

# UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

---

DIVISION DE ESTUDIOS DE POSGRADO  
DE LA FACULTAD DE INGENIERIA

## PLANTEAMIENTO Y PRUEBA DE UN NUEVO METODO PARA REALIZAR EL PROCESO DE DEFUZZIFICACION EN CONTROLADORES DIFUSOS

T E S I S  
QUE PARA OBTENER EL GRADO DE  
MAESTRO EN INGENIERIA (ELECTRICA)  
P R E S E N T A  
PONCIANO JORGE ESCAMILLA AMBROSIO

## DEDICATORIAS

- ◆ A MIS PADRES Y HERMANOS, GRACIAS POR SU APOYO Y COMPRENSIÓN.
  
- ◆ A MI ESPOSA LUPITA, GRACIAS POR TU PACIENCIA Y APOYO.
  
- ◆ A MIS AMIGOS RUBÉN Y ALFREDO, SÉ QUE SIEMPRE SEREMOS AMIGOS.

## AGRADECIMIENTOS

◆ AGRADEZCO A LOS SIGUIENTES CATEDRÁTICOS QUE HICIERON FAVOR DE REVISAR ESTE TRABAJO DE TESIS:

M. ING. JOSE LUIS PÉREZ SILVA

DR. FELIPE LARA ROSANO

DR. TANG XU YU

M. EN C. ARTURO HARO RUIZ

M. EN C. ALBERTO HERRERA BECERRA

◆ UN ESPECIAL AGRADECIMIENTO AL M. JOSE LUIS PÉREZ SILVA, ESPERO SEGUIR CONTANDO CON SU APOYO Y AMISTAD.

◆ A LA FACULTAD DE INGENIERIA, EN ESPECIAL AL DEPARTAMENTO DE ELECTRÓNICA QUE ME OTORGÓ LAS FACILIDADES PARA REALIZAR ESTE TRABAJO DE TESIS.

◆ AL INSTITUTO DE INVESTIGACIONES ELÉCTRICAS POR FACILITARME EL MODELO DINÁMICO DEL PROCESO EN CUESTIÓN.

---

---

# ÍNDICE

---

---

1	INTRODUCCIÓN	1
2	LÓGICA DIFUSA	4
2.1	INTRODUCCIÓN	4
2.2	FILOSOFÍA DE LO DIFUSO	5
2.2.1	Desarrollo Histórico del Tratamiento de la Vaguedad	5
2.2.2	Lógica Difusa y Conjuntos Difusos : Una Forma de Manipular la Vaguedad e Imprecisión	7
2.3	NOTACIÓN, TERMINOLOGÍA Y OPERACIONES BÁSICAS CON CONJUNTOS DIFUSOS	9
2.3.1	Conjuntos Difusos y Terminología	9
2.3.2	Conjunto Teórico de Operaciones Básicas con Conjuntos Difusos	12
2.4	RELACIONES DIFUSAS	15
2.4.1	Operaciones con Relaciones Difusas	18
2.4.2	El principio de Extensión	23
2.5	VARIABLE LINGÜÍSTICA	23
	BIBLIOGRAFÍA	27
3	SISTEMAS DE CONTROL DIFUSO	28
3.1	INTRODUCCIÓN	28
3.2	SISTEMAS DE CONTROL CONVENCIONALES	29
3.2.1	Componentes de un Sistema de Control Convencional	29
3.2.2	Características Fundamentales de un Proceso	29
3.2.3	Formas de Control Convencionales	30
3.2.3.a	Control de Dos Posiciones (Todo-Nada)	30
3.2.3.b	Control Flotante	31
3.2.3.c	Control Proporcional	31
3.2.3.d	Control Integral (de Reposición)	32
3.2.3.e	Control Derivativo	33
3.2.3.f	Control Proporcional más Integral (PI)	33
3.2.3.g	Control Proporcional más Derivativo (PD)	34
3.2.3.h	Control Proporcional más Integral más Derivativo (PID)	35
3.3	DESCRIPCIÓN DE LOS SISTEMAS DE CONTROL DIFUSO (SCD)	36
3.3.1	Fuzzificación (Fuzzification)	38
3.3.2	Evaluación de Reglas (Rule Evaluation)	38
3.3.3	Defuzzificación (Defuzzification)	46
3.3.3.a	Método del Centro de Gravedad o Centroide ( Center Of Area method COA)	46
3.3.3.b	Método del Máximo ( Max Criterion Method MAX)	46
3.3.3.c	Promedio de Influencias (Mean of Maximum Method MOM)	47
3.4	MODELOS DE SISTEMAS DE CONTROL DIFUSO (SCD)	47
3.4.1	Modelo Mamdani	47
3.4.2	Modelo Sugeno-Takagi	48
3.4.3	Modelo Tsukamoto	48
3.5	SISTEMAS DE CONTROL DIFUSO COMO APROXIMADORES UNIVERSALES	49
3.5.1	SCD con Consecuentes Difusos son Aproximadores Universales	49
3.5.2	SCD con Consecuentes no Difusos son Aproximadores Universales	52
	BIBLIOGRAFÍA	54

<b>4. NUEVO MÉTODO PARA REALIZAR EL PROCESO DE DEFUZZIFICACIÓN EN CONTROLADORES DIFUSOS</b>	<b>55</b>
<b>4.1 INTRODUCCIÓN</b>	<b>55</b>
<b>4.2 CENTRAL TERMOELECTRICA DE CICLO COMBINADO</b>	<b>55</b>
4.2.1 Equipos Principales	56
<b>4.3 QUEMADORES POSTERIORES</b>	<b>58</b>
<b>4.4 CONTROL DIFUSO DE LOS QP DE UNA CTCC</b>	<b>60</b>
4.4.1 Planteamiento del Problema	60
4.4.2 Metodología	61
4.4.3 Identificación de las Fuentes de Control	61
4.4.4 Definición de Funciones de Membresía y Conjuntos Difusos	63
4.4.5 Elaboración de Reglas	65
4.4.6 Selección del Método de Defuzzificación	66
4.4.7 Simulación del Sistema de Control Difuso	66
<b>4.4.7.A Modelo Dinámico del Proceso</b>	<b>66</b>
4.4.7.A.1 Descripción	66
4.4.7.A.2 Características del Modelo Dinámico del Proceso	67
4.4.7.A.3 Modelo Matemático de la Válvula	67
4.4.7.A.4 Modelo Matemático del Proceso	67
4.4.7.A.4.1 Principales Fórmulas del Modelo Matemático del Proceso	68
<b>4.4.7.B Algoritmo del SCD</b>	<b>71</b>
<b>4.5 PRUEBAS AL SISTEMA DE CONTROL DIFUSO SIN EL NUEVO MÉTODO DE DEFUZZIFICACIÓN</b>	<b>72</b>
4.5.1 Prueba con una Perturbación Escalón	73
<b>4.6 NUEVO MÉTODO PARA REALIZAR EL PROCESO DE DEFUZZIFICACIÓN EN LOS SCD</b>	<b>74</b>
4.6.1 Simplificaciones del método del centroide	75
4.6.2 Planteamiento del nuevo método para realizar el proceso de defuzzificación en los SCD	76
4.6.3 SCD con el nuevo método de defuzzificación y estructura reducida	77
4.6.4 Pruebas al sistema de control difuso con el nuevo método de defuzzificación	79
4.6.4.a) Prueba con una perturbación escalón	79
4.6.4.b) Prueba con una perturbación rampa 1	80
4.6.4.c) Prueba con una perturbación rampa 2	81
4.6.4.d) Prueba con una perturbación rampa 3	81
<b>BIBLIOGRAFIA</b>	<b>82</b>
<b>GRÁFICAS</b>	<b>83</b>
<b>5 CONCLUSIONES</b>	<b>98</b>
<b>APÉNDICE A</b>	<b>102</b>
<b>ANEXO 1</b>	<b>104</b>
<b>ANEXO 2</b>	<b>106</b>
<b>ANEXO 3</b>	<b>108</b>

---

# CAPÍTULO UNO

## INTRODUCCIÓN

---

### 1 INTRODUCCIÓN

La imprecisión y la incertidumbre son conceptos inherentes a la naturaleza inexacta del razonamiento humano. Debido a ello nuestra manera de interpretar el mundo generalmente es en función de proposiciones vagas, datos inciertos y juicios de valor. Sin embargo, esta imprecisión e incertidumbre no son tomadas en cuenta por la lógica tradicional (crisp logic) en donde sólo existen dos premisas fundamentales: verdadero y falso. En el pasado varios pensadores se percataron de este hecho y se dieron a la tarea de estudiar la vaguedad. Esto con el fin de tratar de crear una estructura matemática que permitiera captar esta característica del lenguaje natural.

El primer pensador que trató seriamente con la vaguedad fue el filósofo estadounidense Sanders Peirce (1839-1914). Peirce decía que todo cuanto existe es continuo y que, por tanto, tal continuidad gobierna el conocimiento. Aseguraba que la vaguedad esta presente en todas partes aún en las palabras. Afirmaba, "*la vaguedad no se puede descartar del mundo de la lógica así como la fricción de la mecánica*" [1]. Posteriormente Bertrand Russell (1872-1970) fue quien comenzó el estudio moderno de la vaguedad. Russell fue aún más lejos que Peirce al afirmar "*la vaguedad claramente es materia de gradación*". Sin embargo, por allá de 1920 el filósofo polaco Jan Lukasiewicz (1878-1955) fue el primero en dar los pasos para formalizar un modelo de la vaguedad. Lukasiewicz inventó la subestructura de los conjuntos difusos, una lógica basada en más valores que cierto y falso. Él estableció, en adición el valor de 1 para verdadero y 0 para falso, el valor de 1/2 para posible. Así pues, una sentencia podía tener cualquiera de esos tres valores. Aparentemente lo anterior fue un paso muy sencillo, sin embargo fue un avance muy importante. Más tarde Max Black (1909-1989) fue el hombre que estableció el camino a seguir en el estudio de la vaguedad. Black planteó sus ideas en su artículo "*Vagueness: An Exercise in Logical Analysis*" publicado en 1937 [2]. En ese artículo Black manifestó estar de acuerdo con Peirce en el sentido de que la vaguedad representaba un continuo. También aceptó estar de acuerdo con Russell en que la vaguedad era objeto de gradación. De hecho, afirmó que lo continuo implicaba grados. Black sugirió que los términos vagos podían formar conjuntos ("conjuntos vagos") e incluso dibujó las primeras curvas de lo que posteriormente serían los conjuntos difusos [2]. Él pensó que con estos conjuntos se podrían hacer operaciones, así como permitirían la existencia de subconjuntos. También observó que un conjunto podía ser un subconjunto parcial de otro. De tal manera, Black había desarrollado una nueva estructura intelectual, pero nunca junto las piezas [1].

El hombre que junto las piezas fue Lotfi A. Zadeh quien en 1965 publicó "Fuzzy Sets" (conjuntos difusos) [3], artículo que marcó un hito histórico y proporcionó a la disciplina su nombre. Con el fin de captar la vaguedad de nuestro razonamiento al interpretar el mundo, en este artículo Zadeh propuso la "*teoría de los conjuntos difusos*". La característica principal de esta teoría es que en ella un elemento ya no se restringe a ser totalmente parte o no de un conjunto, ahora un elemento puede pertenecer parcialmente a un conjunto dado y al mismo tiempo puede parcialmente no pertenecer a dicho conjunto. Es decir, aquí un elemento pertenece a un conjunto en un grado que esta en el rango  $[0, 1]$ . Así, por ejemplo, si un elemento pertenece a un conjunto dado en un grado de 0.7, entonces ese mismo elemento pertenece al complemento del mismo conjunto en un grado de 0.3. Evidentemente, esta es una extensión de la lógica tradicional que incluye a aquella en los extremos. Un grado de pertenencia de cero indica total no pertenencia y un grado de pertenencia de uno indica total pertenencia. Expresado en otras palabras, mientras en la teoría de conjuntos tradicional un conjunto tiene bordes bruscos, los de un conjunto difuso son suaves y se van difuminando. Lo anterior permite caracterizar con conjuntos difusos aquellos conceptos cuyas fronteras no están bien definidas.

A partir de la teoría de los conjuntos difusos se desarrolló la denominada "*lógica difusa*" (fuzzy logic) [4]. En esencia la lógica difusa esta mucho más relacionada al pensamiento humano y al lenguaje natural que a los sistemas lógicos tradicionales. La lógica difusa es la estructura matemática a través de la cual es posible capturar la naturaleza inexacta con la que los humanos perciben el mundo real. Con esta lógica es posible manejar conceptos vagos para caracterizar las variables de un sistema y sus interrelaciones utilizando palabras o proposiciones expresadas en un lenguaje natural o artificial. Así la lógica difusa utiliza conjuntos difusos para diseñar sistemas que son capaces de captar en forma de reglas heurísticas la habilidad que todo ser humano posee para modelar un sistema o proceso utilizando el lenguaje natural.

Durante las pasadas tres décadas la teoría de los conjuntos difusos y la lógica difusa han sido desarrolladas y han alcanzado cierta madurez. Así mismo, se ha dado un considerable desarrollo de las aplicaciones de la teoría de los conjuntos difusos en la solución de un sinnúmero de problemas. En la actualidad la teoría de los conjuntos difusos y la lógica difusa se aplican en una gran variedad de disciplinas, tales como: control automático, procesamiento de señales, manejo de base de

datos, visión computarizada, toma de decisiones y muchas otras. De hecho el mayor éxito se ha alcanzado en el área del control automático. Aquí los denominados sistemas de control difuso (SCD) o simplemente controladores difusos han emergido como una de las áreas de investigación más activas y fructíferas en la aplicación de la teoría de los conjuntos difusos [5] (el 90% de las aplicaciones exitosas reportadas en el Japón son en el área del control automático [6]).

El primer controlador difuso se realizó de manera experimental en 1973. En ese año el Profesor E. Mamdani de la Universidad de Londres empleó, por primera vez, la lógica difusa para diseñar el control automático de una pequeña máquina de vapor. La publicación, en 1975 [7], de los resultados obtenidos por Mamdani generó un gran interés en esta área. Así, después de esta aplicación otras aparecieron. En Holanda V. N. Lemke and W. J. M. Kickert desarrollaron un controlador difuso para regular la temperatura en una planta de tratamiento de agua [1]. En Dinamarca P. M. Larsen y J.-J. Ostergaard trabajaron en un intercambiador de calor difuso [1]. Estas aplicaciones fueron el preámbulo para el desarrollo de la primera aplicación industrial.

La primera aplicación de un SCD en el ámbito industrial fue realizada en Dinamarca por Holmblad y Ostergaard. En 1980 ellos realizaron la instalación de un controlador difuso para controlar un horno industrial de una fábrica de cemento [8]. Esta fue también la primera aplicación comercial de la lógica difusa. Los resultados obtenidos con este controlador demostraron los beneficios que se podrían obtener con un controlador difuso. En este caso, el controlador difuso superó a un operador humano (que anteriormente era el que realizaba el control) en la ejecución del control y además redujo el consumo de combustible. A partir de entonces los SCD se han aplicado en una gran variedad de procesos que van desde el ámbito industrial hasta los electrodomésticos [9-10].

La parte esencial de un SCD es un conjunto de reglas de control expresadas en términos lingüísticos. Estas reglas están relacionadas por los conceptos duales de implicación difusa y regla composicional de inferencias. Básicamente un SCD provee un algoritmo que convierte las estrategias de control expresadas de manera lingüística, y basadas en el conocimiento empírico que se tiene, dentro de una estrategia de control automático. En la práctica se ha observado que en muchos procesos un SCD arroja resultados superiores a los obtenidos con algoritmos de control convencionales. En particular, la metodología de los SCD ha resultado de gran utilidad en los casos en que los procesos son demasiado complejos para ser analizados por técnicas cuantitativas convencionales o cuando las fuentes de información disponibles son interpretadas cualitativamente, de manera inexacta o de manera incierta [11].

El objetivo de este trabajo es el plantear y probar un nuevo método para realizar el proceso de defuzzificación en controladores difusos. Este nuevo método fue el resultado de resolver el problema del control de los quemadores posteriores (QP) de una central termoeléctrica de ciclo combinado (CTCC). En un primer intento se propuso un SCD con dos entradas y una salida para realizar el control de los QP de la CTCC. Los resultados obtenidos de la simulación del SCD, una vez realizada la sintonización (ajustado reglas y conjuntos difusos), mostraron un deficiente desempeño del SCD al compararlo contra los datos existentes de un controlador tradicional tipo PI (acción proporcional más integral), actualmente empleado para tal efecto. Visto lo anterior se buscó la forma de lograr un eficiente control de los QP utilizando un SCD. Se determinó que las posibles causas de la deficiencia del SCD en la ejecución del control eran las características de carga, capacitancia y tiempos de transporte que el proceso involucraba, las cuales causaban que el mismo tuviera una dinámica lenta. La idea era encontrar una solución que tomara en cuenta estas características, pero al mismo tiempo la solución no debía provocar la inclusión de más variables en el SCD, lo cual provocaría un crecimiento exponencial del número de reglas y en consecuencia del tiempo de procesamiento del algoritmo de control. La solución consistió en proponer y probar un nuevo método para realizar el proceso de defuzzificación. Éste es una modificación del método del centro de área utilizando singletons. La característica principal de este nuevo método es la inclusión de elementos de retraso. De esta manera, los efectos de los retrasos inherentes a cualquier proceso industrial son enormemente reducidos.

Así pues con el fin de cumplir el objetivo citado, la organización de este trabajo es como sigue. Los capítulos 2 y 3 están dedicados a sentar las bases teóricas de este trabajo. En el capítulo 2 se exponen los conceptos fundamentales de la lógica difusa. Primeramente se expone la filosofía de lo difuso, se describe el desarrollo histórico del estudio de la vaguedad hasta llegar a la teoría de los conjuntos difusos y a la lógica difusa. Una vez conociendo los orígenes de la lógica difusa ésta se define como un medio para el tratamiento de la vaguedad y la imprecisión. Más adelante se presenta la notación, terminología y las operaciones básicas con conjuntos difusos. También se definen conceptos fundamentales en los SCD, las relaciones difusas, el principio de extensión y la variable lingüística.

En el capítulo 3 en primer instancia se define lo que es un sistema de control, se describen los sistemas de control convencional más ampliamente utilizados, esto como antecedente a la descripción de las tres etapas fundamentales de un sistema de control difuso (SCD): fuzzificación, evaluación de reglas y defuzzificación. Posteriormente se describen los modelos de SCD más utilizados. Finalmente se expone a los SCD como aproximadores universales de funciones.

El capítulo 4 es la parte esencial de este trabajo, inicialmente se describe la central termoeléctrica de ciclo combinado y su principio de operación. Se presentan los equipos principales de la misma. Bajo este entorno se identifican y ubican los quemadores posteriores y se describe su funcionamiento. Una vez descrito el proceso en cuestión se define el problema de control. Posteriormente se presenta la metodología bajo la cual se desarrollo la primer propuesta de SCD (SCD sin el nuevo método de defuzzificación). Bajo la metodología presentada, en esta sección se plantea y desarrolla el primer esquema de SCD. También en esta sección se describe la estructura de simulación y se presentan el modelo dinámico del proceso y el algoritmo de control. Más adelante se describen las pruebas realizadas al SCD, así como se muestran los resultados obtenidos para el caso de una perturbación escalón. Debido a la deficiencia del primer esquema de SCD para alcanzar el objetivo de control, se analizan las posibles causas de este mal desempeño y se propone una solución. Del análisis de los resultados presentados se deduce que es necesario modificar la estructura del SCD para obtener un control eficiente de los QP. Se describe la propuesta de modificación que da origen a un nuevo método para realizar el proceso de defuzzificación. También se plantea una estructura reducida de controlador difuso que utiliza el nuevo método de defuzzificación. Finalmente, el nuevo método de defuzzificación se prueba como parte de los dos esquemas de control difuso planteados. La validación de ambos controladores difusos (con nuevo método de defuzzificación y con nuevo método de defuzzificación y estructura reducida) se realiza mediante la simulación en la ejecución del control de los quemadores posteriores de la central termoeléctrica de ciclo combinado (CTCC). Los resultados obtenidos se comparan contra los datos existentes, correspondientes a un control tradicional tipo proporcional + integral (PI) actualmente empleado para tal fin. Finalmente, en base al análisis de la comparación anterior se determina cual de los controladores es más eficiente en la ejecución del control de los QP.

Por último, en el capítulo 5 se realiza un análisis del nuevo método de defuzzificación y se dan las conclusiones de este trabajo.

## BIBLIOGRAFÍA

- [1] McNeill, Daniel and Freiburger, Paul; *Fuzzy Logic*, Simon & Schuster, USA, 1993.
- [2] Kosko, Bart and Isaka, Satoru; *Lógica borrosa, Investigación y Ciencia*, septiembre, pp. 60-65, 1993.
- [3] Zadeh, L. A., Fuzzy sets, *Information and Control*; Vol. 8, pp. 338-353, 1965.
- [4] Zadeh, L. A., Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Trans. Syst., Man, Cybern.*, vol. 28, pp. 28-44, 1978.
- [5] Maiers, J. and Sherif, Y. S., Applications of Fuzzy Set Theory, *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-15, No. 1, pp. 175-189, 1985.
- [6] Schwartz, T. J., Fuzzy Systems Come to Life in Japan, *IEEE Expert*, February, pp. 77-78, 1990.
- [7] Mamdani, Ebrahim, and Sedrak Assilian. An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller, *International Journal of Man-Machine Studies*, 7, pp. 1-13, 1975
- [8] Holmblad, Peter and Ostergaard, Jens-Jorgen; Control of a Cement Kiln by Fuzzy Logic, en M. M. Gupta and Elie Sanchez, eds., *Fuzzy Information and Decision Processes*, Amsterdam, North-Holland, pp. 398-399, 1982.
- [9] Sugeno, Michio; *Industrial Applications of fuzzy control*, North Holland, 1985.
- [10] Bonissone, P. P. et al; Industrial Applications of Fuzzy Logic at General Electric, *Proceedings of the IEEE*, Vol. 83, No. 3, March 1995.
- [11] Munakata, T. and Jani, Y., Fuzzy Systems: An Overview, *Communications of the ACM*, March, Vol. 37, No. 3, 1994.

