 + 0.8/x_5$

Viejo o Reciente = $0.6/x_2 + 0.6/x_3 + 0.5/x_4 + 0.8/x_5$

Altamente Alto y (Viejo o Reciente)= $0.18/x_3 + 0.18/x_4 + 0.57/x_5$

los resultados son:

CHAPA
LEON
RUIZ

los resultados obtenidos son:

ZECUA
CHAPA
RUIZ

C3.

Lista los empleados cuyo salario sea altamente alto y sean viejos o año de contratación reciente.

En SEQUEL la sintaxis es:

```
SELECT Nombre
FROM Emp
WHERE Salario="ALTAMENTE ALTO" AND (Edad="VIEJO" OR
AñoCont="RECIENTE')
```

Altamente Alto = 0.9/x1 + 0.18/x3 + 0.18/x4 + 0.57/x5

Viejo = 0.5/x4

Reciente = 0.6/x2 + 0.6/x3 + 0.8/x5

Viejo o Reciente = 0.6/x2 + 0.6/x3 + 0.5/x4 + 0.8/x5

Altamente Alto y (Viejo o Reciente)=0.18/x3 + 0.18/x4 + 0.57/x5

los resultados son:

CHAPA
LEON
RUIZ

C4.

Lista los empleado con Salario no alto y edad viejo o año de contratación no reciente

En SEQUEL la consulta se presenta de la siguiente forma:

```
SELECT Nombre
FROM Emp
WHERE Salario="NO ALTO" AND (Edad="VIEJO" OR
AñoCont="NO RECIENTE')
```

Con el interprete de cojuntos difusos:

No Alto = 1.0/x2 + 0.5/x3 + 0.5/x4 + 0.2/x5

Viejo = 0.5/x4

No Reciente = 1.0/x1 + 0.4/x2 + 0.4/x3 + 1.0/x4 + 0.2/x5

Viejo o No Reciente = 1.0/x1 + 0.4/x2 + 0.4/x3 + 1.0/x4 + 0.2/x5

No alto y (Viejo o No Reciente) = 0.4/x2 + 0.4/x3 + 0.5/x4 + 0.2/x5

los resultados quedan:

REYES
CHAPA
LEON
RUIZ

C5.

Esta carita indica lista los empleados cuyo Salario sea no alto y sean jovenes y año de contratación muy reciente.

```
SELECT Nombre
FROM Emp
WHERE  Salario="NO  ALTO"  AND  (Edad="JOVEN"  AND
AñoCont="MUY RECIENTE')
```

No Alto = 1.0/x2 + 0.5/x3 + 0.5/x4 + 0.2/x5

Joven = 1.0/x1 + 0.5/x2 + 0.5/x3 + 1.0/x5

Muy Reciente = 0.36/x2 + 0.36/x3 + 0.64/x5

Joven y Muy Reciente = 0.36/x2 + 0.36/x3 + 0.64/x5

No Alto y (Joven y Muy Reciente) = 0.36/x2 + 0.36/x3 + 0.2/x5

los resultados son:

REYES
CHAPA
RUIZ

5. El sistema de recuperación se basa en el procesamiento de una consulta difusa como es mencionado en la sección 5.4 de este capitulo y los punto 1, 2 y 3 de la sección 5.5 del mismo, que son los componentes principales de un sistema de base de datos para el procesamiento de una consulta difusa (Figura 5.6), aquí se puede observar que la consulta es mediante el lenguaje SEQUEL y el Parser analiza el bloque SELECT, FROM y WHERE, para el

caso de caritas la consulta será el icono cara y el Parser analizara los seis rasgos faciales básicos Figura 5.8.

Podemos observar que las carita forman una sintaxis de consulta, la cual resulta sencilla y que asociada a términos difusos nos ofrece un lenguaje más semejante al pensamiento de los humanos, que es un tanto de información aproximada o imprecisa, que finalmente se comunica al mundo real.