---

## CAPÍTULO DOS

---

### LÓGICA DIFUSA

---

#### 2.1 INTRODUCCIÓN

Los logistas tradicionales tenían la convicción de que las sentencias de los lenguajes naturales podían ser sólo ciertas o falsas, y a lo más, incorporar un tercer valor que se interpretaba como indeterminado. Los logistas modernos han pensado muy seriamente acerca de la semántica formal y otros estudios de la forma del lenguaje. Como los filósofos y psicólogos, los logistas han llegado a determinar el hecho de que los conceptos empleados en el lenguaje natural tienen sus fronteras vagas y su significado es difuso. Dado lo anterior, las sentencias de un lenguaje natural a menudo no son siempre totalmente ciertas, o totalmente falsas y además tienen un cierto sentido, ya que en algunos conceptos son ciertas y en otros falsas.

Para los logistas clásicos, el valor de verdad o falsedad de una sentencia en términos de la teoría clásica de conjuntos se define en base a la pertenencia o no del sujeto a un conjunto perfectamente definido (predicado). Si decimos Pedro es alto, esta sentencia es verdadera si Pedro pertenece al conjunto de hombres altos y falsa si no pertenece. Pero en los lenguajes naturales existe el problema de que el concepto de altura es un concepto relativo y no completamente definido, ya que el concepto de hombre alto no es lo mismo entre los jugadores de un equipo de basquetbol y los pigmeos, por ejemplo.

Considérese además algunas otras situaciones como el hecho de preguntarse acerca del valor de verdad de algunas sentencias del tipo  $X$  es un  $Y$ . Estas podrían ser verdaderas sólo en cierto grado en lugar de ser claramente ciertas o falsas, esto se puede apreciar en los siguientes ejemplos:

Sentencia	Valor de verdad
a) El petirrojo es un pájaro	cierto
b) El pollo es un pájaro	menos cierto que a)
c) El pingüino es un pájaro	menos cierta que b)
d) El murciélago es un pájaro	muy lejos de la verdad
e) La vaca es un pájaro	falso.

Estas mismas sentencias se podrían escribir dando un juicio relativo a la categoría de membresía. Es decir, asignando al miembro (elemento) en cuestión a la categoría en la cual tiene el mayor grado de pertenencia o membresía. Esto es:

Sentencia	Valor de verdad
a) El petirrojo es más pájaro que otra cosa	verdadero
b) El pollo es más pájaro que otra cosa	verdadero
c) El pingüino es más pájaro que otra cosa	verdadero
d) El murciélago es más pájaro que otra cosa	falso
e) La vaca es más pájaro que otra cosa	falso.

Como se aprecia del ejemplo anterior, para la lógica tradicional no es siempre fácil asignar los valores de verdad a las sentencias del lenguaje natural.

Si se parte de la hipótesis fundamental de que todo el conocimiento humano acerca del mundo real está basado en teorías, tomadas aquí en el sentido más general, ya que se incluyen teorías formales y pensamientos intuitivos; y del hecho de que estas teorías son descripciones, y que la descripción para darse requiere de un lenguaje y el lenguaje de la lógica; se debe desarrollar una nueva lógica que contemple las características de imprecisión del lenguaje natural y que permita, con él, describir y teorizar.

Recuérdese que en el sentido más general un lenguaje se define como un conjunto de sentencias compuestas de símbolos acordes con ciertas reglas. Por otro lado, una interpretación es una función que asigna un significado a las palabras de un lenguaje relativo a cierto sistema, y especifica las verdaderas condiciones para las sentencias.

Si se emplea un lenguaje para hacer descripciones, entonces se necesita de alguna lógica, reglas para hacer inferencias, para derivar sentencias conductivas de un conjunto de premisas. Si las reglas tienen las propiedades de preservar la verdad, dando sólo consecuencias ciertas de premisas ciertas, se tendrá una clase de sistema deductivo, en este caso difuso, que se puede identificar con una teoría. De acuerdo con esto, desarrollar una lógica acorde al lenguaje natural que se emplee para

definir los estados actuales permitirá hacer una teoría de él. Si se tiene conocimiento intuitivo o empírico de algunas sentencias que son ciertas, entonces de acuerdo a esta lógica se emplearán algunas reglas de inferencia para obtener nuevas sentencias ciertas.

La interpretación, a menudo, está en estos casos íntimamente relacionada con el lenguaje, cada palabra tiene su propio significado, en este caso, más o menos definido (difuso). Este significado se define como un nuevo concepto, el denominado conjunto difuso, que es una clase con un continuo de grados de membresía. Como se verá posteriormente, la noción de conjunto difuso provee un conveniente punto de partida para la construcción de la estructura conceptual, que es paralela en muchos aspectos a la estructura utilizada en el caso de conjuntos ordinarios, pero que es más general que esta última. Esencialmente, tal estructura (la lógica difusa) provee una forma natural para tratar con problemas en donde la fuente de imprecisión es la ausencia de criterios claramente definidos de clases de membresía (o pertenencia) en lugar de la presencia de variables aleatorias.

A continuación, después de dar un breve desarrollo histórico del tratamiento de la vaguedad, que es el origen de la lógica difusa; se realiza la descripción de los conceptos fundamentales de la lógica difusa, las operaciones básicas en la lógica difusa y se describe un concepto fundamental, la variable lingüística.

## 2.2 FILOSOFÍA DE LO DIFUSO

### 2.2.1 Desarrollo Histórico del Tratamiento de la Vaguedad

El primer pensador que trató seriamente con la vaguedad fue el filósofo estadounidense Sanders Peirce (1839-1914). Peirce aseguraba que todo cuanto existe es continuo y que, por tanto, tal continuidad gobierna el conocimiento. Por ejemplo, decía, los conceptos *tamaño* y *tiempo* son continuos. Aseguraba que la vaguedad está presente en todas partes aún en las palabras. *"Las palabras no dejan de pertenecer súbitamente a puntos en el espectro, mas bien se extienden sobre él"*, sostenía. *"Este tipo de incertidumbre siempre nos afligirá"*, afirmaba, *"la vaguedad no se puede descartar del mundo de la lógica así como la fricción de la mecánica"* [1].

Más tarde Bertrand Russell (1872-1970) también incursionó en este tópico. En 1923 Russell publicó un artículo en el que sostenía que tanto la vaguedad como la imprecisión eran características del lenguaje no de la realidad. *"Todo lenguaje es vago"*, afirmaba; *"por ejemplo, el término metro es claramente un concepto vago, aunque aparentemente sea un término exacto"*. Un metro era una distancia entre dos marcas en una barra de metal que se encontraba en París (metro patrón). Sin embargo, decía, *"esas marcas no son puntos sin dimensión, más bien son señales de algún marcador, así que la distancia entre ellas no es clara. La barra también se puede expandir y contraer con el calor y el frío. Entonces el concepto metro es vago"*, concluía. Russell dio otro ejemplo con el concepto *segundo*: *"un segundo también es un concepto difuso. Un segundo es una fracción de una rotación de la tierra, pero la tierra no es un cuerpo rígido, no todas sus partes giran a la misma velocidad y alguna medida de ésta esta sujeta a error"*. ¿Que sucede con los conceptos cierto y falso? Sin símbolos precisos, éstos también son vagos. De acuerdo a lo anterior, cada proposición tiene un rango no definido de hechos que la harán cierta. *"Este hombre"* puede aplicarse plenamente a un hombre de 40 años, pero, ¿es cierto cuando se aplica a un adolescente o a un Neanderthal?

Russell fue aún más lejos al afirmar *"la vaguedad claramente es materia de gradación"*. Lo anterior se puede deducir claramente de los siguientes ejemplos: un mapa a escala pequeña es más vago que otro a una escala más grande. Una foto borrosa de Juan es más vaga que una foto clara de él. Una persona que vemos a 100 metros es más vaga que una que vemos a 2 metros. Una estrella que vemos con nuestros ojos es más vaga que si la vemos a través de un telescopio.

Durante los siguientes 40 años otros pensadores reconocieron la existencia de la vaguedad. Albert Einstein, por ejemplo, afirmó: *"En la medida en que las leyes de las matemáticas se refieren a la realidad, éstas no son ciertas; y en la medida en que éstas son ciertas, ellas no se refieren a la realidad"* [1]. Sin embargo, el filósofo polaco Jan Lukasiewicz (1878-1955) fue el primero en dar los pasos para formalizar un modelo de la vaguedad. Lukasiewicz inventó la subestructura de los conjuntos difusos, una lógica basada en más valores que cierto y falso. En 1920 publicó un artículo describiendo su nueva creación. En este artículo, se estableció el valor de 1 para verdadero y 0 para falso. En adición se estableció 1/2 para posible. Así pues, una sentencia podía tener cualquiera de esos tres valores. Aparentemente lo anterior fue un paso muy sencillo, sin embargo, de un análisis cuidadoso se obtienen resultados muy interesantes. Lukasiewicz incluso aseveró: *"la lógica cambia desde sus fundamentos si asumimos que, en adición a verdadero y falso, existe algún tercer valor lógico o varios de tales valores"* [1].

La lógica tradicional trata con sentencias como:

*Es verdad que mañana lloverá*

cuya negación es:

*Es verdad que mañana no lloverá.*

Lukasiewicz agregó otro tipo de sentencia:

*Es posible que mañana llueva.*

Esta sentencia tiene un valor de verdad de  $1/2$ . Su negación es:

*Es posible que mañana no llueva.*

Esta sentencia también tiene un valor de verdad de  $1/2$ . Es obvio que  $1/2 = 1/2$ , es decir *SENTENCIA = NO SENTENCIA*. La aseveración y su opuesto son equivalentes.

Para la mayoría de la gente lo anterior no causa conflictos. Si una casa está medio construida, también esta medio no construida. Si es posible que mañana llueva, también es posible que mañana no llueva. Aquí existen contradicciones parciales, las sentencias son parcialmente verdaderas y parcialmente falsas, tales sentencias fuerzan a su opuesto. Si una copa esta medio llena, también esa misma copa esta medio vacía.

En la lógica tradicional la operación de negación define el opuesto. Esta operación convierte una sentencia en su antítesis. En la lógica de cierto (1) y falso (0), una sentencia cierta se convierte en falsa. Representado en una tabla se tiene:

sentencia	negación
1	0
0	1

En la lógica de tres valores de Lukasiewicz sólo hay que agregar una línea a la tabla:

sentencia	negación
1	0
$1/2$	$1/2$
0	1

Como se observa, los valores de la lógica binaria permanecen intactos en las esquinas. Lukasiewicz observó que su lógica también funcionaba bien para la lógica clásica, pues contenía los valores 0 y 1. Así la lógica de tres valores no desplazaba a la lógica tradicional, más bien la *extendía*.

Una vez observado lo anterior, Lukasiewicz no encontró razones para insertar algún valor extra. Sin embargo, pudo agregar cuatro, siete, quince, tantos como hubiera querido. De hecho, pudo agregar un número infinito de valores, todos entre 0 y 1, cada uno indicaría un cierto grado. En otras palabras, cierto y falso podían llenar el vasto rango interno entre 0 y 1. Tal espectro excedía, penso, la lógica de tres valores o cualquier otra lógica [1].

Posteriormente Max Black (1909-1989) fue el hombre que estableció el camino a seguir en el estudio de la vaguedad. Black planteó sus ideas en su artículo "*Vagueness: An Exercise in Logical Analysis*" publicado en 1937. En ese artículo Black manifestó estar de acuerdo con Peirce en el sentido de que la vaguedad representaba un continuo. También aceptó estar de acuerdo con Russell en que la vaguedad es objeto de gradación. De hecho, afirmó que lo continuo implicaba grados. "*Lo continuo no necesariamente es continuo, puede ser discreto como una línea punteada y si los intervalos son lo suficientemente pequeños, entonces nadie notará su discontinuidad*". Por ejemplo, aludía, "*la forma en que se habla el Alemán de villa en villa a través del país (Alemania) va cambiando lentamente, de tal manera que en villas vecinas este cambio es imperceptible. Sin embargo, estos cambios acumulados hacen que personas en los dos extremos del país prácticamente no se puedan entender. Algo similar sucede con los objetos, por ejemplo el concepto "silla". Si se ponen en una línea una serie de objetos cuya apariencia vaya dejando de ser la de una silla en grados pequeños, entonces, esto será imperceptible para alguien que observe dos objetos contiguos. Dado lo anterior, se puede tomar un determinado número*

para cada objeto. Tal número indicará un grado y el objeto "silla" será la colección de esos grados" [1]. ¿Pero grados de que? Black pudo elegir grados de verdad, sin embargo, eligió grados de "uso". El número de cada uno de esos objetos en la fila mostraba el porcentaje en que la gente lo llamaría "silla". Una silla es una silla si la mayoría de la sociedad la nombra silla. Bajo este razonamiento la vaguedad fue considerada por Black como materia de probabilidad. Por ejemplo, si el 67% de una población nombra a un objeto silla, entonces este objeto es una silla en un 67%. Así mismo, la probabilidad de que la siguiente persona que vea el objeto le asigne el nombre de silla es del 67%.

Black sugirió que los términos vagos podían formar conjuntos e incluso dibujó las primeras curvas de lo que posteriormente serían los conjuntos difusos. Él pensó que con estos conjuntos se podrían hacer operaciones, así como permitirían la existencia de subconjuntos [1]. También observó que un conjunto podía ser un subconjunto parcial de otro. De tal manera, Black había desarrollado una nueva estructura intelectual, pero nunca junto las piezas.

En 1965 Lotfi Zadeh publicó su artículo "Fuzzy Sets" [3] en donde junta las partes que Black había dejado dispersas. A diferencia de Lukasiewicz y Black, Zadeh manipuló de manera correcta la vaguedad y estableció la estructura de la lógica difusa. Esta estructura fusionó la lógica de Lukasiewicz con la lógica clásica.

## 2. 2. 2 Lógica Difusa y Conjuntos Difusos : Una Forma de Manipular la Vaguedad e Imprecisión

Si la lógica tradicional se define como la ciencia que trata de los principios formales del razonamiento, entonces la lógica difusa se puede definir como la ciencia que trata de los principios formales del razonamiento aproximado, en donde el razonamiento preciso es un caso extremo y particular [2]. En otras palabras, la lógica difusa es una extensión de la lógica clásica.

En gran medida la lógica difusa fue motivada por la necesidad de encontrar una estructura conceptual para manipular la vaguedad e imprecisión inherentes a la representación que las personas hacen del mundo real y que expresan con el lenguaje natural. Tiene su origen en la teoría de los conjuntos difusos que comenzó a desarrollar L. Zadeh en 1965 [3] como "un sistema que proporciona una vía natural para tratar los problemas en los que la fuente de imprecisión es la ausencia de criterios claramente definidos de tipos de pertenencia". En la actualidad, la lógica difusa constituye una metodología muy adecuada para el tratamiento del conocimiento ordinario (conocimiento del sentido común): para su organización, representación y manipulación.

La lógica difusa descansa en la afirmación de que un concepto generalmente tiene fronteras no bien definidas o vagas. En la lógica tradicional estos conceptos se fuerzan a tener límites bruscos, por supuesto, las personas no utilizan estos límites al expresarse de dichos conceptos. Con su propuesta, Zadeh ofreció un esquema conceptual más apropiado para representar el conocimiento expresado con el lenguaje que el proporcionado por la lógica clásica. Para ello se basó en la premisa de que las percepciones humanas envuelven, en su mayor parte, conjuntos difusos, esto es, clases de objetos en los que la transición de la pertenencia a la no pertenencia es gradual más bien que abrupta. Por ejemplo "muy atractiva", "extremadamente inteligente", "bastante aceptable", "más o menos acertados", "casi verdad", etc. Tales conjuntos vienen determinados (definidos), no como los conjuntos en el sentido clásico: por una definición extensional o intencional (la cual garantiza, por igual, la pertenencia de sus elementos), sino por referencia a un contexto, por un procedimiento *semántico* más bien que *sintáctico*; quedan determinados por referencias a dominios específicos (locales) [4]. De tal forma, la lógica difusa refleja con mayor aproximación la manera en cómo la gente piensa, para ello modela parcialmente lo que decimos con palabras, la forma en que tomamos decisiones, la forma en que reconocemos signos y sonidos. Se piensa que en un futuro también podrá copiar funciones cerebrales, tales como detectar el color y distinguir fenómenos [5].

Lo difuso de una propiedad cae en la ausencia de límites bien definidos del conjunto de objetos a quienes esta propiedad se aplica. De manera más específica, sea  $U$  el campo de referencia, también llamado universo del discurso, que cubre un rango definido de objetos. Ahora considérese un subconjunto  $F$  de  $U$ , donde la transición entre ser miembro y no miembro es gradual en lugar de abrupta. Este subconjunto difuso  $F$  obviamente no tiene límites bien definidos. Por ejemplo,  $F$  podría ser el conjunto de *hombres altos* en una comunidad  $U$ , es decir, el conjunto de hombres que tienen la propiedad de ser altos. Usualmente, existen miembros de  $U$  que son definitivamente altos, otros que no son del todo altos, pero también existen casos que están en el límite. Entonces, un grado de membresía de 1 se asigna a los objetos que pertenecen completamente a  $F$ , en este caso, los hombres que definitivamente son altos. De manera inversa, los objetos que definitivamente no pertenecen a  $F$  tienen asignado un grado de membresía de 0. Obviamente los casos que caen en la frontera (límite) tendrán un grado de membresía entre 0 y 1. En otras palabras, entre más elemento u objeto sea característico de  $F$ , más cerca al valor de 1 estará su grado de membresía. Entonces, el uso de una escala numérica, como el intervalo  $[0,1]$ , permite una representación conveniente del *grado de membresía* (o *grado de pertenencia*) [6].

Sin embargo, es necesario aclarar tres situaciones. Primero, no existen grados de membresía precisos por sí mismos, ellos sólo son índices de tendencias que son subjetivamente asignados por un individuo o por un grupo de ellos. Así, desde un punto de vista psicológico, el grado de membresía no es un objeto primitivo. Éste solamente refleja un ordenamiento sobre los objetos del universo del discurso, inducido por la propiedad asociada con  $F$ , en este caso, *alto*. Este ordenamiento, cuando existe, es más importante que los grados de membresía por sí mismos. Segundo, los grados de membresía no están absolutamente definidos, en la mayoría de los casos son dependientes del contexto. Por ejemplo, Daniel con una altura de solamente 175 cm puede ser clasificado como *hombre alto* por un Mexicano, pero un Suizo difícilmente lo consideraría como un *hombre alto*. Tercero, lo difuso difiere de lo impreciso en que este último concepto se refiere a una ausencia de conocimiento acerca del valor de un parámetro, ejemplo altura, y entonces es expresado en un intervalo de tolerancia nítido o claro. El intervalo es el conjunto de todos los posibles valores de un parámetro. Lo difuso ocurre cuando el intervalo no tiene límites claros o bien definidos, ejemplo el conjunto difuso  $F$ .

De acuerdo con lo anterior, la teoría de los conjuntos difusos y sus ulteriores desarrollos, la lógica difusa y la teoría de la posibilidad, constituyen modelos que resultan especialmente útiles para tratar con la vaguedad y la incertidumbre de manera más "natural" y más "humana" que la lógica y la teoría de conjuntos clásicos. Los sistemas extraídos de la lógica clásica presentan las dificultades de la rigidez y la bivalencia, y resultan, por ello, inservibles para expresar la ambigüedad del significado que se da en el lenguaje natural, base fundamental de los procesos cognoscitivos y de la interacción hombre-máquina.

Ciertamente existen ya otras metodologías, extraídas de la teoría de la probabilidad, para tratar con la incertidumbre: inferencia bayesiana, probabilidad subjetiva, teoría de la evidencia; pero hay cierto tipo de incertidumbre de la que no puede dar cuenta la teoría de la probabilidad. En estas metodologías la incertidumbre reside, entonces, en la aleatoriedad (azar) de los sucesos, aun cuando éstos sean precisos y las proposiciones correspondientes sean inambiguamente verdaderas o falsas. Por ejemplo: "el próximo mes nevará", "lanza el dado y saca un as", etc. Este tipo de incertidumbre constituye fundamentalmente el campo de la teoría de la probabilidad. Pero la incertidumbre puede ser causada por la ambigüedad o por la vaguedad, inherentes, ambas, en el significado de la mayoría de los términos utilizados en el lenguaje ordinario.

Por otro lado, hay que aclarar que a pesar de que en su utilización común la ambigüedad y la vaguedad tienen un mismo significado, esto no es así. Una palabra o frase se dice que es ambigua cuando ésta involucra un conjunto discreto de posibles significados, lo cual produce incertidumbre acerca de cuál es el apropiado en una determinada instancia de uso (contexto). Mientras que una palabra o frase se dice que es vaga si su significado se refiere a un espectro continuo de interpretaciones, bien a causa de la ausencia de límites precisos, como en el caso de los antónimos (*frio-caliente*) y otros conceptos: *valiente*, *alto*, *atractiva*, etc., o bien a causa de una multiplicidad de criterios de uso, lo que conduce a un conjunto de significados que se solapan; por ejemplo, los términos *temperatura* y *color*.

Los predicados ambiguos o vagos inducen conjuntos difusos en el sentido de Zadeh. Los predicados (cuantificadores, calificadores) vagos son intrínsecos en los lenguajes naturales; por tanto, necesariamente producen incertidumbre en los enunciados de los que son constituyentes. Además, este tipo de incertidumbre es de naturaleza no probabilística: no depende del azar; no deviene clarificada con el paso del tiempo o con la testificación. Antes bien, reside en el significado de las palabras; es, pues, inherente en el lenguaje, y, dado que éste es inseparable del pensamiento humano, siempre tendrá lugar en mayor o menor medida y ocupará buena parte en los procesos cognoscitivos. El cálculo de probabilidades sirve para determinar en qué medida cabe esperar (esperanza matemática) que suceda, o no, algo concreto. "Los sistemas difusos permiten medir el grado en que algo está sucediendo ya".

En lo que se refiere a las aplicaciones que involucran el conocimiento del sentido común (o conocimiento ordinario), los modelos difusos resultan más realistas, más humanos y más útiles que los modelos lógicos o matemáticos clásicos, por cuanto que expresan las experiencias humanas, el sentido común, el significado del lenguaje natural, la verdad de las proposiciones, etc., como un sistema continuo, en lugar de hacerlo en un sistema de dos valores [4].

En lo que se refiere al mundo de la ciencia, durante mucho tiempo los científicos han enfrentado el problema de tratar con sistemas complejos. Los sistemas simples como el *péndulo* pueden modelarse con ecuaciones y los sistemas enormemente desorganizados como las moléculas de un gas en una jarra pueden describirse con métodos estadísticos. Sin embargo, las ecuaciones y la estadística fallan al querer modelar sistemas biológicos o humanísticos, es decir, sistemas centrados en el humano o donde éste tiene una gran injerencia. Zadeh sustenta lo anterior al decir "es la tremenda complejidad de los organismos vivos; complejidad cuyo orden de magnitud es más grande que aquel encontrado en el sistema inanimado más complejo hecho u observado por el hombre" [7]. Lo anterior es respaldado por su famoso Principio de incompatibilidad: "En la medida en que crece la complejidad de un sistema, en esa misma medida disminuye nuestra

capacidad para hacer precisos y aún significativos enunciados acerca de su comportamiento, hasta alcanzar un umbral, más allá del cual la precisión y la significancia (o relevancia) resultan, casi siempre, características mutuamente excluyentes”.

El cerebro es un sistema complejo. El lenguaje es un sistema complejo, el esparcimiento de una enfermedad es un sistema complejo, la ley es un sistema complejo. La complejidad domina la vida (el mundo real), las ciencias del comportamiento, la sociedad, el medio ambiente así como la medicina y aún mucho de la tecnología. Tales sistemas frecuentemente tienen un comportamiento impredecible.

Desde el punto de vista de la Ingeniería, muchos sistemas aceptan una descripción exacta en razón de su naturaleza. Razonable precisión existe en los sistemas lineales, los sistemas estables y en los sistemas no variantes en el tiempo. Sin embargo otros carecen de ella. Estos incluyen, por ejemplo, los sistemas descentralizados y los sistemas lentamente variables en el tiempo. Así, en el área de la ingeniería, la aplicación más exitosa de la lógica difusa se ha dado en los denominados sistemas de control difuso (SCD).

Así, antes de describir los SCD, es necesario establecer una notación y definir los conceptos fundamentales de la lógica difusa. Lo cual se realiza a continuación.

## 2.3 NOTACIÓN, TERMINOLOGÍA Y OPERACIONES BÁSICAS CON CONJUNTOS DIFUSOS

### 2.3.1 Conjuntos Difusos y Terminología

Sea  $U$  una colección de objetos denotados genéricamente por  $\{u\}$ , que pueden ser discretos o continuos. A  $U$  se le denomina *universo del discurso* y  $u$  representa el elemento genérico de  $U$ . Por ejemplo, un universo del discurso  $U$ , puede ser el conjunto de todos los números reales; el conjunto de enteros  $0, 1, 2, \dots, 100$ ; el conjunto de todos los residentes de una ciudad; el conjunto de objetos en un cuarto; etc. Los universos del discurso generalmente se denotan con los símbolos  $U, V, W, X, Y$  y  $Z$ , con o sin subíndices o superíndices. Un conjunto difuso en  $U$  generalmente se denota con las letras mayúsculas  $A, B, C, D, E, F, G, H$ , con o sin subíndices o superíndices.

#### Definición 2.1: Conjunto Difuso.

Un conjunto difuso  $F$  en el universo del discurso  $U$  está caracterizado por una función de membresía  $\mu_F: U \rightarrow [0, 1]$  que asocia a cada elemento de  $u$  en  $U$  un número  $\mu_F(u)$  en el intervalo  $[0, 1]$ . Un conjunto difuso se puede ver como una generalización o extensión del concepto de conjunto dado por la teoría clásica de conjuntos. Recuérdese que en esta última la definición de la función de membresía sólo puede tomar dos valores  $\{0, 1\}$ .

El conjunto difuso  $F$  en  $U$  se puede representar como el conjunto de pares ordenados del elemento genérico  $u$  y su respectivo grado de membresía o grado de pertenencia dado por  $\mu_F$ , esto es:

$$F = \{(u, \mu_F(u)) / u \in U\} \quad (2.1)$$

Cuando  $U$  es continuo, el conjunto difuso  $F$  puede representarse de manera consistente como:

$$F = \int \mu_F(u) / u \quad (2.2)$$

en donde la integral denota la operación de unión.

Cuando  $U$  es discreto, el conjunto difuso  $F$  se puede representar como:

$$F = \sum_{i=1}^n \mu_F(u_i) / u_i \quad (2.3)$$

en donde  $\mu_F(u_i)$  es el grado de membresía de  $u_i$  en  $F$  y el símbolo  $\Sigma$  representa unión en lugar de suma aritmética.

**Definición 2. 2: Soporte, punto de cruce y singleton.**

El soporte  $Supp(F)$  del conjunto difuso  $F$  es el conjunto nítido (no difuso) de todos los puntos  $u$  de  $U$  tales que  $\mu_F(u) > 0$ , esto es:

$$Supp(F) = \{u \in U / \mu_F(u) > 0\} \quad (2.4).$$

En particular, el elemento  $u$  en  $U$  para el que  $\mu_F(u) = 0.5$  se denomina *punto de cruce*.

Un conjunto difuso  $F$  cuyo soporte es un único punto en  $U$  con  $\mu_F(u) = 1.0$  se denomina *singleton*.

El conjunto difuso  $F$  se dice que es *compacto* si su soporte es finito.

En la literatura de la teoría de los conjuntos y sistemas difusos se han reportado varios métodos para describir un conjunto difuso, a continuación se darán dos ejemplos y para mayor detalle referirse a las referencias [8] y [9] en donde se muestran diversas formas de representar un conjunto difuso.

**Ejemplo 2. 1:** Sea el universo del discurso el intervalo  $[0, 100]$  con  $u$  interpretado como el concepto *edad*. Un conjunto difuso en  $U$  llamado *viejo* se puede definir por medio de una función de membresía tal como:

$$\mu_F(u) \begin{cases} 0 & \text{para } 0 \leq u \leq 50 \\ \left(1 + \left(\frac{u-50}{5}\right)^{-2}\right)^{-1} & \text{para } 50 \leq u \leq 100 \end{cases} \quad (2.5)$$

y el conjunto difuso *viejo* se puede representar como:

$$viejo = \int_{50}^{100} \left(1 + \left(\frac{u-50}{5}\right)^{-2}\right)^{-1} / u \quad (2.6).$$

En este caso, el soporte del conjunto difuso *viejo* es el intervalo  $[50, 100]$ ,  $Supp(viejo) = [50, 100]$ ; y el punto de cruce es  $u = 55$ . La representación gráfica del conjunto difuso *viejo* se puede apreciar en la figura 2.1.

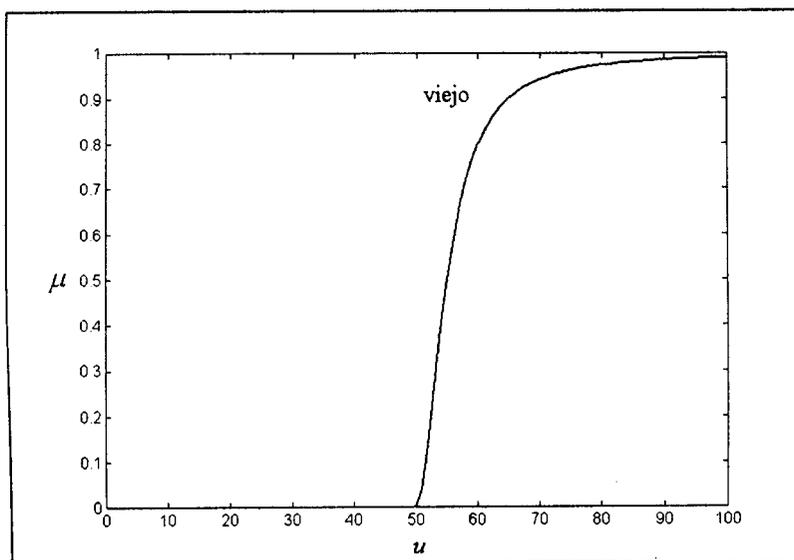


Figura 2. 1 Representación gráfica del conjunto difuso *viejo*.

**Ejemplo 2. 2:** Considérese el universo del discurso  $U = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$  y el conjunto difuso *entero cercano a cero*. Este conjunto se puede representar como:

$$entero\ cercano\ a\ cero = 1.0/0 + 0.9/1 + 0.75/2 + 0.5/3 + 0.2/4 + 0.1/5.$$

De la misma forma, el conjunto difuso *varios* podría definirse como:

$$\text{varios} = 0.5/3 + 0.8/4 + 1.0/5 + 1.0/6 + 0.8/7 + 0.5/8.$$

En muchas aplicaciones el grado de membresía  $\mu_F(u)$  puede ser interpretado como un grado de "compatibilidad" de  $u$  con el concepto representado por  $F$ . Por ejemplo, en el caso del conjunto difuso *viejo* definido por la expresión (2.6), el grado en que el valor numérico de 60 años es compatible con el concepto "viejo" es de  $\mu_{\text{viejo}}(60)=0.8$ . En otros casos  $\mu_F(u)$  se puede interpretar como el grado de la posibilidad de  $u$  dado  $F$ . Cuando  $\mu_F(u)$  es considerado como un grado de compatibilidad o posibilidad, la función  $\mu_F: U \rightarrow [0, 1]$  puede referirse como función de compatibilidad o función de posibilidad, respectivamente [10].

Es importante hacer notar que el significado asignado a un valor numérico particular de la función de membresía es de naturaleza subjetiva y dependiente del contexto de definición. Por otra parte, una vez definido el conjunto difuso, éste no tiene nada de difuso; es decir, a través de una función continua o discreta, el conjunto queda *bien definido*. Entonces el término "difuso" puede verse justamente como el nombre seleccionado para distinguir la clase de conjuntos donde  $\mu_F(u)$  puede tomar todo el rango de valores entre 0 y 1 inclusive.

**Definición 2. 3: Conjunto corte- $\alpha$ .**

Conjunto *corte- $\alpha$*  es el conjunto nítido de los elementos que pertenecen al conjunto difuso  $F$  en al menos el grado  $\alpha$ , esto es:

$$F_\alpha = \{u \in F / \mu_F(u) \geq \alpha\} \quad (2.7).$$

El conjunto *corte- $\alpha$*  firme se define como:

$$F_\alpha = \{u \in F / \mu_F(u) > \alpha\} \quad (2.8).$$

**Definición 2. 4: Ancho de un conjunto difuso.**

El *ancho* de un conjunto difuso convexo  $F$  con soporte  $S(F)$  esta definido por:

$$\text{width}(F) = \sup(S(F)) - \inf(S(F)) \quad (2.9)$$

donde sup e inf denotan las operaciones matemáticas del supremo y el infimo, respectivamente. Estas operaciones se definen como sigue:

$$\begin{aligned} \alpha = \sup(F) \text{ si y sólo si } \forall u \in F : u \leq \alpha \text{ y } \forall \varepsilon > 0 \exists u \in F : u > \alpha - \varepsilon \\ \beta = \inf(F) \text{ si y sólo si } \forall u \in F : u \geq \beta \text{ y } \forall \varepsilon > 0 \exists u \in F : u < \beta + \varepsilon \end{aligned} \quad (2.10).$$

Si el soporte de  $F$  es acotado, entonces las operaciones sup e inf se pueden reemplazar por el máximo y el mínimo respectivamente.

**Definición 2. 5: Núcleo.**

El núcleo de un conjunto difuso  $F$  es el conjunto nítido que contiene todos los valores con grado de membresía igual a 1, esto es:

$$\text{núcleo}(F) = \{u \in U / \mu_F(u) = 1\} \quad (2.11).$$

Si solamente existe en  $U$  un punto con grado de membresía igual a uno, entonces a este punto se le denomina *valor pico* de  $F$ .

**Definición 2.6: Altura de un conjunto difuso.**

La *altura* de un conjunto difuso  $F$ , denotada por  $\text{hgt}(F)$ , esta dada por el grado de membresía supremo sobre toda  $u$  en  $U$ , esto es:

$$hgt(F) = \sup_{u \in U} \mu_F(u) \quad (2.12).$$

Es decir, la altura es el máximo grado de membresía dado por  $\mu_F$ .

Un conjunto difuso  $F$  se dice que es *normal* si  $hgt(F)=1$  y es *subnormal* si  $hgt(F)<1$ .

**Definición 2.7: Conjunto difuso convexo.**

Se dice que un conjunto difuso es *convexo* si su función de membresía no tiene depresiones. De manera formal, un conjunto difuso es convexo si y sólo si:

$$\forall u_1, u_2 \in U, \forall \lambda \in [0,1]: \mu_F(\lambda \cdot u_1 + (1-\lambda) \cdot u_2) \geq \min(\mu_F(u_1), \mu_F(u_2)) \quad (2.13).$$

Lo anterior significa que la función de membresía es, por ejemplo, creciente, decreciente, o en forma de campana de gauss. También es fácil ver que un conjunto difuso  $F$  es convexo si y sólo si sus cortes- $\alpha$  son convexos en el sentido de las matemáticas clásicas. En la figura 2.2 se muestran dos conjuntos difusos, uno convexo y el otro no.

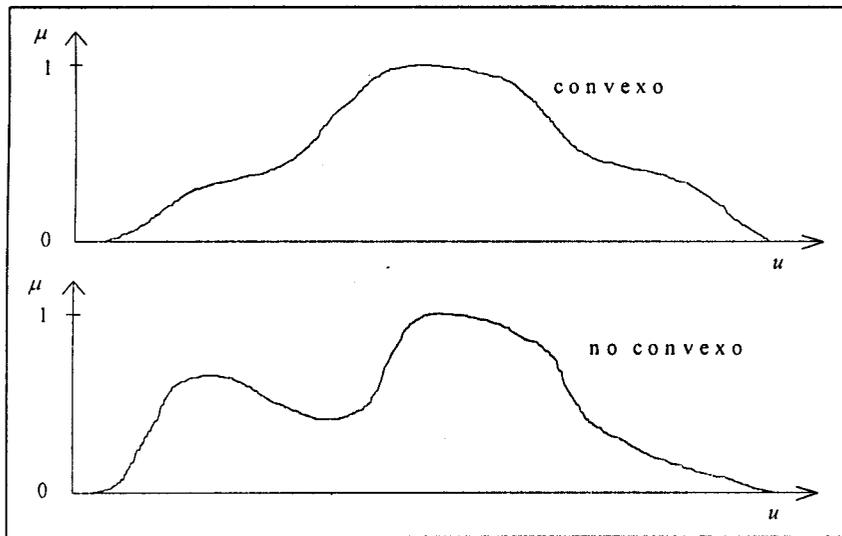


Figura 2.2 Ejemplo de un conjunto difuso convexo y otro no convexo.

**Definición 2.8: Número difuso.**

Un *número difuso*  $F$  en el universo de discurso continuo  $U$ , es un conjunto difuso que es normal y convexo.

**Definición 2.9: Cardinalidad de un conjunto difuso.**

La *cardinalidad* de un conjunto difuso  $F$  esta definida como:

$$|F| = \sum_{u \in F} \mu_F(u) \quad (2.14)$$

y cardinalidad relativa se define como:

$$\|F\| = \frac{|F|}{|U|} \quad (2.15)$$

donde  $|U|$  es la cardinalidad del universo del discurso  $U$ .

**2.3.2 Conjunto Teórico de Operaciones Básicas con Conjuntos Difusos**

Debido a que los conjuntos difusos se denotan en términos de sus funciones de membresía, las operaciones con conjuntos difusos también se definen vía sus funciones de membresía. Sean  $A$  y  $B$  dos conjuntos difusos en el universo del discurso  $U$ , con funciones de membresía  $\mu_A$  y  $\mu_B$  respectivamente.

**Definición 2. 10: Conjunto vacío.**

Un conjunto difuso está vacío si y sólo si  $\mu_A(U) = 0$ ; y es universal si y sólo si  $\mu_A(U) = 1 \forall u \in U$ .

Las nociones de igualdad y contención de conjuntos difusos se derivan inmediatamente de la teoría clásica de conjuntos.

**Definición 2. 11: Igualdad de conjuntos.**

Dos conjuntos difusos  $A$  y  $B$  son iguales ( $A = B$ ) si y sólo si  $\forall u \in U: \mu_A(u) = \mu_B(u)$ .

**Definición 2. 12: Subconjunto.**

$A$  es subconjunto difuso  $B$  ( $A \subseteq B$ ) si y sólo si  $\forall u \in U: \mu_A(u) \leq \mu_B(u)$ .

En la teoría clásica de conjuntos la unión, intersección y complemento son operaciones simples definidas sin ambigüedad. Esta claridad de definición viene dada por el hecho de que las operaciones lógicas *AND*, *OR* y *NOT* tienen una semántica bien definida basada en una lógica de proposiciones. Por ejemplo, " $\phi$  *AND*  $\psi$ ", en la lógica de proposiciones es verdadera si y sólo si *ambas* expresiones  $\phi$  y  $\psi$  son verdaderas. En la teoría de los conjuntos difusos la definición de estas operaciones no es tan simple, esto se debe a la utilización de grados de pertenencia, Zadeh propuso [10] las siguientes definiciones para la intersección, unión y complemento, respectivamente:

$$\forall u \in U: \mu_{A \cap B}(u) = \min(\mu_A(u), \mu_B(u)), \quad (2.16)$$

$$\forall u \in U: \mu_{A \cup B}(u) = \max(\mu_A(u), \mu_B(u)), \quad (2.17)$$

$$\forall u \in U: \mu_{\bar{A}}(u) = 1 - \mu_A(u). \quad (2.18)$$

Las definiciones anteriores son una extensión muy simple de las operaciones clásicas con conjuntos. Sin embargo, se pueden realizar otras extensiones. De manera general, normas triangulares (T-norma y T-conorma o S-norma) se utilizan para representar las operaciones de intersección, unión y complemento. Una norma triangular o T-norma se puede considerar como el operador de intersección más general.

**Definición 2. 13: T-norma.**

Una *norma triangular* o *T-norma* " $t$ " es una función binaria en el intervalo unitario  $[0, 1]$  que puede representar la operación de intersección si satisface las siguientes condiciones:

$$T-1: a t b = b t a;$$

$$T-2: (a t b) t c = a t (b t c);$$

$$T-3: a \leq c \text{ y } b \leq d \text{ implica que } a t b \leq c t d;$$

$$T-4: a t 1 = a;$$

con  $a, b, c, d \in [0, 1]$ .

De T-3 se deduce que para cualquier  $a, 0 t a \leq 0 t 1$ ; de T-4 se deduce que  $0 t a = 0$ . Las T-normas más importantes son:

$$\text{Mínimo: } M(a, b) = \min(a, b) \quad (2.19)$$

$$\text{Producto algebraico: } \prod(a, b) = a * b \quad (2.20)$$

$$\text{Lukasiewictz: } W(a, b) = \max(0, a + b - 1) \quad (2.21)$$

$$\text{Degenerada: } Z(a, b) = \begin{cases} a & \text{si } b = 1 \\ b & \text{si } a = 1 \\ 0 & \text{en otro caso} \end{cases} \quad (2.22).$$

Para cualquier T-norma se verifica que  $Z \leq t \leq M$ . En particular, para las T-normas citadas, se cumple que  $Z \leq W \leq \prod \leq M$ . Las T-normas más utilizadas son el mínimo (min) y el producto.

Una conorma triangular o S-norma se puede considerar como el operador de unión más general.

**Definición 2. 14: T-conorma o S-norma.**

Una *conorma triangular* o *S-norma* " $s$ " es una función binaria en el intervalo unitario  $[0, 1]$  que puede representar la

operación de unión si satisface las siguientes condiciones:

- S-1:  $a s b = b s a$ ;
- S-2:  $(a s b) s c = a s (b s c)$ ;
- S-3:  $a \leq c$  y  $b \leq d$  implica que  $a s b \leq c s d$ ;
- S-4:  $a s 0 = a$ ;

con  $a, b, c, d \in [0, 1]$ .

De las condiciones S-3 y S-4 se deduce que para cualquier  $a$ ,  $1 s a = 1$ . Algunas de las S-normas más importantes son:

$$\text{Máximo: } M^*(a, b) = \max(a, b) \quad (2.23)$$

$$\text{Suma probabilística: } \Pi^*(a, b) = a + b - a * b \quad (2.24)$$

$$\text{Lukasiewictz: } W^*(a, b) = \min(a + b, 1) \quad (2.25)$$

$$\text{Degenerada: } Z^*(a, b) = \begin{cases} a & \text{si } b = 0 \\ b & \text{si } a = 0 \\ 1 & \text{en otro caso} \end{cases} \quad (2.26).$$

Para cualquier S-norma se cumple que  $M^* \leq s \leq Z^*$ . Las S-normas más utilizadas son el máximo y la suma probabilística.

La relación general entre las T y las S normas esta dada por un equivalente de la ley de De-Morgan, esto es:

$$a t b = 1 - ((1 - a) s (1 - b)) \quad (2.27).$$

Si una T y S normas satisfacen esta propiedad, entonces una es la conjugada de la otra. Ejemplo, el mínimo y el máximo.

El complemento se denota por  $c(a)$ , donde  $c$  es la operación complemento más general.

**Definición 2.15: Complemento.**

La operación *complemento* "c" debe al menos satisfacer las siguientes condiciones:

- C-1:  $c(0) = 1$ ;
- C-2:  $a < b$  implica  $c(a) > c(b)$ ;
- C-3:  $c(c(a)) = a$ .

De aquí,  $c(1) = c(c) = 0$ . El operador complemento más empleado es "menos uno", esto es:

$$\mu_{\bar{A}}(u) = 1 - \mu_A(u) \quad (2.28)$$

con  $u \in U$ . Esta operación corresponde a la operación lógica NOT.

Algunas operaciones que son frecuentemente utilizadas en la teoría del control difuso son las siguientes:

**Definición 2.16: Suma acotada.**

La *suma acotada* de A y B ( $A \oplus B$ ) esta definida por:

$$A \oplus B = \int_U \min(1, \mu_A(u) + \mu_B(u)) / u \quad (2.29)$$

donde + es la suma aritmética.

**Definición 17: Resta acotada.**

La *resta acotada* de A y B ( $A \ominus B$ ) esta definida por:

$$A \ominus B = \int_U \max(0, \mu_A(u) - \mu_B(u)) / u \quad (2.30)$$

donde  $-$  es la resta aritmética.

**Definición 2. 18: Producto algebraico.**

El producto algebraico de dos conjuntos difusos  $A$  y  $B$  ( $A * B$ ) está definido por:

$$A * B = \int_U \mu_A(u) * \mu_B(u) / u \quad (2.31).$$

**Definición 2. 19: Involución.**

La involución o potencia  $\alpha$  del conjunto difuso  $A$ , donde  $\alpha$  es un número real, está definida por:

$$A^\alpha = \int_U (\mu_A(u))^\alpha / u \quad (2.32).$$

Las operaciones de *concentración* ( $CON$ ) y *dilatación* ( $DIL$ ) son casos especiales de  $A^\alpha$ .