El planteamiento icónico de caritas, en éste trabajo es tratado bajo dos aspectos: 1) el visual, que como se menciono en el capítulo 1, las gráficas o imágenes son de gran potencial, ya que nos ofrecen una percepción rápida, y que generalmente como medio de comunicación nos dan un significado completo de lo que se observa; y las caritas en este caso tienen gran aceptación, y 2) el aspecto enfocado a a resolver problemas de tipo difuso, ésto es, con éste modelo gráfico se plantea un sintaxis, la cual mediante rasgos faciales forma la consulta en base a términos o variables lingüísticas (capítulo 2 y 3). Estos dos aspectos se conjugan dando como resultado un modelo como es mostrado en la Figura 5.8, el que es resuelto mediante un sistema computacional en el ambiente de base de datos.

Ahora, como presentación a éste modelo gráfico, en el siguiente punto muestro una sesión con el Editor de caritas.

5.7 Sesión con Editor de caritas de Chernoff.

Esta parte la dedico principalmente a mostrar el modelo gráfico de caritas para formular una sintaxis de consulta. Lo que a continuación presento son pantallas que componen el Editor, así como algunos ejemplos de como se presenta una consulta con la sintaxis gráfica de rasgos faciales.

El Editor de Caritas de Chernoff es un sistema que está desarrollado en el lenguaje de programación Pascal. La pantalla 1 nos muestra las opciones de:

- Base de Datos, en donde se da la estructura de la base de datos, a la que se le asignarán los rasgos faciales.

- Asignación de Rasgos, en ésta opción se asignarán las variables de la base de datos con los rasgos faciales de la carita.

- En la opción Consulta se formará la sintaxis con los rasgos de la carita.

- Y la última opción es salir.

Si seleccionamos la opción Asignación de Rasgos (Pantalla 2) nos muestra un desplegado de los archivos que se pueden plantear como de consulta, en ésta pantalla se pide el nombre de tal archivo.

Una vez dado el nombre del archivo estructura, se muestra la pantalla en donde se lleva a cabo la asignacion de los rasgos faciales (Pantalla 3 y 4), ya asignados los rasgos faciales al archivo estructura podemos entonces, pasar a la opción de Consulta (Pantalla 5), en donde se despliega las variables con los rasgos faciales asignados y que nos permitirán hacer la sintaxis de consulta de acuerdo a las etiquetas lingüísticas.

Por ejemplo en la Pantalla 6, se muestra la variable Edad, la cual se representa por medio de la inclinación de la ceja, en este caso la variable Edad toma el valor lingüístico Joven y si vemos la inclinación de la ceja que se muestra en ésta pantalla podemos observar la variación, que se da de acuerdo a la etiqueta lingüistica que represente, aquí podemos ver que cuando la ceja este totalmente hacia arriba representa la etiqueta completamente, a la que se le une el valor de Joven. Ahora, en la Pantalla 7, se muestra la inclinación de la ceja hacia abajo, y en este caso representa la etiqueta no completamente, lo que indicaría la negación de la sintáxis de consulta de la Pantalla 6.

En la Pantalla 8, se muestra la variable promedio con la amplitud de la frente, aquí podemos ver que la variación de tal rasgo representa una etiqueta, la primera carita representa excelente, la segunda muy excelente, la tercera altamente excelente y la cuarta carita representa a la etiqueta completamente excelente, si observamos cada una de éstas caritas notaremos la variación que presentan.

En la Pantalla 9 y 10, se representa la variable sueldo con la curvatura de la boca, en este caso el valor lingüístico que se toma es alto y las distintas caritas muestran una etiqueta unida a éste valor lingüístico. Podemos ver que la cuarta carita de la pantalla 9 presenta la etiqueta completamente y la tercera carita de la pantalla 10 representa la etiqueta no completamente.

En la Pantalla 11, podemos ver que la forma de la cara representa a la variable peso y los valores lingüísticos que puede asumir ésta variable son robusto, normal y delgado, entonces cada una de las caritas mostradas en orden representan a tales valores, y podemos notar que la tercer carita representa al valor lingüístico delgado. Y así, podemos continuar haciendo distintas sintaxis e incluso combinar rasgos, como en la pantalla 12, en donde se representa la variable peso con la forma de la cara y la variable sueldo con la curvatura de la boca.

EDITOR DE CARITAS DE CHERNOFF

Base de Datos   Asignacion de Rasgos   Consulta   Salir

Pantalla 1

EDITOR DE CARITAS DE CHERNOFF

DATOS.EST
ARCHI1.EST

Nombre Archivo Estructura: [datos   .EST]

Asignando Rasgos Faciales ...