**Definición 2. 20: Concentración.**

La *concentración* del conjunto difuso  $A$  ( $CON(A)$ ) está definida por:

$$CON(A) = A^2 \quad (2.33).$$

**Definición 2. 21: Dilatación.**

La *dilatación* del conjunto difuso  $A$  ( $DIL(A)$ ) está definida por:

$$DIL(A) = A^{0.5} \quad (2.34).$$

## 2.4 RELACIONES DIFUSAS

En los sistemas difusos las relaciones o mapeos se realizan entre variables difusas definidas en diferentes universos del discurso. Estas relaciones se efectúan a través de sentencias condicionales difusas o implicaciones lingüísticas de la forma:

$$A \Rightarrow B \text{ ó "Si } A(u) \text{ Entonces } B(v)" \quad (2.35).$$

Antes de definir una relación difusa es necesario definir el producto cartesiano de conjuntos difusos. El producto cartesiano de los conjuntos difusos  $A$  y  $B$  denotado por  $A \times B$  se define por:

$$A \times B = \int_{U \times V} \mu_A(u) \text{ t } \mu_B(v) / (u, v) \quad (2.36)$$

donde  $U \times V$  es la colección de pares ordenados  $(u, v)$ , tales que  $U \times V = \{(u, v) / u \in U, v \in V\}$ . De manera general se tiene la siguiente definición:

**Definición 2. 22: Producto Cartesiano.**

Si  $A_1, \dots, A_n$  son conjuntos difusos definidos en los universos del discurso  $U_1, \dots, U_n$ , respectivamente, entonces el producto cartesiano de  $A_1, \dots, A_n$  es un conjunto difuso en el espacio producto  $U_1 \times U_2 \times \dots \times U_n$  con las funciones de membresía: con  $t = \min$

$$\mu_{A_1 \times A_2 \times \dots \times A_n}(u_1, u_2, \dots, u_n) = \min\{\mu_{A_1}(u_1), \mu_{A_2}(u_2), \dots, \mu_{A_n}(u_n)\} \quad (2.37);$$

ó con  $t = \text{producto algebraico}$

$$\mu_{A_1 \times A_2 \times \dots \times A_n}(u_1, u_2, \dots, u_n) = \{\mu_{A_1}(u_1) * \mu_{A_2}(u_2) * \dots * \mu_{A_n}(u_n)\} \quad (2.38).$$

Ahora se tienen las condiciones para definir una relación difusa. Una relación difusa de la forma  $A \Rightarrow B$ , denotada por  $R$ , del conjunto  $A \subset U$  al conjunto  $B \subset V$  se define como un subconjunto difuso del producto cartesiano  $U \times V$ .  $R$  se caracteriza por la función de membresía de 2 variables  $\mu_R(u, v)$  dada por:

$$\begin{aligned} R = A \times B &= \int_{U \times V} \mu_R(u, v)/(u, v) \\ &= \int_{U \times V} \mu_A(u) \text{ t } \mu_B(v)/(u, v) \end{aligned} \quad (2.39).$$

Utilizando las T-normas del mínimo y del producto,  $R$  se define como:

$$\mu_R(u, v) = \mu_{A \times B}(u, v) = \begin{cases} \min(\mu_A(u), \mu_B(v)) & \text{mínimo} \\ \mu_A(u) * \mu_B(v) & \text{producto} \end{cases} \quad (2.40),$$

para toda  $u \in U$  y  $v \in V$ .

La función de membresía  $\mu_R(u, v)$  corresponde a la fuerza de conexión o correlación entre  $u$  y  $v$  en el mapeo o implicación.

Si los universos del discurso son discretos, entonces la ecuación (2.39) se transforma en:

$$A \times B = \sum_{u, v} \mu_A(u) \text{ t } \mu_B(v)/(u, v) \quad (2.41),$$

para cada operador intersección  $\text{t}$  y  $U = \{u_1, u_2, \dots, u_n\}$ ;  $V = \{v_1, v_2, \dots, v_n\}$ . De manera general:

**Definición 2. 23: Relación Difusa R.**

Una  $n$ -aria relación difusa  $R$ , es un subconjunto difuso en  $U_1 \times U_2 \times \dots \times U_n$  dado por:

$$R_{U_1 \times \dots \times U_n} = \int_{U_1 \times \dots \times U_n} \mu_R(u_1, \dots, u_n)/(u_1, \dots, u_n) \in U_1 \times \dots \times U_n \quad (2.42).$$

Dados los conjuntos difusos finitos expresados de manera discreta como  $A = \{a_1, a_2, \dots, a_m\}$  y  $B = \{b_1, b_2, \dots, b_n\}$ , una relación difusa  $R = A \times B$  se puede expresar con la matriz  $m \times n$ :

$$M_R = \begin{bmatrix} \mu_R(a_1, b_1) & \mu_R(a_1, b_2) & \dots & \mu_R(a_1, b_n) \\ \mu_R(a_2, b_1) & \mu_R(a_2, b_2) & \dots & \mu_R(a_2, b_n) \\ \vdots & \vdots & \vdots & \vdots \\ \mu_R(a_m, b_1) & \mu_R(a_m, b_2) & \dots & \mu_R(a_m, b_n) \end{bmatrix} \quad (2.43).$$

A la matriz anterior se le denomina matriz difusa. Los elementos de la matriz difusa son valores que están dentro del intervalo  $[0, 1]$ , dado que  $\mu_R$  tiene valores dentro de éste rango.

**Ejemplo 2. 3:** Sea el universo del discurso discreto  $U = \{0, 20, 40, 60, 80, 100\}$  y los conjuntos difusos etiquetados como *rápido* y *lento*:

$$\begin{aligned} A = \text{rápido} &= 0/0 + 0/20 + 0.3/40 + 0.7/60 + 1/80 + 1/100 \\ B = \text{lento} &= 1/0 + 0.7/20 + 0.3/40 + 0/60 + 0/80 + 0/100, \end{aligned}$$

considérese la siguiente sentencia condicional difusa: "Si  $B$  es *lento* entonces  $A$  es *rápido*". La matriz difusa de la relación  $B \times A$  es de la forma dada por (2.43). Si se utiliza la T-norma del mínimo la matriz difusa esta dada por:

$$\begin{aligned}
 B \times A &= \begin{bmatrix} \min(\mu_B(a_1), \mu_A(b_1)) & \min(\mu_B(a_1), \mu_A(b_2)) & \cdots & \min(\mu_B(a_1), \mu_A(b_6)) \\ \min(\mu_B(a_2), \mu_A(b_1)) & \min(\mu_B(a_2), \mu_A(b_2)) & \cdots & \min(\mu_B(a_2), \mu_A(b_6)) \\ \vdots & \vdots & \vdots & \vdots \\ \min(\mu_B(a_6), \mu_A(b_1)) & \min(\mu_B(a_6), \mu_A(b_2)) & \cdots & \min(\mu_B(a_6), \mu_A(b_6)) \end{bmatrix} \\
 &= \begin{bmatrix} 0 & 0 & 0.3 & 0.7 & 1 & 1 \\ 0 & 0 & 0.3 & 0.7 & 0.7 & 0.7 \\ 0 & 0 & 0.3 & 0.3 & 0.3 & 0.3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
 \end{aligned}$$

Mientras que si se utiliza el producto entonces la matriz difusa esta dada por:

$$\begin{aligned}
 B \times A &= \begin{bmatrix} \mu_B(a_1) * \mu_A(b_1) & \mu_B(a_1) * \mu_A(b_2) & \cdots & \mu_B(a_1) * \mu_A(b_6) \\ \mu_B(a_2) * \mu_A(b_1) & \mu_B(a_2) * \mu_A(b_2) & \cdots & \mu_B(a_2) * \mu_A(b_6) \\ \vdots & \vdots & \vdots & \vdots \\ \mu_B(a_6) * \mu_A(b_1) & \mu_B(a_6) * \mu_A(b_2) & \cdots & \mu_B(a_6) * \mu_A(b_6) \end{bmatrix} \\
 &= \begin{bmatrix} 0 & 0 & 0.3 & 0.7 & 1 & 1 \\ 0 & 0 & 0.21 & 0.49 & 0.7 & 0.7 \\ 0 & 0 & 0.09 & 0.21 & 0.3 & 0.3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
 \end{aligned}$$

De forma gráfica las 2 relaciones se muestran en la figura 2. 3. De aquí se puede observar que la T-norma del producto algebraico produce una superficie de relación más suave que la arrojada por el mínimo.

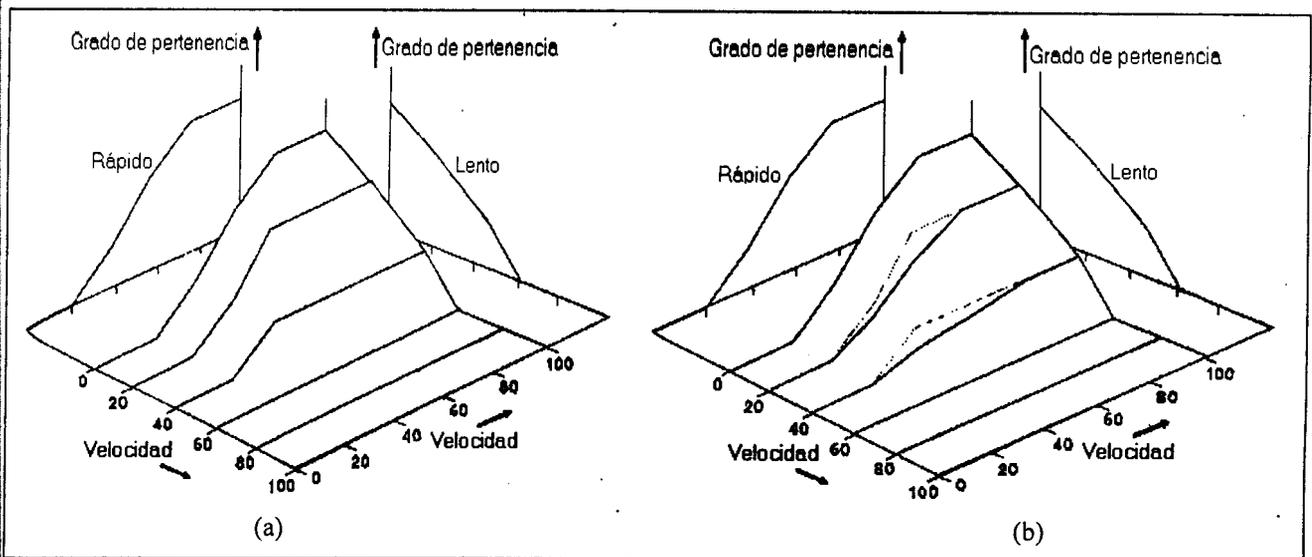


Figura 2. 3 (a) Relación  $B \times A$  utilizando la T-norma del mínimo, (b) Relación  $B \times A$  utilizando la T-norma del producto.

### 2.4.1 Operaciones con Relaciones Difusas

Las relaciones difusas son muy importantes en la teoría del control difuso, ya que con ellas se pueden describir las interacciones que se dan entre las variables de un proceso. Antes de describir como se efectúa lo anterior, primero es necesario dar algunas definiciones que se refieren a las operaciones que se pueden realizar con las relaciones difusas. Las primeras dos operaciones son la intersección y la unión:

**Definición 2. 24: Intersección de relaciones difusas.**

Sean  $R$  y  $S$  dos relaciones difusas binarias definidas en  $U \times V$ . La intersección de  $R$  y  $S$  esta definida por:

$$\forall (u, v) \in U \times V: \mu_{R \cap S}(u, v) = \min(\mu_R(u, v), \mu_S(u, v)) \quad (2.44).$$

En lugar del mínimo (min) se puede utilizar cualquier T-norma.

**Definición 2. 25: Unión de relaciones difusas.**

La unión de  $R$  y  $S$  esta definida por:

$$\forall (u, v) \in U \times V: \mu_{R \cup S}(u, v) = \max(\mu_R(u, v), \mu_S(u, v)) \quad (2.45).$$

En lugar del máximo (max) se puede utilizar cualquier T-conorma o S-norma.

Obviamente las 2 definiciones anteriores se pueden extender para el caso de relaciones n-arias.

**Ejemplo 2. 4:** Sean las relaciones binaras  $R$  y  $S$  definidas por:

$$R = \text{"}u \text{ considerablemente más largo que } v\text{"} = \begin{bmatrix} 0.8 & 1 & 0.1 & 0.7 \\ 0 & 0.8 & 0 & 0 \\ 0.9 & 1 & 0.7 & 0.8 \end{bmatrix}$$

$$S = \text{"}v \text{ muy cerca de } u\text{"} = \begin{bmatrix} 0.4 & 0 & 0.9 & 0.6 \\ 0.9 & 0.4 & 0.5 & 0.7 \\ 0.3 & 0 & 0.8 & 0.5 \end{bmatrix}$$

La intersección de las relaciones  $R$  y  $S$  denotada como "u considerablemente más largo que v y v muy cerca de u", esta dada por:

$$R \cap S = \begin{bmatrix} 0.4 & 0 & 0.1 & 0.6 \\ 0 & 0.4 & 0 & 0 \\ 0.3 & 0 & 0.7 & 0.5 \end{bmatrix}$$

La unión de las relaciones  $R$  y  $S$  denotada como "u considerablemente más largo que v o v muy cerca de u", es la relación:

$$R \cup S = \begin{bmatrix} 0.8 & 1 & 0.9 & 0.7 \\ 0.9 & 0.8 & 0.5 & 0.7 \\ 0.9 & 1 & 0.8 & 0.8 \end{bmatrix}$$

En el ejemplo anterior, para la operación de intersección se utilizó la T-norma del mínimo, mientras que para la unión se utilizó la T-conorma del máximo. Es necesario aclarar que se pudo utilizar cualquier T-norma para realizar la operación de intersección, así como cualquier T-conorma para realizar la operación de unión.

Dos operaciones con conjuntos difusos y relaciones difusas que son de enorme importancia son la *proyección* y la *extensión cilíndrica*. La operación de proyección lleva una relación ternaria a una relación binaria, o una relación binaria a un conjunto difuso, o un conjunto difuso a un sólo valor concreto.

Para definir la operación de proyección sea  $R$  una relación en  $U = \times_{i=1}^n U_i$ ; sea  $(i_1, \dots, i_k)$  una subsecuencia de  $(1, \dots, n)$  y sea  $(j_1, \dots, j_l)$  la subsecuencia complementaria de  $(1, \dots, n)$ . Sea  $V = \times_{m=1}^k U_{i_m}$ .

**Definición 2. 26: Proyección.**

La *proyección* de  $R$  en  $V$  esta definida por:

$$proj R \text{ en } V = \int_V \sup_{u_{j_1}, \dots, u_{j_l}} \mu_R(u_1, \dots, u_n) / (u_{i_1}, \dots, u_{i_k}) \quad (2.46).$$

En el caso más simple de dos variables, sea  $R$  una relación definida en  $U \times V$ , la proyección de  $R$  en  $V$  esta dada por:

$$proj R \text{ en } V = \int_V \sup_u \mu_R(u, v) / v \quad (2.47).$$

En lugar del supremo, que es necesario utilizar cuando  $U$  y  $V$  son continuos, el máximo se utiliza cuando  $U$  y  $V$  son discretos.

**Ejemplo 2. 5:** Considérese la relación dada por:

$$R = \text{"}u \text{ considerablemente más largo que } v\text{"} = \begin{bmatrix} 0.8 & 1 & 0.1 & 0.7 \\ 0 & 0.8 & 0 & 0 \\ 0.9 & 1 & 0.7 & 0.8 \end{bmatrix}.$$

La proyección de  $R$  sobre  $U$  significa que:

$u_1$  toma el valor del grado de pertenencia más grande de las tuplas  $(u_1, v_1), (u_1, v_2), (u_1, v_3), (u_1, v_4)$ ; es decir, el valor máximo de pertenencia encontrado en el primer renglón de la matriz difusa.

$u_2$  toma el valor del grado de pertenencia más grande de las tuplas  $(u_2, v_1), (u_2, v_2), (u_2, v_3), (u_2, v_4)$ ; es decir, el valor máximo de pertenencia encontrado en el segundo renglón de la matriz difusa.

$u_3$  toma el valor del grado de pertenencia más grande de las tuplas  $(u_3, v_1), (u_3, v_2), (u_3, v_3), (u_3, v_4)$ ; es decir, el valor máximo de pertenencia encontrado en el tercer renglón de la matriz difusa.

De tal forma, el conjunto difuso que se obtiene es:

$$proj R \text{ en } U = 1/u_1 + 0.8/u_2 + 1/u_3 .$$

De la misma forma , la proyección de  $R$  sobre  $V$  se realiza buscando el grado de pertenencia que sea máximo en cada una de las cuatro columnas. Lo anterior da como resultado el siguiente conjunto difuso:

$$proj R \text{ en } V = 0.9/v_1 + 0.8/v_2 + 0.7/v_3 + 0.8/v_4 .$$

También se puede realizar la proyección total de esta relación. En este caso el valor de membresía o pertenencia máximo es igual a 1. Lo anterior también se puede ilustrar con una tabla como sigue:

	$v_1$	$v_2$	$v_3$	$v_4$	Proj en $U$
$u_1$	0.8	1	0.1	0.7	1
$u_2$	0	0.8	0	0	0.8
$u_3$	0.9	1	0.7	0.8	1
Proj en $V$	0.9	1	0.7	0.8	Proyección total 1

La operación de proyección casi siempre se utiliza en combinación con la *extensión cilíndrica*. La extensión cilíndrica es más o menos la operación opuesta de la proyección. La extensión cilíndrica extiende un conjunto difuso a una relación binaria, una relación binaria a una relación ternaria y así sucesivamente. La extensión cilíndrica tiene la siguiente función: sea  $A$  un conjunto difuso definido en  $U$  y sea  $R$  una relación difusa definida en  $U \times V$ , obviamente no es posible realizar la intersección de  $A$  y  $R$ , pero si se extiende  $A$  a  $U \times V$ , entonces lo anterior se hace posible. Sea  $S$  una relación en  $V = \times_{m=1}^k U_{im}$ , como se mencionó en la definición de proyección, también sea  $U = \times_{i=1}^n U_i$  de la misma definición. De aquí la definición de la extensión cilíndrica es:

**Definición 2. 27: Extensión cilíndrica.**

La *extensión cilíndrica* de  $S$  a  $U$  se define como:

$$ce(S) = \int \mu_S(u_1, \dots, u_k) / (u_1, \dots, u_n) \quad (2.48).$$

En el caso de dos variables, sea  $F$  el conjunto difuso definido en  $V$ , la extensión cilíndrica de  $F$  en  $U \times V$  es el conjunto de todas las tuplas  $(u, v) \in U \times V$  con grado de membresía o pertenencia igual a  $\mu_F(v)$ , esto es:

$$ce(F) = \int_{U \times V} \mu_F(v) / (u, v) \quad (2.49).$$

De aquí se deduce que  $proj\ ce(s)$  en  $V = S$ , pero, en general,  $ce(proj R$  en  $V) \neq R$ .

**Ejemplo 2. 6:** Considérese el siguiente conjunto difuso:

$$Pr\ o\ j\ de\ R\ en\ U = 1/u_1 + 0.8/u_2 + 1/u_3$$

que se obtuvo en el ejemplo anterior. La extensión cilíndrica de este conjunto en el dominio  $U \times V$  esta dada por:

$$ce(A) =$$

	$v_1$	$v_2$	$v_3$	$v_4$
$u_1$	1	1	1	1
$u_2$	0.8	0.8	0.8	0.8
$u_3$	1	1	1	1

Si ahora se considera el conjunto difuso resultante de proyectar  $R$  en  $V$ , como se obtuvo en el ejemplo anterior:

$$B = Pr\ o\ j\ de\ R\ en\ V = 0.9/v_1 + 1/v_2 + 0.7/v_3 + 0.8/v_4$$

entonces la extensión cilíndrica de  $B$  en  $U \times V$  es:

$$ce(A) =$$

	$v_1$	$v_2$	$v_3$	$v_4$
$u_1$	0.9	1	0.7	0.8
$u_2$	0.9	1	0.7	0.8
$u_3$	0.9	1	0.7	0.8

La combinación de los conjuntos difusos y las relaciones difusas, con la ayuda de las operaciones de extensión cilíndrica y la proyección, dan origen a otra operación denominada *composición*. La composición denotada por  $\circ$  se define como:

**Definición 2. 28: Composición.**

Sea  $A$  un conjunto difuso definido en  $U$  y  $R$  una relación difusa definida en  $U \times V$ . La *composición* de  $A$  y  $R$ , que da como resultado el conjunto difuso  $B$  en  $V$ , esta dada por:

$$B = A \circ R = Pr\ o\ j\ (ce(A) \cap R) \text{ en } V \quad (2.50).$$

Si la operación de intersección se realiza con la T-norma del mínimo y en la operación de proyección se utiliza el máximo, entonces se tiene:

$$\mu_B(v) = \max_u \min(\mu_A(u), \mu_R(u, v)) \quad (2.51).$$

Esta operación es conocida como la composición *max-min*. Si la operación de intersección se realiza con la T-norma del producto algebraico y en la operación de proyección se utiliza el máximo, entonces se tiene:

$$\mu_B(v) = \max_u \mu_A(u) * \mu_R(u, v) \quad (2.52).$$

A esta operación se le conoce como composición *max-product*.