Pantalla 2

EDITOR DE CARITAS DE CHERNOFF

datos.EST

VARIABLE	RASGO	TERMINOS
NOMBRE	Ninguno	
SEXO	Ninguno	
EDAD	Inclinacion Ceja	Joven,Viejo
DIRECCION	Ninguno	
CARRERA		
EDOCIVIL		
NACIONALIDAD		
TELEFONO		
TITULONOM		
INSTITU		
PUESTO		
RFC	Ninguno	
PROMEDIO	Amplitud Frente	Excelente,Bueno,Regular,Malo
SUELDO	Curvatura Boca	Alto,Regular,Bajo

Ins Acepta Rasgo DnArrow Cambia Campo End Fin Asignacion PgDn Pag. Sig.

Pantalla 3

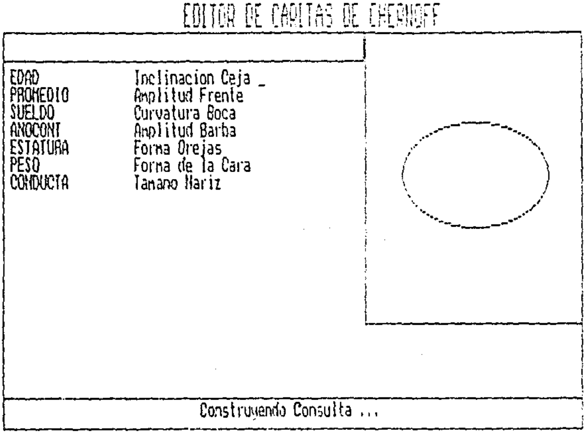
EDITOR DE CARITAS DE CHERNOFF

datos.EST

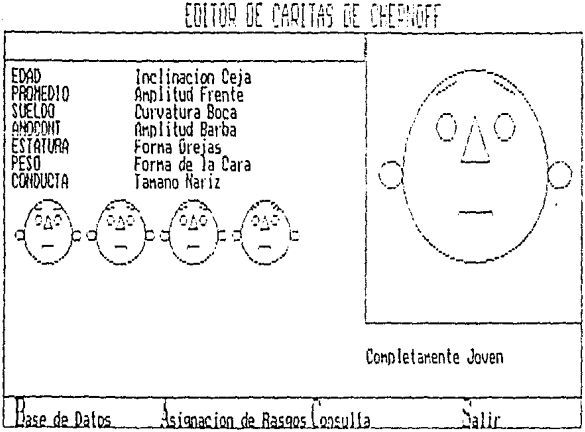
VARIABLE	RASGO	TERMINOS
ANOCONT	Amplitud Barba	Reciente
ESTATURA	Forma Orejas	Alto,Regular,Bajo
PESO	Forma de la Cara	Robusto,Normal,Delgado
CONDUCTA	Tamano Nariz	Excelente,Bueno,Regular,Malo
ASISTENCIA		

Ins Acepta Rasgo DnArrow Cambia Campo End Fin Asignacion PgDn Pag. Sig.