**Ejemplo 2. 7:** Considérese que se tiene el siguiente universo del discurso denominado *altura* (expresado en centímetros):

$$U = \{170.0, 172.5, 175.0, 177.5, 180.0, 182.5, 185.0\}$$

y considérese la siguiente relación expresada de manera lingüística como:

*R*: Mario es un poco más alto que Juan

en donde *un poco más alto* es una relación difusa binaria con la función de membresía  $\mu_R(x, y)$  que denota el grado en que *x* es un poco más alto que *y*. Esta función de membresía se describe por medio de la siguiente matriz:

185.0	0.0	0.1	0.4	0.7	1.0	0.7	0.4
182.5	0.1	0.4	0.7	1.0	0.7	0.4	0.1
180.0	0.4	0.7	1.0	0.7	0.4	0.1	0.0
y 177.5	0.7	1.0	0.7	0.4	0.1	0.0	0.0
175.0	1.0	0.7	0.4	0.1	0.0	0.0	0.0
172.5	0.7	0.4	0.1	0.0	0.0	0.0	0.0
170.0	0.4	0.1	0.0	0.0	0.0	0.0	0.0
	170.0	172.5	175.0	177.5	180.0	182.5	185.0
	<i>x</i>						

Considérese también el hecho de que “Mario es algo alto”, descrito por:

$$M = 0.0/170.0 + 0.1/172.5 + 0.4/175.0 + 0.7/177.5 + 0.9/180.0 + 1.0/182.5 + 1.0/185.0$$

Para determinar la estatura de Juan es necesario realizar la composición de *M* y *R*, *M* o *R*. Para realizar lo anterior, primero se efectúa la extensión cilíndrica de *A*, esto es:

185.0	0.0	0.1	0.4	0.7	0.9	1.0	1.0
182.5	0.0	0.1	0.4	0.7	0.9	1.0	1.0
180.0	0.0	0.1	0.4	0.7	0.9	1.0	1.0
y 177.5	0.0	0.1	0.4	0.7	0.9	1.0	1.0
175.0	0.0	0.1	0.4	0.7	0.9	1.0	1.0
172.5	0.0	0.1	0.4	0.7	0.9	1.0	1.0
170.0	0.0	0.1	0.4	0.7	0.9	1.0	1.0
	170.0	172.5	175.0	177.5	180.0	182.5	185.0
	<i>x</i>						

El siguiente paso es realizar la intersección de las dos matrices anteriores, para ello se empleara la T-norma del mínimo, así el resultado es una nueva matriz. Finalmente, la altura de Juan esta dada por la proyección de  $R \cap ce(M)$  en *y*. Enseguida se muestra la matriz  $R \cap ce(M)$  y allí mismo se muestra la proyección sobre *y*, esto es:

								Proj en y
185.0	0.0	0.1	0.4	0.7	0.9	0.7	0.4	0.9
182.5	0.0	0.1	0.4	0.7	0.7	0.4	0.1	0.7
180.0	0.0	0.1	0.4	0.7	0.4	0.1	0.0	0.7
y 177.5	0.0	0.1	0.4	0.4	0.1	0.0	0.0	0.4
175.0	0.0	0.1	0.4	0.1	0.0	0.0	0.0	0.4
172.5	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.1
170.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.1
	170.0	172.5	175.0	177.5	180.0	182.5	185.0	
	x							

de tal forma que el conjunto difuso  $J$  que corresponde a la estatura de Juan es:

$$J = 0.1/170.0 + 0.1/172.5 + 0.4/175.0 + 0.4/177.5 + 0.7/180.0 + 0.7/182.5 + 0.9/185.0$$

que muestra que "Juan es un poco más bajo que Mario".

Ahora supóngase que existen dos relaciones  $R$  y  $S$ , donde  $R$  esta definida en  $U \times V$  y  $S$  esta definida en  $V \times W$ . En este caso no es posible realizar la intersección de  $R$  y  $S$  porque éstas están definidas en diferentes dominios. Para realizar la intersección, primero es necesario extender ambas relaciones al dominio  $U \times V \times W$ . Una vez realizado lo anterior, entonces será posible realizar la intersección de  $R$  y  $S$ . Por último, la intersección deberá ser proyectada sobre  $U \times W$ . Formalmente, la intersección de  $R$  y  $S$  esta dada por:

$$\text{Proj de } (ce(R) \cap ce(S)) \text{ en } V \times W \quad (2.53).$$

Teóricamente, las operaciones se suceden como sigue:

$$R = \int_{U \times V} \mu_R(u_i, v_j)/(u_i, v_j) \quad S = \int_{V \times W} \mu_S(v_j, w_k)/(v_j, w_k) \quad (2.54)$$

La extensión de estas dos relaciones al dominio  $U \times V \times W$  esta dada por:

$$ce(R) = \int_{U \times V \times W} \mu_R(u_i, v_j)/(u_i, v_j, w_k) \quad ce(S) = \int_{U \times V \times W} \mu_S(v_j, w_k)/(u_i, v_j, w_k) \quad (2.55).$$

La intersección de  $ce(R)$  y  $ce(S)$  es igual a:

$$ce(R) \cap ce(S) = \int_{U \times V \times W} \min(\mu_R(u_i, v_j), \mu_S(v_j, w_k))/(u_i, v_j, w_k) \quad (2.56).$$

La proyección de esta nueva relación ternaria al espacio  $U \times W$  esta dada por:

$$\text{proj de } (ce(R) \cap ce(S)) \text{ en } U \times W = \int_{U \times W} \sup_j \min(\mu_R(u_i, v_j), \mu_S(v_j, w_k))/(u_i, w_k) \quad (2.57).$$

Esta última expresión que resulta de la combinación de las operaciones del mínimo y el supremo es extremadamente importante en la teoría de los conjuntos difusos. Ésta se denota como composición *sup-min* "o", así  $R \circ S$  esta dada por:

$$R \circ S = \int_{U \times W} \sup_j \min(\mu_R(u_i, v_j), \mu_S(v_j, w_k))/(u_i, w_k) \quad (2.58).$$

En lugar del supremo, que es necesario utilizar cuando  $U$ ,  $V$  y  $W$  son continuos, el máximo se utiliza cuando  $U$ ,  $V$  y  $W$  son discretos:

$$R \circ S = \int_{U \times W} \max_j \min(\mu_R(u_i, v_j), \mu_S(v_j, w_k)) / (u_i, w_k) \tag{2.59}$$

Esta expresión es conocida como composición *max-min*.

### 2.4.2 El Principio de Extensión.

Un concepto muy importante en la teoría de los conjuntos difusos es el *principio de extensión*. El principio de extensión provee una metodología general para combinar conceptos no difusos y conceptos difusos de todo tipo; por ejemplo, para combinar conjuntos difusos y relaciones, así como también las operaciones de las matemáticas tradicionales con conjuntos difusos. Los conjuntos difusos también se pueden interpretar como números difusos. En este caso se puede utilizar el principio de extensión para sumar o multiplicar dichos números difusos.

Sean  $A_1, \dots, A_n$  conjuntos difusos definidos en  $U_1, \dots, U_n$ , respectivamente; y sea  $f$  una función no difusa  $f: U_1 \times \dots \times U_n \rightarrow V$ . El objetivo es extender  $f$  de tal forma que opere sobre  $A_1, \dots, A_n$  para obtener como resultado un conjunto difuso  $F$  en  $V$ . Lo anterior es posible de realizar utilizando la composición sup-min como sigue:

**Definición 2.29: Extensión.**

La *extensión* de  $f$ , operando sobre  $A_1, \dots, A_n$  da como resultado la siguiente función de membresía  $F$ :

$$\mu_F(v) = \sup_{\substack{u_1, \dots, u_n \\ f(u_1, \dots, u_n) = v}} \min(\mu_{A_1}(u_1), \dots, \mu_{A_n}(u_n)) \tag{2.60}$$

cuando  $f^{-1}(v)$  existe. En cualquier otra forma  $\mu_A(v) = 0$ .

Otra forma de obtener el resultado anterior es:

$$\mu_F(v) = \int_{U_1 \times \dots \times U_n} \min(\mu_{A_1}(u_1), \dots, \mu_{A_n}(u_n)) / f(u_1, \dots, u_n) \tag{2.61}$$

En el caso de dos variables y considerando un dominio discreto o compacto, la ecuación (2.60) se transforma en:

$$\mu_{f(A_1, A_2)}(v) = \max_{\substack{u_1, u_2 \\ v = f(u_1, u_2)}} \min(\mu_{A_1}(u_1), \mu_{A_2}(u_2)) \tag{2.62}$$

De esta manera  $f$  se extiende del dominio de los números reales al dominio de los números difusos.

### 2.5 VARIABLE LINGÜÍSTICA

Cuando se describe el comportamiento de sistemas humanísticos, es decir, sistemas centrados en el humano o donde éste tiene una gran injerencia, generalmente se utilizan palabras en lugar de números para caracterizar los valores de las variables así como las relaciones existentes entre ellas. Por ejemplo, se podría caracterizar la edad, inteligencia, la habilidad para relacionarse con otras personas y la apariencia de una persona en palabras como sigue:

edad	muy joven
inteligencia	elevada
relación con otras personas	no muy amigable
apariencia	algo atractiva.

Las tareas básicas desarrolladas por los humanos no requieren de un alto grado de precisión en su ejecución. El cerebro humano toma provecho de esta tolerancia para codificar la información, relevante en la toma de decisiones o de la tarea específica en cuestión, dentro de etiquetas de conjuntos difusos que contienen una relación aproximada con respecto a los datos preliminares. Así pues, el flujo de información que llega al cerebro vía los sentidos visual, auditivo, tacto, etc., eventualmente se reduce a un flujo mínimo necesario para desarrollar una tarea específica con un grado mínimo de precisión.

Es claro que el uso de palabras en lugar de números implica un grado inferior de precisión en la caracterización de los valores de una variable. En algunos casos, se elige ser impreciso porque no existe la necesidad para tener un alto grado de precisión. Sin embargo, en la mayoría de los casos, la imprecisión es forzada a pesar de lo que se desee; esto debido a que no existen unidades de medición para los atributos de un objeto y no existe un criterio cuantitativo para representar los valores de tales atributos como puntos en una escala graduada.

Desde esta perspectiva, la función principal de las variables lingüísticas es proveer una estructura sistemática para caracterizar, de manera aproximada, los fenómenos complejos o no bien definidos. Esta caracterización se realiza mediante el uso de palabras o sentencias expresadas en un lenguaje natural o sintético. Dichas palabras o sentencias describen los valores y las interrelaciones de las variables de un fenómeno.

De lo anterior se deduce que la esencia del aprovechamiento lingüístico es que se sacrifica precisión para ganar significancia. Esto hace posible analizar de manera aproximada esos sistemas humanísticos y también mecánicos que son demasiado complejos para la aplicación de técnicas convencionales. La primera característica del aprovechamiento lingüístico y del razonamiento aproximado es el uso de las llamadas variables lingüísticas en lugar o además de variables numéricas para caracterizar los fenómenos complejos.

Ahora se está en condiciones de definir una variable lingüística.

**Definición 2. 30: Variable lingüística.**

Una variable lingüística es aquella cuyos valores no son números sino palabras o sentencias (proposiciones) expresadas en un lenguaje natural o sintético (artificial). Así, por ejemplo, si se considera la variable *Edad*, entonces ésta puede verse como una variable numérica cuyo rango de valores es el intervalo  $[0, 100]$ ; o también puede verse como una variable lingüística que puede tomar los valores *joven*, *viejo*, *no muy joven*, *no muy viejo*, por ejemplo. Cada uno de estos valores (palabras o sentencias) se puede interpretar como una etiqueta de un subconjunto difuso del universo del discurso  $U = [0, 100]$ , cuya variable base  $u$ , es el valor numérico genérico de *Edad* tal como se puede apreciar en la figura 2.4.

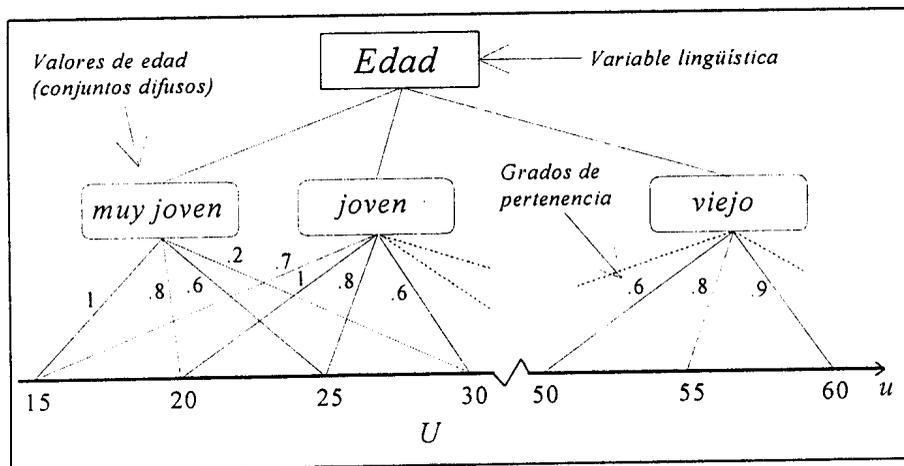


Figura 2. 4 Estructura de la variable lingüística Edad.

De manera más formal, una variable lingüística se caracteriza por la quinteta  $(x, T(x), U, G, M)$  [6] donde:

$x$  es el nombre de la variable.

$T(x)$  es el conjunto de términos o valores de  $x$ , es decir, el conjunto de nombres de los valores lingüísticos de  $x$ . Cada valor es una variable difusa denotada genéricamente por  $X$  y cuyos valores están dentro de un rango sobre el universo del discurso  $U$ .

$U$  es el universo del discurso, que está asociado con la variable base  $u$ .

$G$  es una regla sintáctica para generar los nombres de los términos de  $x$ , es decir, los valores de  $x$ .

$M$  es la regla semántica que asocia a cada valor de  $x$  su significado  $M(x)$ , donde  $M$  es un subconjunto difuso de  $U$ . Es decir, esta regla define un procedimiento algorítmico para calcular el significado de cada valor lingüístico.

Comúnmente los valores de una variable lingüística tal como *Edad* se conforman a partir de uno o más términos primarios (que son etiquetas o calificativos de conjuntos difusos primarios), junto con una colección de modificadores y conectivos que permiten generar un valor lingüístico compuesto a partir de los términos primarios. Generalmente el número

de términos primarios es dos, en donde uno es el antónimo del otro. Por ejemplo, en el caso de la variable Edad, los términos primarios son joven y viejo, con viejo siendo el antónimo de joven.

Es preciso aclarar que en el concepto de variable lingüística el significado de los términos primarios es dependiente del contexto, es decir, se asume que el significado de estos términos se especifica a priori; mientras que el significado de los modificadores y conectivos no. De tal manera, una vez que el significado de los términos primarios se ha calibrado o especificado en un contexto dado, entonces se puede calcular a través de la regla semántica  $M$  el significado de los términos compuestos tales como no muy joven, no muy viejo, no muy joven y no muy viejo etc.

En general el valor de una variable lingüística es un término compuesto  $x = x_1 x_2 \dots x_n$ , que es una concatenación de los términos individuales  $x_1 x_2 \dots x_n$ . Estos términos individuales pueden dividirse en cuatro categorías:

- 1) Términos primarios, son etiquetas de subconjuntos difusos específicos del universo del discurso.
- 2) La negación  $NO$  y los conectivos  $Y$  y  $O$ .
- 3) Bordes, tales como muy, mucho, más o menos, etc.
- 4) Marcadores, tales como paréntesis.

De tal manera que el conjunto de términos  $T(x)$  comprende los términos primarios junto con los términos generados a partir de varias combinaciones de los términos primarios.

Por ejemplo, en el caso de Edad (hablando de personas) una posible lista de sus términos, es decir los valores de la variable lingüística serían:

joven	viejo	no joven o viejo
no joven	no viejo	no muy joven y no muy viejo
muy joven	muy viejo	joven o viejo
algo joven	algo viejo	
más o menos joven	más o menos viejo	
extremadamente joven	extremadamente viejo.	

Algo que es importante señalar es que la mayoría de las variables lingüísticas tienen la misma estructura que la dada para Edad. Por ejemplo, si se reemplaza joven por alto y viejo por bajo, se podría obtener la lista de valores lingüísticos de la variable lingüística Estatura (para el caso de personas).

Una variable lingüística puede caracterizarse con una gramática atribucional que genere el conjunto de términos de la variable y que provea un procedimiento simple para calcular el significado de un valor lingüístico compuesto en términos de los conjuntos difusos primarios, es decir, los términos primarios que aparecen en su constitución. Para ilustrar lo anterior, se puede considerar la gramática atribucional mostrada abajo, en donde  $S, B, C, D$  y  $E$  no son terminales;  $NO, Y, a$  y  $b$  son terminales;  $a$  y  $b$  son los términos primarios (también los conjuntos difusos primarios); símbolos subscriptos de los conjuntos difusos que se etiquetan con el correspondiente no terminal, con  $L \underline{\Delta}$  izquierdo (perteneciente al antecedente).  $R \underline{\Delta}$  derecho (perteneciente al consecuente); y una producción de la forma:

$$S \rightarrow S Y B : S_L = S_R \cap B_R$$

significa que el conjunto difuso que es el significado del antecedente  $S$ , es la intersección de  $S_R$ , que es el significado del consecuente  $S$ , y  $B_R$ , que es el significado del consecuente  $B$ .

$$S \rightarrow B : S_L = B_R$$

$$S \rightarrow S Y B : S_L = S_R \cap B_R$$

$$B \rightarrow C : B_L = C_R$$

$$B \rightarrow NO C : B_L = C'_R (\underline{\Delta} \text{ complemento de } C_R)$$

$$C \rightarrow S : C_L = S_R$$

$$C \rightarrow D : C_L = D_R$$

$$C \rightarrow E : C_L = E_R$$

$$D \rightarrow muy D : D_L = D^2_R$$

$$D \rightarrow muy E : E_L = E^2_R$$

$$\begin{aligned} D \rightarrow a & : D_L = a \\ D \rightarrow b & : E_L = b \end{aligned}$$

Con esta gramática se podría generar la siguiente lista de valores lingüísticos:

<i>a</i>	<i>b</i>	<i>a Y b</i>
<i>NO a</i>	<i>NO b</i>	<i>NO a Y b</i>
<i>muy a</i>	<i>muy b</i>	<i>NO a Y NO b</i>
<i>NO muy a</i>	<i>NO muy b</i>	<i>NO muy a Y NO muy b</i>
<i>NO muy muy a</i>	<i>NO muy muy b</i>	

Para calcular el significado de un término compuesto de la forma  $x=hu$ , donde  $h$  es un borde y  $u$  es un término con un significado específico; ejemplo,  $x=muy\ viejo$ , donde  $h=muy$  y  $u=viejo$ . Un borde  $h$  se puede ver como un operador que transforma el conjunto difuso  $M(u)$  (que representa el significado de  $u$ ), al conjunto difuso  $M(hu)$ . Es decir, un borde sirve como una función generadora de un conjunto más grande de valores de una variable lingüística partiendo de una pequeña colección de términos primarios. En la definición del borde  $h$  como un operador se pueden emplear algunas operaciones básicas tales como la *concentración* y la *dilatación*. Así por ejemplo el borde *muy* en esencia actúa como un intensificador, generando un subconjunto del conjunto sobre el cual opera. Una simple operación que tiene esta propiedad es la *concentración (CON)*. Lo anterior sugiere que *muy x*, donde  $x$  es un término, se puede definir como el cuadrado de  $x$ , es decir:

$$muy\ x \triangleq x^2$$

más explícitamente:

$$muy\ x \triangleq \int_u \mu_x^2(y) / y.$$

Un ejemplo de lo anterior puede apreciarse en la figura 2.5 donde:

$$y = \text{viejo} \begin{cases} = 0 & \text{para } x \leq \alpha \\ = 2 \left( \frac{x - \alpha}{\gamma - \alpha} \right)^2 & \text{para } \alpha \leq x \leq \beta \\ = 1 - 2 \left( \frac{x - \alpha}{\gamma - \alpha} \right)^2 & \text{para } b \leq x \leq \gamma \\ = 1 & \text{para } x \leq \gamma \end{cases}$$

con  $\alpha=40$ ,  $\beta=55$  y  $\gamma=70$ .

entonces:

$$muy\ viejo = CON(y) = y^2$$

De tal manera que, por ejemplo, si el grado de membresía (o pertenencia) de Juan en la clase *viejo* es 0.8; entonces su grado de membresía en la clase *muy viejo* es 0.64.

Hasta aquí se han descrito los conceptos fundamentales de la lógica difusa, estos representan las bases para poder realizar cualquier aplicación. Sin lugar a dudas el área donde la lógica difusa ha tenido un mayor éxito, en cuanto a aplicaciones se refiere, es en la teoría del control automático. Por lo anterior en el siguiente capítulo se abordará este tema.

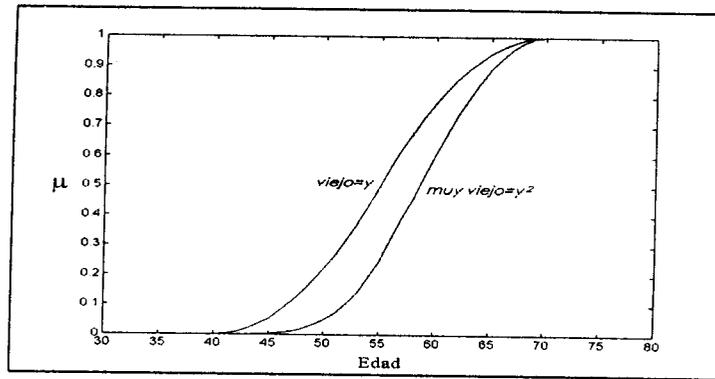


Figura 2. 5 Conjuntos difusos viejo y muy viejo.

### BIBLIOGRAFÍA

- [1] McNeill, Daniel and Freiberger, Paul; *Fuzzy Logic*, Simon & Schuster, USA, 1993.
- [2] Zadeh, L. A., *Fuzzy Logic*, *Computer*, April, 1988, pp. 83-93.
- [3] Zadeh, L. A., *Fuzzy Sets*, *Information and Control*, Vol. 8, 1965, pp. 338-353.
- [4] Velarde, L. J., *Ingeniería del conocimiento y conocimiento ordinario*, Universidad de Oviedo.
- [5] Kosko, B., *Neural Networks and Fuzzy Systems*, Prentice Hall, USA, 1992.
- [6] Driankov, D., et al; *An Introduction to Fuzzy Control*, Springer-Verlag, USA, 1993.
- [7] Zadeh, L. A., Outline of a New Approach to the Analysis of Complex Systems and Decision Processes, *IEEE Trans. Syst., Man., Cybern.*, vol. SMC-3, no. 1, Jan. 1973.
- [8] Harris, C. J., Moore, C. G. and Brown, M., *Intelligent Control Aspects of Fuzzy Logic and Neural Nets*, World Scientific, England, 1995.
- [9] Jyh-Shing, Roger Jang and Chuen-Tsai, Sun., Neuro-Fuzzy Modeling and Control, *Proceedings of the IEEE*, Vol. 83, No. 3, March, 1995, pp 378-406.
- [10] Zadeh, L. A., *Theory of Fuzzy Sets*, Memorandum No. UCB/ERL M77-1, Electronic Research, Laboratory, College of Engineering, University of California, Berkeley, USA, 4 January 1977.

---

## CAPÍTULO TRES

---

# SISTEMAS DE CONTROL DIFUSO

---

### 3.1 INTRODUCCIÓN

Un **sistema de control de procesos** se define como un arreglo de unidades integradas entre sí de manera sistemática y racional. Estas unidades se encargan de monitorear de manera continua las variables que representan los objetivos de control y las compara con el valor deseado (punto de ajuste) para determinar, de esta manera, la desviación. Para corregir la desviación el sistema produce una señal de control que, mediante un actuador final, la reduce a cero o a un valor deseado [1].

En un principio el control se realizaba de forma manual con la constante intervención de un operador humano; éste utilizaba sus sentidos para medir la evolución del proceso y efectuar acciones que redujeran las desviaciones que se presentaban respecto al comportamiento deseado, acciones que generalmente obedecían a su intuición y experiencia. Es evidente que esta manera de controlar, que en algunos casos daba resultados aceptables para lo que se pretendía, dejaba mucho que desear, tomando en cuenta que cualquier desviación en las condiciones de operación de un proceso puede producir grandes pérdidas, tanto por exceso de consumo de materias primas o energía como por variaciones en la calidad del producto obtenido. Por otra parte las exigencias del desarrollo industrial, en donde muchos de los procesos se encuentran encadenados e interactúan entre sí, obligó a acelerar la evolución del control de procesos, viéndolo no sólo como el gobierno de un proceso individual sino como todo un conjunto de ellos que tienen que ser coordinados y vistos de manera global.

Para poder evolucionar, el control de procesos necesitó del desarrollo de sensores y de las técnicas de medida y de tratamiento de señales. En un inicio tales sensores eran mecánicos, posteriormente se consiguió convertir las señales mecánicas en eléctricas de tipo analógico y por último se ha llegado a convertirlas en señales digitales lográndose, con ello, tratar numéricamente sus valores.

Respecto al equipo de control, los primeros dispositivos fueron también mecánicos con reguladores neumáticos; con ellos se realizaban las tres acciones típicas del control: proporcional, integral y derivativa, dando lugar al control PID, de uso universal y que aún hoy en día sigue dando buenos resultados en los problemas de control. Con el auge de la electrónica se introdujeron reguladores electrónicos que contaron con una mayor precisión y en los que se podía incluir una serie de operaciones para tratamiento de la señal.

Al aparecer las computadoras estas fueron consideradas como sistemas de vigilancia en donde se adquirían, almacenaban y trataban datos; para producir señales de alarma, para realizar el cálculo de variables que no se podían medir y para elaborar informes de la evolución del proceso. Cuando se observó que la computadora era capaz de responder correctamente se le encargó la tarea central de control prescindiendo y substituyendo, incluso, a los reguladores analógicos.

La información referente a las condiciones de operación del proceso entra en la computadora de manera digital, proveniente de sensores que generan señales discretas, o señales analógicas que son convertidas a digitales. Esta información es comparada con los valores deseados para producir las órdenes de control, calculadas según el algoritmo correspondiente a la ley de control que se desee utilizar, estas órdenes son enviadas a los actuadores correspondientes para efectuar la regulación del proceso. Como se observa, la computadora se convirtió en el corazón y cerebro del sistema de control al ejecutar la ley de control, que ya no se restringe al clásico PID como sucedía en los reguladores analógicos, sino que ahora se puede llevar a cabo cualquier ley de control que pueda programarse.

Por otro lado, con la aparición y difusión del microprocesador los controladores han dejado a un lado la electrónica analógica para emplear la digital y la informática. El uso de un microprocesador como núcleo de un controlador hace que todas las operaciones, cálculos y algoritmos que antes se realizaban por hardware de forma rígida ahora se lleven a cabo por rutinas, mucho más flexibles, y permite la inclusión de algoritmos de control mucho más complejos y otros tratamientos de señal que anteriormente resultaba imposible de realizar.

La evolución anterior permitió desarrollar sistemas de control que no se restringieran más a las formas convencionales (PI, PID, etc.) en donde el modelo del proceso y las leyes de control se expresan en términos de ecuaciones diferenciales. Así surgió la idea de construir un controlador capaz de captar la experiencia de operarios y expertos; y que además funcionara de manera automática. Este controlador debía ser capaz de transformar de alguna manera el conocimiento heurístico expresado en sentencias lingüísticas a un modelo matemático que fuera capaz de producir resultados aceptables. La herramienta para implementar este nuevo controlador fue la lógica difusa, y al controlador se le denominó sistema de control difuso (SCD).

A continuación, después de describir los componentes de un sistema de control convencional y las características fundamentales de un proceso, se da una descripción de las formas de control convencional más utilizadas. Posteriormente, se realiza una descripción detallada de los sistemas de control difuso.

## 3.2 SISTEMAS DE CONTROL CONVENCIONALES

### 3.2.1 Componentes de un Sistema de Control Convencional

Los componentes de un sistema de control convencional se muestran en la figura 3.1.

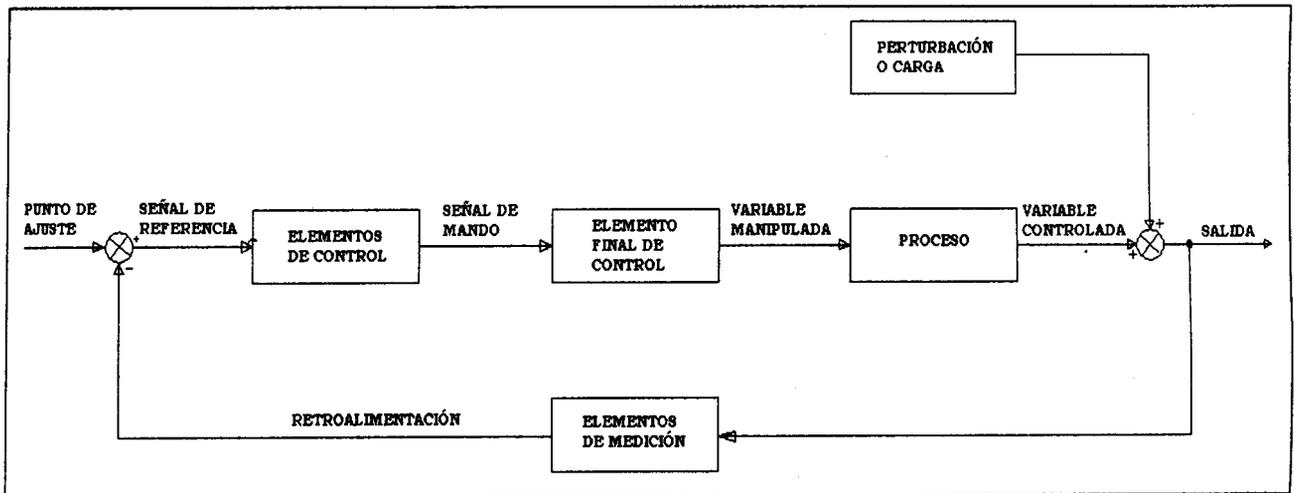


Figura 3.1 Componentes básicos de un sistema de control convencional.

Los diferentes componentes de un sistema de control, pueden definirse de la siguiente manera [2]:

**Proceso:** En la ingeniería de control, proceso se define como cualquier operación o serie de ellas que producen un resultado final deseado. Estas operaciones pueden ser mecánicas, eléctricas, reacciones químicas o cualquier combinación de ellas.

**Elementos de medición:** Es el conjunto de dispositivos de medición que tienen como finalidad el determinar y transmitir el valor de la variable controlada.

**Elementos de control:** Son los dispositivos mediante los cuales se puede efectuar una acción correctiva.

**Elemento final de control:** Es cualquier actuador o unidad de potencia que directamente modifica el valor de la variable manipulada.

**Perturbación:** Es cualquier señal que provoque cambios en las condiciones de operación del proceso.

### 3.2.2 Características Fundamentales de un Proceso

Todo proceso posee ciertas características fundamentales de retardo y demora en los cambios de sus variables, que dificultan la acción de control. Estos efectos de retardo y demora se denominan **retrasos del tiempo del proceso** y son provocados por las siguientes **propiedades del proceso** [2]:

**Capacidad:** Se refiere a la cantidad máxima de energía o material que puede ser almacenado entre los límites de una parte determinada del equipo, medida en unidades de cantidad.

**Capacitancia:** Es una medida de las características propias del proceso para mantener o transferir una cantidad de energía o de material con relación a una cantidad unitaria de alguna variable de referencia. En un proceso, una capacitancia grande es favorable para mantener constante la variable controlada a pesar de los cambios de carga que puedan presentarse. Sin embargo, esta misma propiedad hace que sea más difícil cambiar el valor de la variable controlada a uno nuevo, e introduce un retardo importante entre una variación del agente de control y el nuevo valor que toma la variable controlada.

**Resistencia:** Es la oposición parcial o total que el proceso presenta al flujo de energía o de material. Se mide en unidades de potencia requeridas para producir una unidad de cambio de flujo.

**Tiempo de transporte o tiempo muerto:** Es cualquier demora definida entre dos acciones relacionadas, durante el cual el controlador no ejecuta ninguna acción. Se mide en unidades de tiempo.

**Autorregulación:** Es una reacción sostenida, inherente a un proceso, que ayuda o se opone al establecimiento del equilibrio.

Además de tomar en cuenta las propiedades anteriores, cuando se realiza la automatización de un proceso también se debe tomar en cuenta los **cambios de carga** en el mismo. La carga del proceso es la cantidad total del fluido o agente de control que el proceso requiere en cualquier momento para mantener unas condiciones de trabajo equilibrado. Los cambios de carga, necesarios para mantener el equilibrio del proceso, son generados por cambios en la variable controlada debido a alteraciones en las condiciones de operación. Las siguientes son las causas más comunes de cambios de carga:

- 1) Mayor o menor demanda del fluido de control por el medio controlado.
- 2) Variaciones en la calidad del fluido de control.
- 3) Cambios en las condiciones ambientales.

### 3.2.3 Formas de Control Convencionales

La forma de control es la manera en que el sistema de control hace correcciones, relativas a una desviación. Visto de manera práctica, se considera que la forma de control de un controlador es la manera en que éste cambia su salida con relación a una señal de error. Se debe aclarar que la forma de control que se aplica a un proceso es el resultante de las características de operación de todos los elementos funcionales que componen el sistema de control.

Existen varias formas de control convencionales que pueden ser aplicadas individualmente o en distintas combinaciones, ello dependiendo de las características y desventajas propias de cada una, así como del proceso a controlar. Por lo general, mientras más difícil sea controlar un proceso, más complicada será la forma de control que mejor se adapte al control del proceso en cuestión.

Las siguientes son las formas de control convencionales más ampliamente utilizadas:

- a) Control de dos posiciones (todo-nada).
- b) Control flotante.
- c) Control proporcional.
- d) Control integral (de reposición).
- e) Control derivativo.
- f) Control proporcional más integral (PI).
- g) Control proporcional más derivativo (PD).
- h) Control proporcional más integral más derivativo (PID).

En los siguientes apartados se expondrá brevemente el funcionamiento de cada una de las formas de control citadas.

#### 3.2.3. a Control de Dos Posiciones (Todo-Nada)

En el control de dos posiciones el elemento final de control es llevado rápidamente desde una de dos posiciones fijas, a la otra, tomando la variable controlada, un único valor. A esta forma de control también se le conoce como dentro-fuera (on-off), ya que en las dos posiciones el elemento final de control generalmente esta abierto o cerrado, respectivamente.

Usualmente, este tipo de control se emplea con una banda diferencial o zona neutra en la que el elemento final de control permanece en su última posición para valores de la variable comprendidos dentro de la banda diferencial. De esta manera, los ajustes de control se realizan variando el punto de ajuste y la gama diferencial.

Este tipo de control funciona adecuadamente si el proceso tiene una velocidad de reacción lenta y posee un tiempo de retardo mínimo. Sin embargo, es evidente que con esta clase de control no pueden alcanzarse condiciones estables de energía de entrada y salida, y la variable controlada oscilará continuamente hacia arriba o hacia abajo del punto de ajuste. En la figura 3.2 se puede apreciar la curva de respuesta representativa de este tipo de control aplicado a un intercambiador de calor.

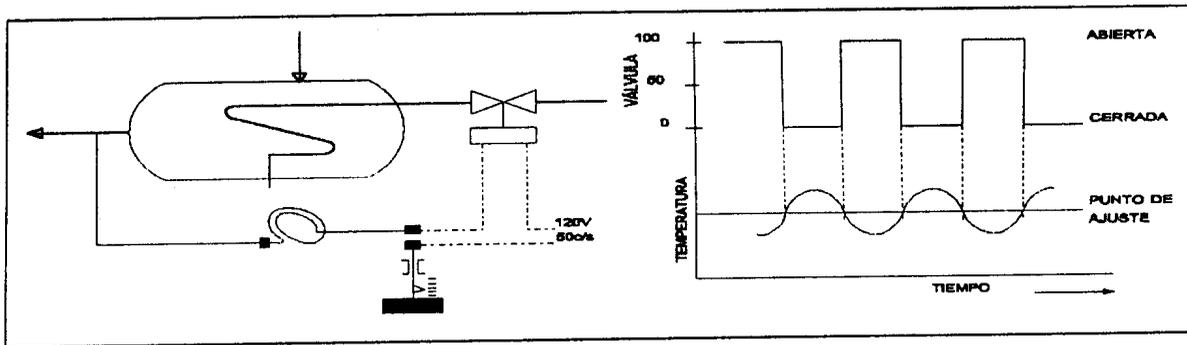


Figura 3. 2 Control de dos posiciones.

### 3. 2. 3. b Control Flotante

Una variación del control de dos posiciones es el control flotante, en el cual el elemento final de control se mueve más lentamente de una posición a la otra. En este tipo de control, existe una relación predeterminada entre la desviación y la razón del desplazamiento del elemento final de control. También puede existir una zona neutra en la cual el elemento final de control permanece en su última posición hasta que la variable controlada ha cambiado en una cantidad prescrita.

En el control flotante, al igual que en el control de dos posiciones, la variable controlada tiende a oscilar, pero en este caso las oscilaciones pueden reducirse al mínimo, si se elige adecuadamente la velocidad del elemento final de control para compensar las características del proceso.

Esta forma de control es adecuada cuando los cambios de carga en el proceso son lentos, pero si existen retardos de tiempo importantes o si los cambios de carga son muy rápidos, entonces ya no lo es. En la figura 3.3 se puede apreciar la forma en que funciona este tipo de control.

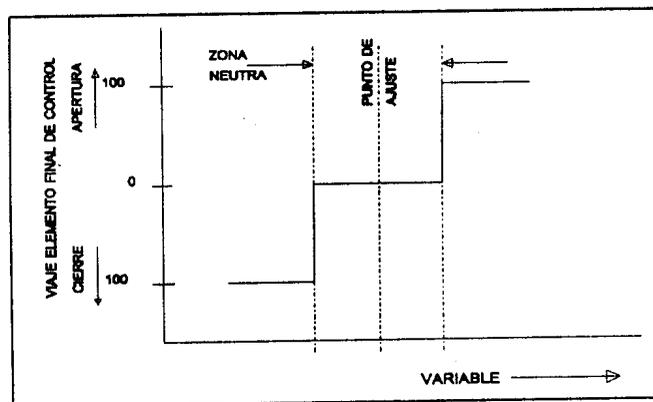


Figura 3. 3 Control flotante.

### 3. 2. 3. c Control Proporcional

En esta forma de control existe una relación lineal continua entre el valor de la variable controlada y la posición del elemento final de control, tal como se muestra en la figura 3.4. En otras palabras, el elemento final de control se mueve la misma cantidad para cada unidad de desviación; esto resulta multiplicando la señal de error por una constante proporcional, como se indica en la siguiente ecuación:

$$U_o(t) = K_p(t) U_i(t) \quad (3.1)$$

donde:  $U_o(t)$  = Señal de salida.

$U_i(t)$  = Señal de error.

$K_p$  = Ganancia proporcional.

Un término importante en este tipo de control es la denominada **banda proporcional (BP)**, la cual se define como el porcentaje de variación de la variable controlada necesario para producir una carrera completa del elemento final de control (ver figura 3.4). Generalmente, la banda proporcional se expresa en tanto por ciento de su campo de medida total. Así, por ejemplo,

si un controlador tiene una banda proporcional de 20, significa que el 20% de la señal de entrada permisible provocará un recorrido completo (sobre su rango) del elemento final de control. La fórmula de la banda proporcional es:

$$BP = \frac{100 \%}{G_c} \quad (3.2)$$

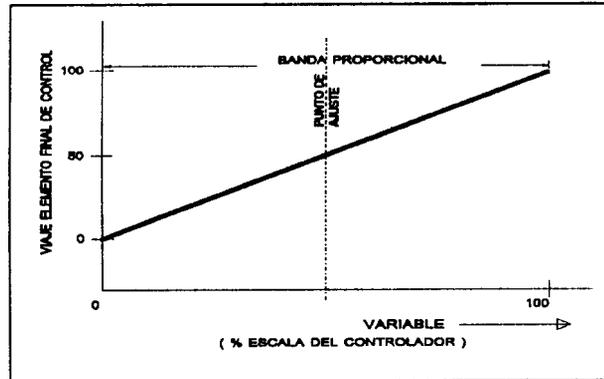


Figura 3. 4 Control proporcional.

En la ecuación anterior se introduce el concepto de *ganancia del controlador* ( $G_c$ ) que es el resultante de dividir el cambio en la salida del controlador entre el cambio a la entrada del mismo.

Este tipo de control responde únicamente a la cantidad de la desviación y no considera su duración, lo que da como resultado que siempre exista un error o desviación permanente (offset). En la figura 3.5 se gráfica la respuesta de un controlador proporcional a un cambio de carga escalonado para el ejemplo del intercambiador de calor (figura 3.2).

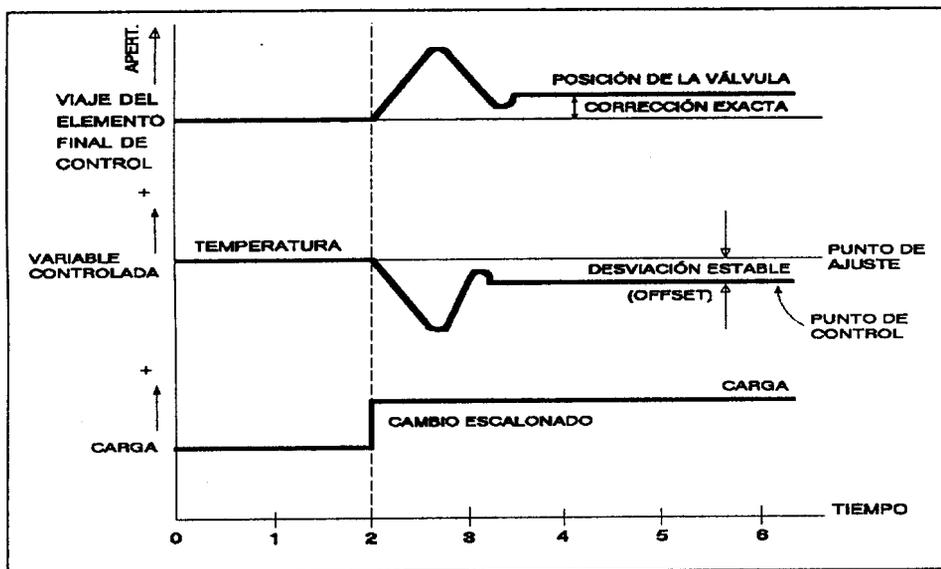


Figura 3. 5 Respuesta de un controlador proporcional a un cambio de carga escalonado.

### 3. 2. 3. d Control Integral (de Reposición)

La forma de control integral es aquella en donde el elemento final de control se posiciona de acuerdo con una función integral en el tiempo de la variable controlada. En términos matemáticos esto es:

$$U_o(t) = K_i \int_0^t U_i(t)dt + U_o(0) \quad (3.3)$$

donde:  $U_o(t)$  = Salida del controlador.

$K_i = 1/T_i$  Coeficiente integral.

$T_i$  = Constante de tiempo de acción integral.

$U_o(t)$  = Valor acumulado.

$U_i(t)$  = Señal de error.

Analizando la ecuación anterior, se deduce que cuando la variable de entrada es cero no se modifica la variable de salida, ya que en este caso se anula la integral y el valor de salida es igual al valor acumulado  $U_o(t)$ . Si la variable de entrada es pequeña, entonces la variable de salida se modifica lentamente; si esta es grande, entonces su variación es súbita. La rapidez de respuesta esta determinada por la constante de acción integral ( $T_i$ ), pues cuanto mayor sea la constante de tiempo, tanto menor será el incremento de la variable de salida por unidad de tiempo para una variable de entrada determinada.

El control integral corrige por completo la desviación permanente que se presenta en el control proporcional, ya que su salida varía de acuerdo a la cantidad y duración de la desviación; es decir, la corrección de la posición del elemento final de control continúa mientras perdure la desviación y por lo tanto se produce una corrección exacta, para cualquier cambio de carga. Sin embargo, la respuesta de esta forma de control es lenta debido a que los cambios de signo en la señal de error no son considerados hasta que se cancele el error previo, momento en el cual se lleva a cabo la inversión en la acción de corrección. En la figura 3.6 se gráfica la respuesta de un control integral a un cambio de carga escalonado.

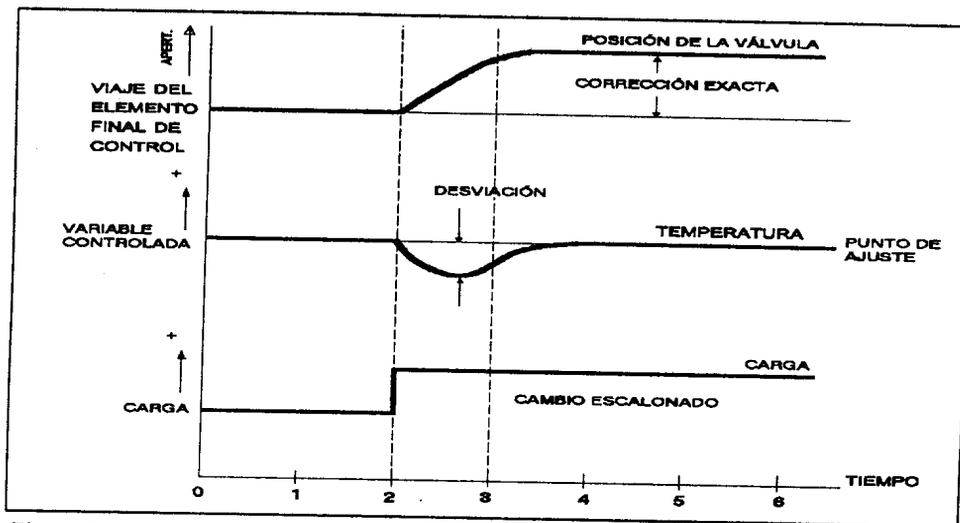


Figura 3. 6 Respuesta de un controlador integral a un cambio de carga escalonado.