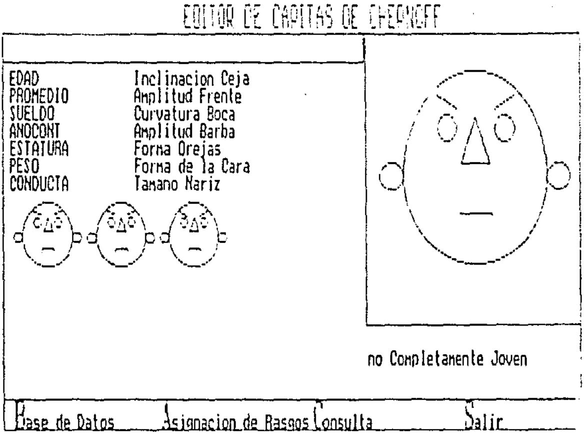
Pantalla 4



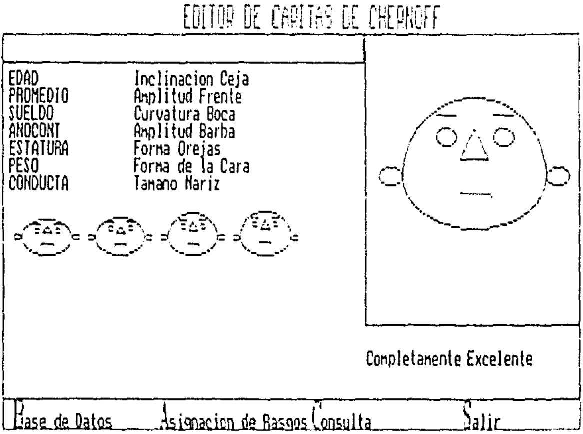
Pantalla 5



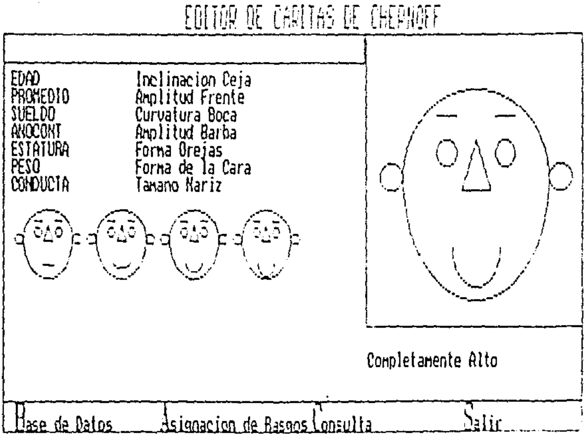
Pantalla 6



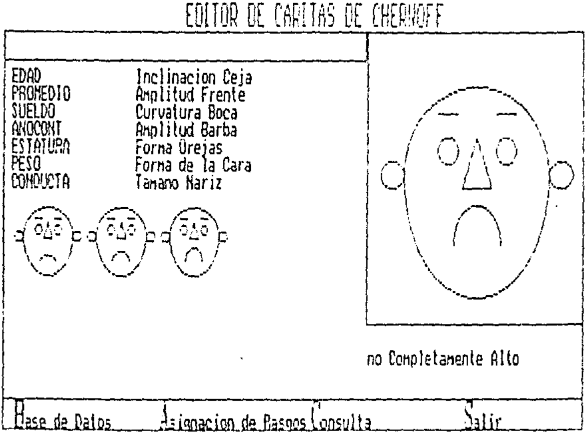
Pantalla 7



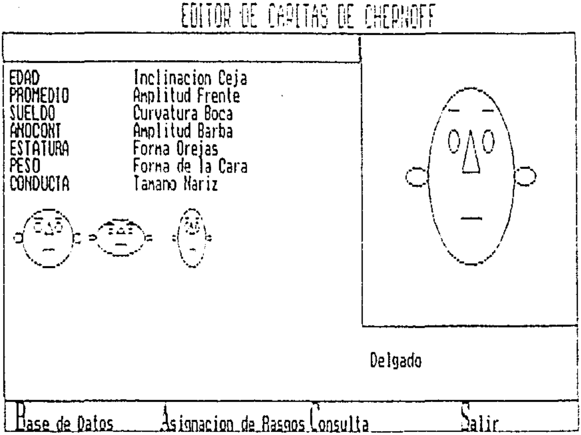
Pantalla 8



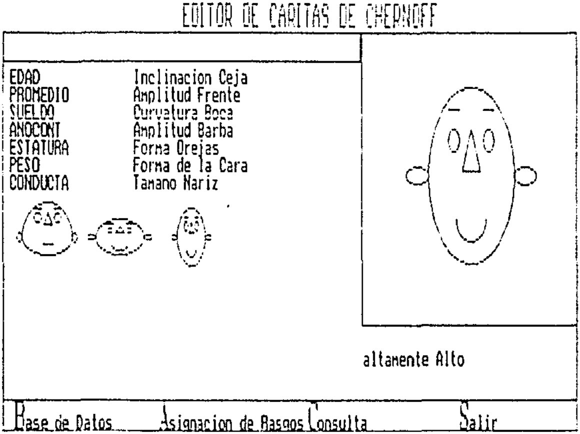
Pantalla 9



Pantalla 10



Pantalla 11



Pantalla 12

CONCLUSIONES

La conclusión principal de este trabajo consiste en proponer las caritas de Chernoff como una representación gráfica para el procesamiento de consultas difusas. El antecedente es que las caritas de Chernoff sirven para representar clases de datos multivariados. En este trabajo tratamos con iconos de caritas como un medio de programación visual para plantear consultas a una base de datos. Las caritas funcionan como operadores de clasificación de datos.