### 3. 2. 3. e Control Derivativo

El control derivativo es aquel en donde la acción correctiva es igual a una constante multiplicada por la derivada del error. Su expresión matemática es la siguiente:

$$U_o(t) = T_d \frac{d}{dt} U_i(t) \quad (3.4)$$

donde:  $U_o(t)$  = Salida del controlador.

$T_d$  = Constante de tiempo derivativo en minutos.

$U_i(t)$  = Señal de error.

Este tipo de control no puede aplicarse solo, ya que un error estable o constante tendría una derivada igual a cero y por lo tanto no habría acción correctiva. El control derivativo es adecuado cuando se desea aplicar una acción correctiva larga, al tiempo que se esta formando el error, pues el valor de salida del controlador es proporcional a la velocidad de variación de la señal de error.

### 3. 2. 3. f Control Proporcional más Integral (PI)

Como se vio anteriormente, el control proporcional por sí sólo responde únicamente a la cantidad de desviación y no toma en cuenta la duración de esta, lo que da origen a una desviación permanente (offset). Para corregir este defecto, se combina el control proporcional con el integral. El primero se encarga de que el sistema de control adquiera la velocidad de respuesta necesaria para actuar de inmediato cuando aparezca el error; con el segundo, el sistema adquiere la característica de eliminar la desviación permanente, ya que la corrección de la posición del elemento final de control continuará mientras perdure el error. Por lo tanto la combinación del control proporcional con el control integral da como resultado la obtención de un control sin

error bajo todas las condiciones de carga. En la figura 3.7 puede apreciarse la respuesta de este tipo de control para un cambio de carga escalonado.

La ecuación característica de esta forma de control es la siguiente:

$$U_o(t) = K_p U_i(t) + \frac{1}{T_i} \int_0^t U_i(t) dt + U_o(0) \quad (3.5)$$

donde:  $U_o(t)$  = Salida del controlador.

$K_p$  = Ganancia proporcional.

$T_i$  = Constante de tiempo de acción integral.

$U_o(0)$  = Valor alcanzado por la salida antes de presentarse el error.

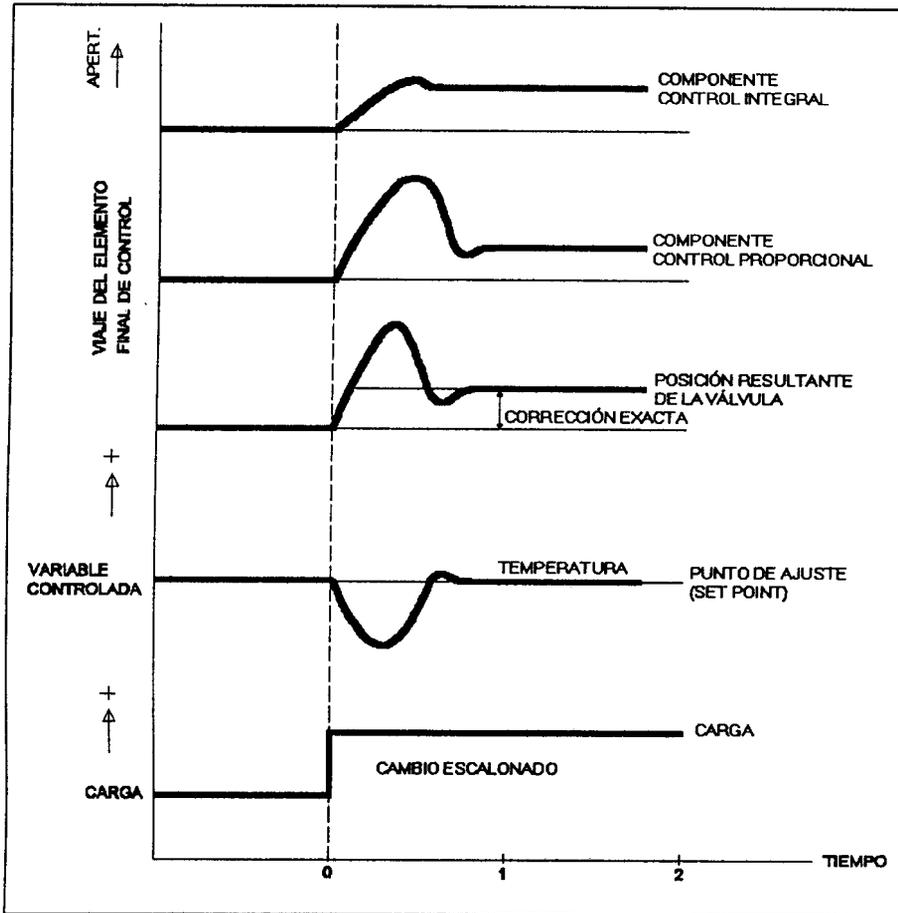


Figura 3. 7 Respuesta de un controlador Proporcional más Integral (PI) a un cambio de carga escalonado.

### 3. 2. 3. g Control Proporcional más Derivativo (PD)

En esta forma de control existe una relación lineal continua entre la velocidad de variación de la variable controlada y la posición del elemento final de control. En otras palabras, entre más rápido varíe la variable controlada, más rápido se moverá el elemento final de control intentando corregir la desviación.

Debido a que este tipo de control tiene como característica el oponerse a todas las variaciones, provoca que el sistema de control sea muy estable; sin embargo, no logra eliminar la desviación permanente, característica del control proporcional. En la figura 3.8 se ilustra la respuesta de esta forma de control a un cambio de carga escalonado.

La expresión matemática de esta forma de control es la siguiente:

$$U_o(t) = K_p U_i(t) + K_p T_v \frac{d}{dt} U_i(t) \quad (3.6)$$

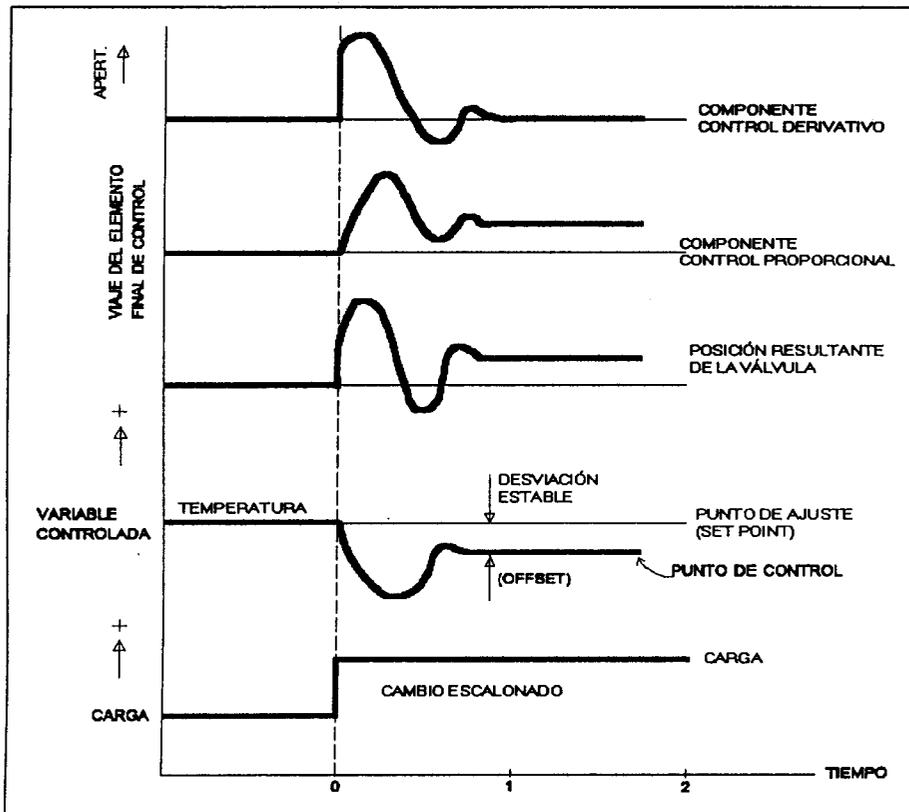


Figura 3. 8 Respuesta de un control Proporcional más Derivativo (PD) a un cambio de carga escalonado.

### 3. 2. 3. h Control Proporcional más Integral más Derivativo (PID)

La combinación de las formas de control proporcional, integral y derivativo permite diseñar un sistema que aprovecha las ventajas que cada una ofrece:

- El control proporcional permite corregir la posición del elemento final de control en una cantidad proporcional a la desviación.
- El control integral permite eliminar la desviación estable originada por el control proporcional, ya que corrige la posición del elemento final de control en un coeficiente proporcional a la desviación.
- El control derivativo permite reducir la oscilación del sistema de control y alinea, de manera rápida, a la variable controlada.

La ecuación característica de esta forma de control es la siguiente:

$$U_o(t) = \frac{1}{T_i} \int_0^t U_i(t) dt + K_p \left[ 1 + \frac{T_v}{T_n} \right] U_i(t) + T_d \frac{d[U_i(t)]}{dt} + U_o(0) \quad (3.7).$$

Como se observa, la ecuación se integra por cuatro sumandos, el primero corresponde a la parte integral; el segundo, corresponde a la parte proporcional; el tercero, corresponde a la parte derivativa; y el cuarto, corresponde al valor acumulado.

Al aplicar esta forma de control, la variable controlada tiende a desviarse menos del punto de ajuste (set point) y retorna de manera más rápida a dicho valor. Este tipo de control se emplea principalmente en procesos con grandes retrasos entre la señal de entrada y la respuesta del sistema. En la figura 3.9 se gráfica la respuesta de un controlador PID para un cambio de carga escalonado.

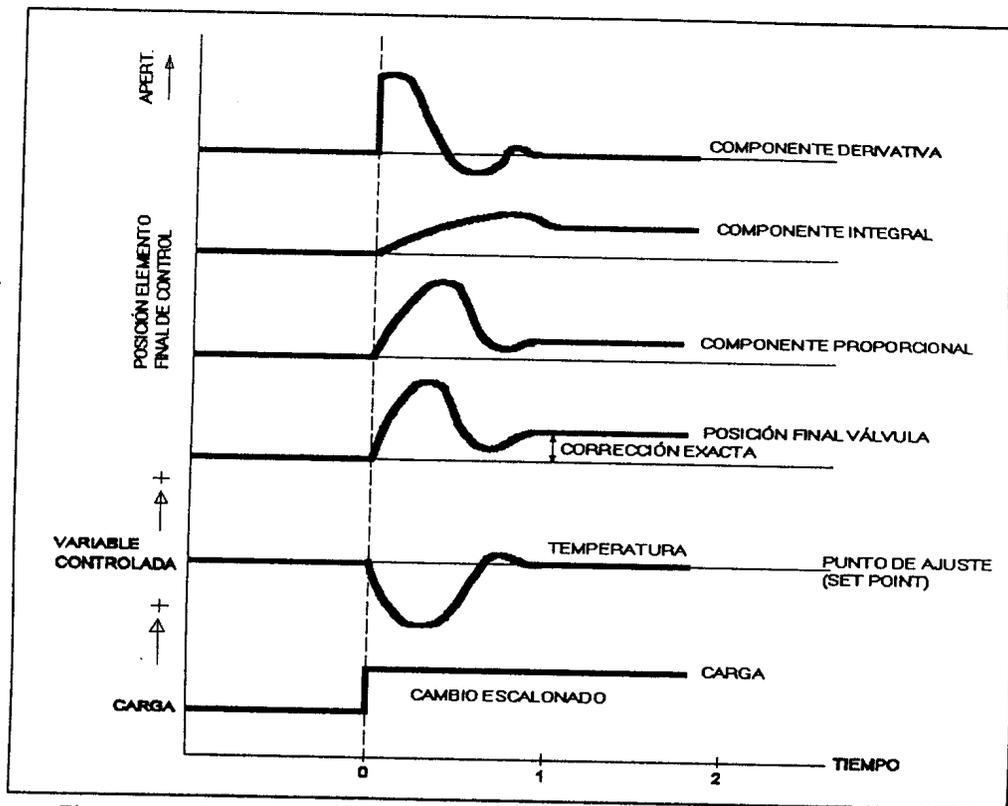


Figura 3.9 Respuesta de un controlador PID a un cambio de carga escalonado.

### 3.3 DESCRIPCIÓN DE LOS SISTEMAS DE CONTROL DIFUSO (SCD)

En los sistemas de control tradicional más utilizados (PI, PID) el control se ejerce aplicando una fórmula matemática directa, tal como se vio en la sección anterior. Sin embargo, en muchos casos las leyes de control del proceso en cuestión son difíciles de obtener o la relación matemática resultante es demasiado compleja. Por otra parte, puede ser que las condiciones de operación del proceso no sean totalmente conocidas y la experiencia del operador influya mucho en la ejecución del control. Debido a que la lógica difusa crea sus propias fuentes de control al combinar reglas y conjuntos difusos, ésta permite al diseñador construir sistemas de control aún cuando su entendimiento del modelo matemático del sistema en cuestión sea incompleto [3].

En los casos en que existe una alta no linealidad entre las variables de un proceso, en que el modelo matemático del proceso en cuestión es incompleto, éste no existe, o existe pero es difícil de entender, es cuando la lógica difusa describe el proceso y las leyes de control por medio de "reglas de sentido común" que se refieren a cantidades indefinidas (conjuntos difusos). Con reglas condicionales de la forma "si..., entonces..." se describen las estrategias de control y se crea un algoritmo difuso donde las variables objeto de control se transforman en variables lingüísticas. A la parte condicional de la regla se le denomina antecedente y a la parte de la acción se le denomina conclusión (o consecuente) [4]. El antecedente de una regla de control difuso expresa el estado en el que se deben encontrar las variables del proceso para activarla, mientras que su conclusión especifica los cambios que deben llevarse a cabo cuando se activa dicha regla. Expresado en términos de la ingeniería de control, el antecedente se relaciona con el estado actual de la variable de control, mientras que la conclusión se relaciona con el incremento o decremento que debe sufrir la variable manipulada para reducir la desviación o error.

El algoritmo difuso se define como un conjunto de reglas que describen un mapeo entre las variables de entrada y las variables de salida del sistema de control difuso (SCD). De esta manera, al tomar decisiones en función de los diferentes estados de un proceso se manejan términos cualitativos en lugar de términos cuantitativos.

Si un SCD tiene  $n$  entradas  $(x_1, \dots, x_n)^T$  y  $m$  salidas  $(z_1, \dots, z_m)^T$ , entonces éste puede equivalentemente representarse como  $m$  SCD cada uno con  $n$  entradas y una salida [4]. Por lo anterior, en este trabajo se considerará sólo sistemas de  $n$  entradas y una salida (MISO multiple input single output). El algoritmo difuso para un sistema MISO está compuesto por reglas de la forma:

$$r_j : \text{si } (x_1 \text{ es } A_1^j) \text{ AND } (x_2 \text{ es } A_2^j) \text{ AND } \dots \text{ AND } (x_n \text{ es } A_n^j), \text{ entonces } (z \text{ es } B^j) \quad (3.8)$$

donde:  $x_i$  ( $i=1,2,\dots,n$ ) son las variables de entrada al SCD,  $z$  es la variable de salida del SCD;  $A_j^i$  y  $B^j$  ( $j=1,2,\dots,m$ ;  $m$  es el número de reglas) son valores lingüísticos (conjuntos difusos) definidos para las variables de entrada y la variable de salida en los universos del discurso  $U_1, U_2, \dots, U_n$  y  $V$ , respectivamente. Estos valores lingüísticos representan términos vagos tales como: *pequeño*, *mediano* ó *grande*, y están caracterizados por las funciones de membresía  $\mu_{A_j^i}(x_i)$  y  $\mu_{B^j}(z)$ . De manera general, una regla difusa mapea el antecedente, formado por la intersección (AND) de los  $n$  valores lingüísticos dados por las variables ( $x_i$  es  $A_j^i$ ), al consecuente formado por un valor lingüístico dado por la variable ( $z$  es  $B^j$ ).

La regla (3.8) se puede reescribir de la forma:

$$R_j : \text{si } (\mathbf{x} \text{ es } A^j), \text{ entonces } (z \text{ es } B^j) \quad (3.9)$$

donde  $A^j$  es el  $j$ -ésimo conjunto difuso multivariable formado por la intersección difusa de los conjuntos difusos individuales  $A_i^j$  ( $A^j = A_1^j \text{ AND } A_2^j \text{ AND } \dots \text{ AND } A_n^j$ );  $\mathbf{x} = (x_1, \dots, x_n)^T$  y  $r_j$  representa la  $j$ -ésima regla. El algoritmo difuso consiste de un conjunto de estas reglas conectadas utilizando un operador de unión (OR difusa) para formar una base de reglas. Por ejemplo, para el sistema considerado, la base de reglas es de la forma:

$$\begin{aligned} r_1 &: \text{si } (\mathbf{x} \text{ es } A^1), \text{ entonces } (z \text{ es } B^1) \\ r_2 &: \text{OR si } (\mathbf{x} \text{ es } A^2), \text{ entonces } (z \text{ es } B^2) \\ &\vdots \\ r_m &: \text{OR si } (\mathbf{x} \text{ es } A^m), \text{ entonces } (z \text{ es } B^m) \end{aligned} \quad (3.10)$$

Evidentemente, para implementar el algoritmo difuso se deben definir los conjuntos difusos que representen los valores lingüísticos; así como las funciones a utilizar para realizar las operaciones difusas AND, si ( $\cdot$ ) entonces ( $\cdot$ ) y OR. Es necesario aclarar, que una vez definidos los conjuntos difusos así como los operadores, la relación entre las variables de entrada y la variable de salida del SCD queda nítidamente especificada. Es decir, esta relación deja de ser vaga o incierta, ahora es una función multivariable no lineal y determinística. En este sentido, la lógica difusa funciona como una interface entre la interpretación imprecisa que se hace del mundo real (fuera del SCD), expresada en términos lingüísticos, y la representación matemática que se hace del mismo dentro del controlador difuso [5].

La estructura básica de un SCD consta de tres etapas fundamentales: *fuzzificación (fuzzification)*, *evaluación de reglas (rule evaluation)*, y *defuzzificación (defuzzification)*<sup>1</sup>. En la figura 3.10 se muestra la estructura básica de un SCD.

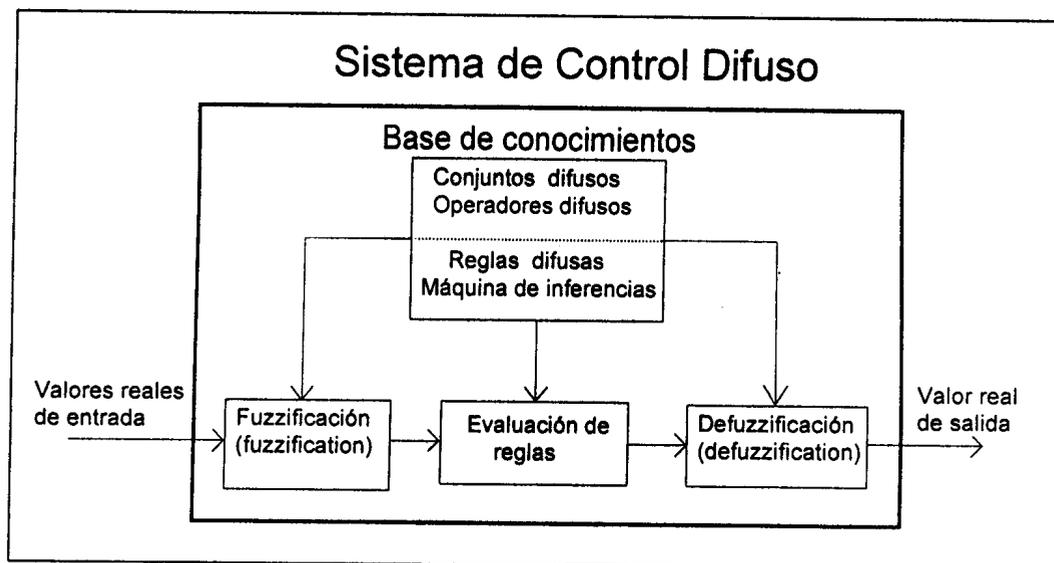


Figura 3. 10 Estructura básica de un sistema de control difuso.

A continuación se hace una breve descripción de cada una de las etapas que componen un SCD.

<sup>1</sup> La traducción al español de las palabras fuzzification y defuzzification generalmente se realiza como emborronamiento y desemborronamiento, pero el autor considera más adecuados los términos fuzzificación y defuzzificación; los cuales han sido utilizados, últimamente, en varios textos en español.

### 3.3.1 Fuzzificación (Fuzzification)

En la mayoría de los controladores difusos las entradas son valores reales (también nombrados en la literatura de los SCD como valores nítidos) o no difusos, sin embargo, estos valores al ser obtenidos por mediciones pueden estar corrompidos por ruido. De la misma manera, las entradas podrían ser sentencias lingüísticas tales como *x es grande*. Para representar la incertidumbre asociada con cada entrada se utiliza el procedimiento conocido como fuzzificación.

El proceso de **fuzzificación (fuzzification)** consiste en tomar el valor nítido de una variable de entrada al SCD y obtener su representación difusa, es decir, su función de membresía. De manera formal, considerando un sistema MISO:  $U \subset R^n \rightarrow R$ , donde  $U$  es compacto, la etapa de *fuzzificación* es el proceso a través del cual se realiza un mapeo de un espacio de entradas  $U \subset R^n$  a conjuntos difusos definidos en  $U$ , donde un conjunto difuso está caracterizado por una función de membresía  $\mu_F: U \rightarrow [0, 1]$ , y una etiqueta lingüística  $F$  [6], tal como *pequeño*, *muy pequeño*, *grande*, *muy grande*, etc. Por ejemplo, si  $X$  es una variable de entrada al SCD y  $X = x_0 \in U$ , entonces la salida de la etapa de fuzzificación es un conjunto difuso en  $U$ ,  $F = \text{fuzzifier}(x_0)$ ; en donde el operador *fuzzifier* transforma el valor de entrada  $x_0$  en un valor lingüístico o conjunto difuso.

Principalmente existen 2 tipos de fuzzificación:

a) Fuzzificación singleton:

$$F(x) = \begin{cases} 1 & \text{si } x = x_0 \\ 0 & \text{en otro caso} \end{cases} \quad (3.11).$$

b) Fuzzificación aproximada:

$$F(x) \neq 0 \quad \text{si y sólo si} \quad |x - x_0| < \delta \quad (3.12)$$

en donde  $\delta$  es un parámetro que se determina de acuerdo al contexto de cada aplicación.