El primer resultado es un nuevo modelo visual basado en iconos de caritas, que pueden ser creadas con un editor gráfico que maneja variaciones en los rasgos como son: boca, nariz, cejas y orejas. Con esto se pretende dar una sintaxis de tipo visual con un semántica que se asocia a operadores de alto orden.

Las caritas editadas son interpretadas mediante un analizador sintáctico para formular una consulta de tipo difuso. El poder procesar una consulta difusa es el segundo resultado que se tiene en este trabajo. Los rasgos se asocian a etiquetas lingüísticas y mediante un evaluador difuso se pueda recuperar un subconjunto difuso de un conjunto de tuplas, de la base de datos.

La recuperación se lleva a cabo considerando a los datos asociados que cumplan con el mayor grado de pertenencia y mandar una salida de datos ordenados.

Actualmente existe un sistema denominado Manipulador de Conjuntos Difusos[34], que consiste en representar y manejar conjuntos difusos por medio de un Parser descendente recursivo, que reconoce expresiones del tipo mencionado en los capítulos 2 y 3, este programa está códificado en lenguaje C y puede ser una parte del interprete ya mencionado.

Para llevar a cabo el tratamiento de consultas difusas matemáticamente en la teoría de Conjuntos Difusos o también conocido como Lógica Difusa; presento definiciones, conceptos y ejemplos que dan el soporte de lo que es llamado variables lingüísticas, que finalmente estas nos permiten elaborar las consultas difusas y que en combinación con las caritas nos muestran un nuevo tipo de sintaxis con una semántica que asocia a operadores de alto orden, permitiendo un procesamiento más inteligente con una clasificación automática de datos. Las variaciones de los rasgos faciales en formas, posiciones y colores nos llevan a tener un sistema con más opciones y porqué, no, un sistema más agradable.

BIBLIOGRAFIA

[1] Beniger, James R. y Robyn, Dorothy L. "Quantitative Graphics in Statistics: A Brief History." American Statistician , 32:1, 1978.

[2] Fienberg, S. E., "Graphical Methods in Statistics." Technical report No. 304, Dept. of Applied Statistics, University of Minnesota, pp. 44.

[3] Kruskal, William. "Visions of Mapas and Graphs." Proceedings of the International Symposium on Computer assisted Cartography, 1975, pp. 27 (1977).

[4] Royton, E. "Studies in the History of Probability and Statistics", A Note on the History of the Graphical Presentation of Data", Biometrika, Pts. 3 y 4 (diciembre 1956) pp. 241.

[5] Glinert, E. y Tanimoto, S. "PICT: An Interactive Graphical Programming Environment". Computer IEEE, Noviembre 1984. pp. 7-25.

[6] Chapa Vergara S., "Lenguaje de Flujograma como un Lenguaje de Consulta a una Base de Datos (Un Estudio Comparativo)".

[7] Guzmán A. A., "The File Descriptor: Use of a Descriptive Tool to Retrive General Queries to Files.", Reporte Técnico No. 16, Serie Amarilla, Depto. de Ingeniería Eléctrica, CINVESTAV-IPN, mayo 10, 1985.

[8] Dickinson G. C., Statistical mapping and the presentation of statistics, Edward arnold, Second edition.

[9] Schmid C. F., Statistical Graphics: Design Principles y Practrices, Wiley New York, 1983.

[10] James M., Diagramming Techniques for analysis and Programmes, Prentice-Hall, 1985.

[11] James M., Fourth-Generation Languajes, Vol.I, Prentice-Hall, 1985.

[12] Wang Peter C. C., Graphical Representation of Multivariate Data, Symposium on Graphical Represetation of Multivariate Data, Naval Postgraduate School, Monterey, Calif., Academic Press, 1978.

[13] Chatfield, C., Introduction to Multivariate Analysis, Science Paperbacks, 1980.

[15] Zadeh L. A., Semantics of context-free languajes, Math. Syst. Theory, 2, 127-145 (1968).

[16] Zadeh L. A., Quantitative fuzzy semantics, Inf. Sci. 3, 159-176 (1971).

[17] Knuth D., Semantics of context-free languages, Math. Syst. Theory, 2, 127-145 (1968).

[18] Aho A. V., and Ullman J. D., The Theory of Parsing, Translation and Compiling. Prentice-Hall, Englewood Hills, N. J., 1973.

[19] Codd E.F., A Data Base Sublanguage Founded on the Relational Calculus, Proc 1971 ACM SIGFIDET Workshop on Data Description, Access and Control, Available from ACM.

[20] Chamberlin D. D. and Boyce R. F., SEQUEL: A Structured English Query Language, pp. 249-264. Proc. ACM SIGFIDET, Ann Arbor, Michigan, May 1974.

[21] Held G. D., Stonebraker M. R. and Wong E., INGRES- A Relational Data Base System, Proc. AFIPS 1975 NCC, AFIPS Press, Montvale, N.J.

[22] Zloof M. M., Query-By-Example: The Invocation and Definition of Tables and Forms, Proc. Very Large Data Bases 1975, Framingham, Massachusetts, Available from ACM.

[23] Date C. J., An Introduction to Data Base Systems. Addison-Wesley, Reading, Mass. (1975).

[24] Newell A. and Simon H. A., Computer science as empirical inquiry: symbols an search, Comm. ACM 1976, 19, 113-126.

[25] Zadeh L. A., A fuzzy-set-theoretic interpretation of linguistic hedges. J. Cybern 1972, 2,4-34.

[26] Zadeh L. A., A fuzzy sets. J. Inform. Control 1965 8,338-353.

[27] Zadeh L. A., The concept of a linguistc variable and its application to aproximate reasoning-I. Inform. Sci. 1975,8, 199-249.

[28] Salton G., Automatic Information Organizationan Retrieval. McGraw-Hill, New York (1968).

[29] Tahani V., A general conceptual framework for Informationretrieval systems. Ph.D. Thesis, University of California, Berkeley, CA, 1972.

[30] Zadeh L. A., The concept of a linguistc variable and its application to aprocimate reasoning-II. Inform. Sci. 1975, 8, 301-358.

[31] Astrahan M. M. and Chamberlin D. D., Implementation of a structured English Query Lenguaje. Comm. ACM 1975, 18, 580-588.

[32] Codd E. F., A relational model of data for large shared data banks. Comm. ACM 1970,13, 377-387.

[33] Chapa Vergara S. y Zecua Fernández C., Caritas como Representación Iconica para el Procesamiento de Consultas Difusas en Lenguaje de Flujograma. Quinta Conferencia Internacional UNISYS y la UNAM 1989.

[34] Chapa Vergara S. y González Ambriz E., Manipulador de Conjuntos Difusos. CINVESTAV-IPN. (Por Publicar).

[35] Chapa Vergara S., Programación Automática a partir de Descriptores de Flujo de Información. Tesis de Doctorado. CINVESTAV-IPN, México, D. F., 1991.

APENDICE A

CONSULTAS / REGISTROS	VIEJO	JOVEN	ALTO	RECIENTE	MUY ALTO	ALTAMENTE ALTO	NO ALTO	NO RECIENTE	MUY RECIENTE	NO VIEJO	NO JOVEN
X1	0.0	1.0	1.0	0.0	0.9	0.9	0.0	1.0	0.0	1.0	0.0
X2	0.0	0.5	0.0	0.6	0.0	0.0	1.0	0.4	0.36	1.0	0.5
X3	0.0	0.5	0.5	0.6	0.25	0.13	0.5	0.4	0.36	1.0	0.5
X4	0.5	0.0	0.5	0.0	0.25	0.18	0.5	1.0	0.0	0.5	1.0
X5	0.0	1.0	0.8	0.8	0.64	0.57	0.2	0.2	0.64	1.0	0.0

Base de Datos de Compatibilidades