En la mayoría de las aplicaciones reportadas en la literatura [7] la función  $F = \text{fuzzifier}(x_0)$  toma la especial forma de  $F = x_0$  para cada valor medido de la variable en cuestión. Es decir, el valor concreto en la entrada se mapea a un singleton definido por el punto  $x_0 \in U$ .

Cuando la entrada al SCD es una medición, la fuzzificación más utilizada es la singleton. Cuando las entradas están corrompidas por ruido, la fuzzificación más utilizada es la aproximada. Por ejemplo, podría utilizarse un conjunto difuso triangular en donde el vértice corresponda a la media de las mediciones hechas y el ancho de la base ( $\delta$ ) sería una función de la desviación estándar. El propósito de la etapa de fuzzificación, en este caso, es interpretar cada valor medido de las variables de entrada como una aproximación difusa más realista que el propio valor medido.

### 3.3.2 Evaluación de Reglas (Rule Evaluation)

Desde un punto de vista general, la evaluación de reglas involucra una base de reglas difusa de la forma dada en (3.10) y una *máquina de inferencias difusa*. La máquina de inferencias es un interprete de la base de reglas, su tarea consiste en obtener una conclusión difusa empleando una lógica de toma de decisiones, llamada "*razonamiento aproximado*".

El razonamiento aproximado es un mecanismo de inferencias difuso utilizado para obtener una conclusión difusa a partir de un conjunto de reglas de la forma "*si..., entonces...*"; y una o más condiciones. La conclusión difusa corresponde a los valores que han ingresado al sistema de control y que han pasado, en primera instancia, por la etapa de fuzzificación.

El mecanismo de inferencias empleado en los controladores difusos es generalmente mucho más sencillo que el utilizado en los sistemas expertos convencionales. En los controladores difusos la conclusión de cada regla no se aplica como antecedente de alguna otra, por lo cual, no se emplea el mecanismo de encadenamiento para realizar una inferencia. En un sistema difuso se emplea un mecanismo de inferencias de búsqueda hacia adelante y en un sólo sentido (one-level forward data-driven inference) [4].

Antes de describir como se efectúa el razonamiento aproximado es necesario establecer la manera en que se implementa una regla y una base de reglas en el ámbito de la lógica difusa. El modo de implementar lógicamente una regla de control difuso de la forma:  $r_j: \text{si } (x \text{ es } A^i), \text{ entonces } (z \text{ es } B^j)$ ; es a través de una *implicación difusa o relación difusa* (fuzzy relation)

[8]. El grado en que la variable  $x$  esta relacionada con la variable  $z$  se representa por la relación difusa  $r_j$ , cuya función de membresía  $\mu_{r_j}(x, z)$  definida en el espacio producto  $A_1 \times \dots \times A_n \times B$  esta dada por:

$$\mu_{r_j}(x, y) = t (\mu_{A_i}(x), \mu_{B_j}(z)) \tag{3.13}$$

donde  $t$  es una T-norma, generalmente se emplea el mínimo (min) o el producto algebraico; pero cualquier T-norma podría aplicarse. La operación dada por (3.13) consiste en realizar la intersección de las  $(n+1)$  funciones de membresía que representan los valores lingüísticos de las variables de entrada y la variable de salida del SCD. El conjunto difuso  $\mu_{r_j}(x, z)$  representa la *confianza* en que la salida sea  $z$  dado que la entrada es  $x$  para la  $j$ -ésima regla. Nótese que el conjunto difuso resultante es un subconjunto del espacio producto  $U_1 \times \dots \times U_n \times V$  ya que  $A_1 \in U_1, A_2 \in U_2, \dots, A_n \in U_n, B \in V$ .

Si existen  $m$  reglas que mapean  $x$  en  $z$ , entonces existen  $m$  relaciones que pueden conectarse con el operador unión (OR) para formar una base de reglas  $BR$ , esto es:

$$\mu_{BR}(x, y) = \bigcup_j \mu_{r_j}(x, z) \tag{3.14}$$

Esta operación se define como:

$$\mu_{BR}(x, z) = s (\mu_{r_1}(x, z), \mu_{r_2}(x, z), \dots, \mu_{r_m}(x, z)) \tag{3.15}$$

donde  $s$  es una conorma triangular o S-norma. Las S-normas más utilizadas son el máximo (max) y la suma probabilística. La unión de todas las relaciones individuales forman una superficie de relación en el espacio entradas-salida que representa cómo pares individuales de entradas-salida están relacionados, y cómo, empleando el operador composición, éstos se pueden utilizar para inferir una función de membresía (conjunto difuso) de salida dado un conjunto particular de entradas.

Para ilustrar cómo se crea la superficie de relación considérese un sistema de una entrada, una salida. Cada una de estas variables tiene definidos cuatro conjuntos difusos que representan los términos lingüísticos *casi cero*, *pequeño*, *mediano* y *grande*; tal como se muestra en la figura 3. 11.

Sea la base de reglas para este sistema la que se da enseguida:

- si ( $x$  es casi cero), entonces ( $z$  es casi cero)
- OR si ( $x$  es pequeño), entonces ( $z$  es pequeño)
- OR si ( $x$  es mediano), entonces ( $z$  es mediano)
- OR si ( $x$  es grande), entonces ( $z$  es grande).

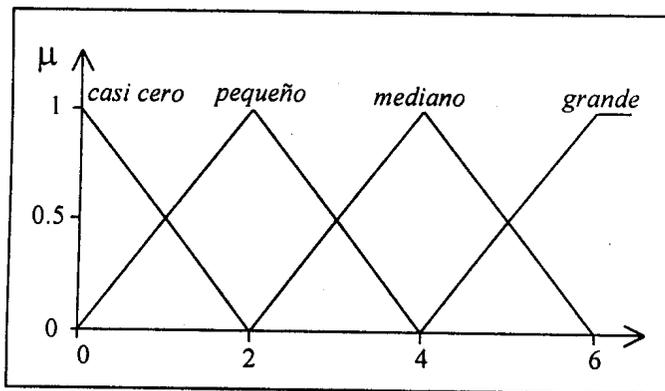


Figura 3. 11 Funciones de membresía para la entrada y la salida.

Si se utiliza la T-norma del producto y la S-norma de la suma, entonces la superficie de relación difusa sería la mostrada en la figura 3. 12.

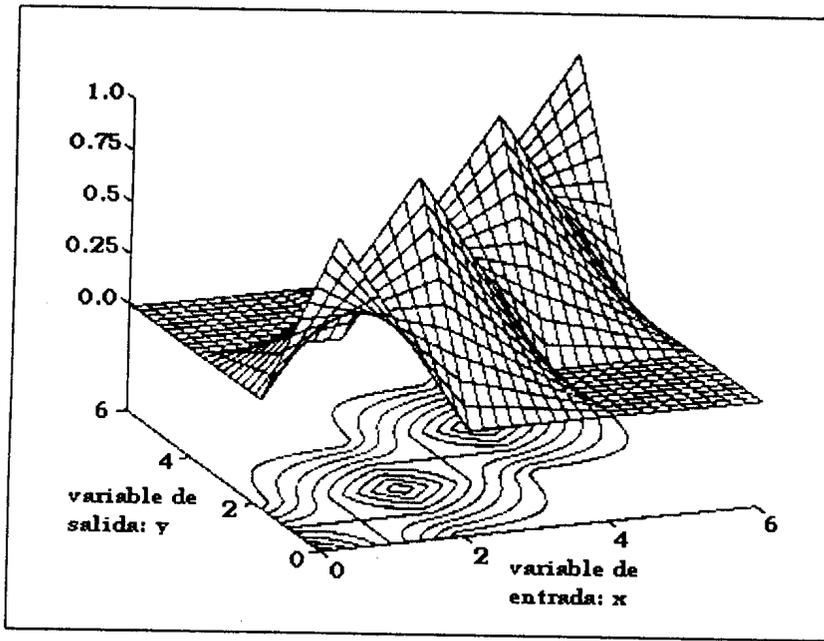


Figura 3.12 Superficie de la relación difusa y contorno asociado para un sistema de una entrada-una salida.

Una vez que se ha formado la base de reglas, ésta puede utilizarse para generar un conjunto difuso de salida para un conjunto dado de entradas. Para ello se utiliza el procedimiento conocido como *regla composicional de inferencias*. Sea  $\mu_{BR}(x, z)$  una base de reglas, entonces el conjunto difuso de salida  $\mu_B(z)$  inducido por el conjunto difuso de entrada  $\mu_A(x)$  esta dado por:

$$\mu_B(z) = \mu_A(x) \circ \mu_{BR}(x, z) \quad (3.16)$$

donde  $\circ$  representa el operador composición, el cual esta definido por:

$$\mu_B(z) = s(t(\mu_A(x), \mu_{BR}(x, z))) \quad (3.17)$$

donde  $s$  es una S-norma y  $t$  es una T-norma que operan tal como se describió en el capítulo 2. Es evidente que se pueden generar varias definiciones para la composición dependiendo de los operadores que se elijan para  $s$  y para  $t$ . Cuando se elige a  $s$  como el operador integral y a  $t$  como el operador producto, entonces la composición esta dada por:

$$\mu_B(z) = \int \mu_A(x) * \mu_{BR}(x, z) dx \quad (3.18)$$

que requiere la realización de una integral n-dimensional para un conjunto difuso arbitrario de entrada. La integral también puede aproximarse con una rutina de integración numérica (discreta).

Si  $s$  y  $t$  se eligen como el operador *max* y *min* respectivamente, entonces se define la regla composicional de inferencias *max-min*, también conocida como *razonamiento correlación mínimo* [4]:

$$\mu_B(z) = \max_x (\min(\mu_A(x), \mu_{BR}(x, z))) \quad (3.19)$$

que también involucra dificultades numéricas dado que se tiene que realizar una optimización no lineal en  $n$  dimensiones para obtener el máximo sobre  $U_1 \times \dots \times U_n$ .

Si se elige a  $t$  como el operador producto y a  $s$  como el operador *max*, entonces se tiene la regla composicional de inferencias *max-product*, también conocida como *razonamiento correlación producto* [4], definida por:

$$\mu_B(z) = \max_x (\mu_A(x) * \mu_{BR}(x, z)) \quad (3.20)$$

Utilizando la regla composicional de inferencias se formaliza el procedimiento del razonamiento aproximado. La regla básica para realizar inferencias en la lógica tradicional es el modus ponens, de acuerdo al cual se puede inferir la verdad de una proposición  $B$  a partir de la verdad de una proposición  $A$  y la implicación  $A \rightarrow B$ . Por ejemplo, si se define  $A$  como "el jitomate esta rojo" y  $B$  como "el jitomate esta maduro", entonces si la sentencia "el jitomate esta rojo" es verdadera la sentencia "el jitomate esta maduro" también es verdadera. Este concepto se ilustra de la siguiente manera:

$$\begin{array}{ll} \text{premisa 1 (hecho):} & x \text{ es } A \\ \text{premisa 2 (regla):} & \text{si } x \text{ es } A \text{ entonces } z \text{ es } B \\ \hline \text{consecuencia (conclusión):} & z \text{ es } B. \end{array} \quad (3.21).$$

Sin embargo, en el razonamiento humano el modus ponens generalmente se emplea de manera aproximada. Por ejemplo, si se tiene la misma regla "si el tomate esta rojo, entonces esta maduro" pero ahora se sabe que "el tomate esta más o menos rojo", entonces se puede inferir que "el tomate esta más o menos maduro". Esto se puede escribir de la siguiente forma:

$$\begin{array}{ll} \text{premisa 1 (hecho):} & x \text{ es } A' \\ \text{premisa 2 (regla):} & \text{si } x \text{ es } A \text{ entonces } z \text{ es } B \\ \hline \text{consecuencia (conclusión):} & z \text{ es } B' \end{array} \quad (3.22)$$

en donde la sentencia  $A'$  esta cercana a la sentencia  $A$  y  $B'$  a la sentencia  $B$ . Cuando  $A$ ,  $B$ ,  $A'$  y  $B'$  representan conjuntos difusos en universos del discurso apropiados la inferencia, arriba citada, se denomina modus ponens generalizado (generalized modus ponens GMP) ya que tiene al modus ponens como un caso particular [4].

Utilizando el GMP se plantea enseguida el procedimiento para realizar inferencias considerando, en primera instancia, el caso de una regla con un antecedente; posteriormente se analizará una regla con dos antecedentes, y por último, se considerará un sistema con dos entradas y  $m$  reglas.

Sea un sistema SISO (single input-single output); sean  $A$  y  $A'$  valores lingüísticos (conjuntos difusos) de la variable de entrada  $x$ , cuyo universo del discurso es  $U$ . Sean  $B$  y  $B'$  valores lingüísticos de la variable de salida  $z$ , cuyo universo del discurso es  $V$ . Asíumase que la implicación difusa  $A \rightarrow B$  esta expresada como la relación difusa  $R$  en  $U \times V$ . De acuerdo a lo anterior, el conjunto difuso  $B'$  inducido por la sentencia "x es  $A'$ " y la regla difusa "si  $x$  es  $A$ , entonces  $z$  es  $B$ " esta dado por:

$$\mu_{B'}(z) = s(t(\mu_{A'}(x), \mu_R(x, z))) \quad (3.23)$$

o de manera equivalente:

$$B' = A' \circ R = A' \circ (A \rightarrow B) \quad (3.24)$$

La ecuación (3.23) puede reducirse. Antes, para una fácil manipulación de los operadores  $t$  y  $s$ , éstos se substituyen con los símbolos de intersección  $\wedge$  y de unión  $\vee$  respectivamente. De tal forma la ecuación citada se reescribe como:

$$\mu_{B'}(z) = \vee_x [\mu_{A'}(x) \wedge \mu_R(x, z)] \quad (3.25).$$

De esta forma, substituyendo  $\mu_R(x) = \mu_A(x) \wedge \mu_B(z)$  en (3.25), la reducción se realiza como sigue:

$$\begin{aligned} \mu_{B'}(z) &= \vee_x [\mu_{A'}(x) \wedge \mu_A(x) \wedge \mu_B(z)] \\ &= \vee_x [\mu_{A'}(x) \wedge \mu_A(x)] \wedge \mu_B(z) \\ &= w \wedge \mu_B(z) \end{aligned} \quad (3.26).$$

Si en (3.26) se utilizan los operadores  $\max$  y  $\min$  para  $\vee$  y  $\wedge$ , respectivamente; entonces se tiene que primeramente se encuentra el grado de compatibilidad "w" entre  $A$  y  $A'$  ( $w = \vee_x [\mu_{A'}(x) \wedge \mu_A(x)]$ ). Posteriormente, el conjunto difuso que representa  $B'$  resulta de realizar un corte al conjunto  $B$  en el punto dado por  $w$ , tal como se muestra en la figura 3.13. En

dicha figura  $w$  corresponde al área sombreada en la parte del antecedente y  $B'$  representa el área sombreada en la parte del consecuente. El término  $w$  también se puede interpretar como el *grado de activación o disparo* de la regla difusa dada como  $R$ .

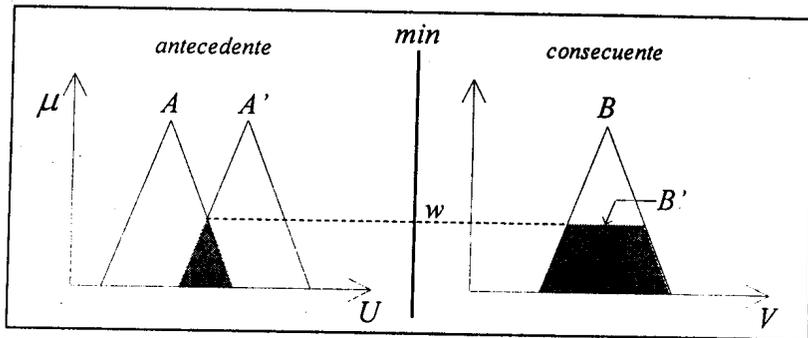


Figura 3. 13 Razonamiento aproximado para una regla con un antecedente y un consecuente.

Ahora, sea un sistema MISO; sea una regla con dos antecedentes de la forma "si  $x$  es  $A$  y  $y$  es  $B$  entonces  $z$  es  $C$ ", el problema correspondiente para el razonamiento aproximado es:

premisa 1 (hecho):	$x$ es $A'$ y $y$ es $B'$	
premisa 2 (regla):	si $x$ es $A_1$ y $y$ es $B_1$ entonces $z$ es $C_1$	(3.27)
consecuencia (conclusión):	$z$ es $C'$	

donde  $x, y, z$  son variables lingüísticas representando dos variables de entrada y una de salida, respectivamente;  $A_1, B_1, C_1$  son valores (conjuntos difusos) de las variables  $x, y, z$  en los universos de discurso  $U, V, W$

La regla difusa en la premisa dos, mostrada arriba, se puede escribir como " $A \times B \rightarrow C$ " (en el análisis se han eliminado los subíndices para evitar confusiones). La traslación de esta regla a su forma lógica es la relación difusa  $R$ , cuya función de membresía (FM) se especifica de la siguiente manera:

$$\begin{aligned} \mu_R(x, y, z) &= \mu_{(A \times B) \rightarrow C}(x, y, z) \\ &= \mu_A(x) \wedge \mu_B(y) \wedge \mu_C(z) \end{aligned} \quad (3.28)$$

y el resultado  $C'$  se expresa como:

$$C' = (A' \times B') \circ (A \times B \rightarrow C) \quad (3.29)$$

La FM resultante es:

$$\begin{aligned} \mu_{C'}(z) &= \vee_{x,y} [\mu_{A'}(x) \wedge \mu_{B'}(y)] \wedge [\mu_A(x) \wedge \mu_B(y) \wedge \mu_C(z)] \\ &= \vee_{x,y} [\mu_{A'}(x) \wedge \mu_{B'}(y) \wedge \mu_A(x) \wedge \mu_B(y)] \wedge \mu_C(z) \\ &= \{\vee_x [\mu_{A'}(x) \wedge \mu_A(x)]\} \wedge \{\vee_y [\mu_{B'}(y) \wedge \mu_B(y)]\} \wedge \mu_C(z) \end{aligned} \quad (3.30)$$

haciendo:

$$\begin{aligned} w_A &= \vee_x [\mu_{A'}(x) \wedge \mu_A(x)] \\ w_B &= \vee_y [\mu_{B'}(y) \wedge \mu_B(y)] \end{aligned} \quad (3.31)$$

y substituyendo (3.30) en (3.29) se tiene:

$$\mu_{C'}(z) = w_A \wedge w_B \wedge \mu_C(z) \quad (3.32)$$

haciendo  $w = w_A \wedge w_B$ , queda finalmente:

$$\mu_{C'}(z) = w \wedge \mu_C(z) \quad (3.33)$$

en (3.31)  $w_A$  representa el grado de compatibilidad entre  $A$  y  $A'$ ;  $w_B$  representa el grado de compatibilidad entre  $B$  y  $B'$ . En (3.32)  $w$  representa el grado de activación o disparo de la regla  $R$  [9]. Una interpretación gráfica de este resultado se muestra en la figura 3.14.

La interpretación de múltiples reglas generalmente se realiza como la unión de las relaciones correspondientes a cada regla difusa, por ejemplo, sea el siguiente hecho y reglas:

premisa 1 (hecho):	$x$ es $A'$ y $y$ es $B'$	
premisa 2 (regla 1):	si $x$ es $A_1$ y $y$ es $B_1$ entonces $z$ es $C_1$	
premisa 3 (regla 2):	si $x$ es $A_2$ y $y$ es $B_2$ entonces $z$ es $C_2$	(3.34)
<hr/>		
consecuencia (conclusión):	$z$ es $C'$	

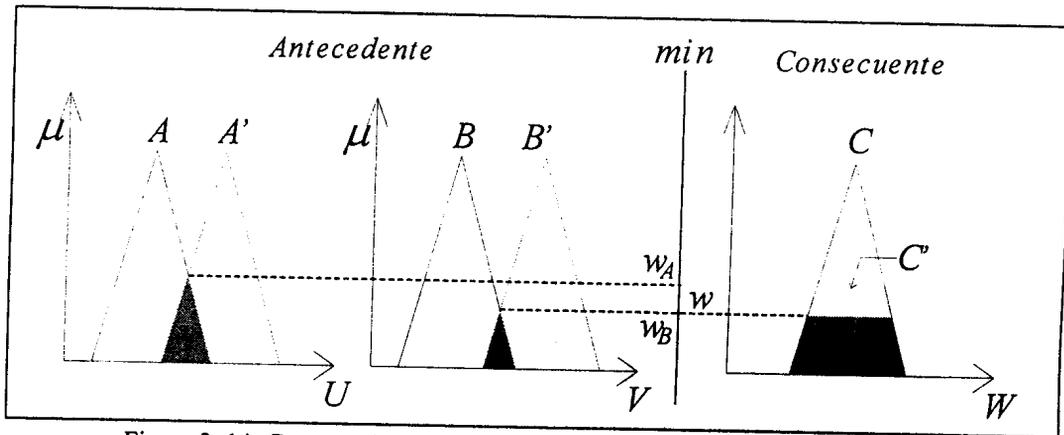


Figura 3.14 Razonamiento aproximado para una regla con dos antecedentes.

La conclusión  $C'$  se obtiene de la siguiente manera, sean  $R_1 = (A_1 \times B_1) \rightarrow C_1$  y  $R_2 = (A_2 \times B_2) \rightarrow C_2$ . Dado que la composición es distributiva sobre la unión se obtiene:

$$\begin{aligned} C' &= (A' \times B') \circ (R_1 \vee R_2) \\ &= [(A' \times B') \circ R_1] \vee [(A' \times B') \circ R_2] \\ &= C'_1 \vee C'_2 \end{aligned} \quad (3.35)$$

De la ecuación (3.32), encontrada en el ejemplo anterior, se tiene que:

$$\mu_{C'_1}(z) = w_1 \wedge \mu_{C_1}(z) \quad (3.36)$$

$$\mu_{C'_2}(z) = w_2 \wedge \mu_{C_2}(z) \quad (3.37)$$

finalmente:

$$\mu_{C'}(z) = [w_1 \wedge \mu_{C_1}(z)] \vee [w_2 \wedge \mu_{C_2}(z)] \quad (3.38)$$

donde  $w_1$  y  $w_2$  indican los grados de activación de las reglas 1 y 2, respectivamente.  $\mu_{C_1}(z)$  y  $\mu_{C_2}(z)$  son las FM de las conclusiones difusas  $C_1$  y  $C_2$ , respectivamente.

De acuerdo a este resultado, para el caso en el que se tiene una base de  $m$  reglas de la forma:

$$\begin{array}{ll}
 \text{premisa 1 (hecho):} & x \text{ es } A' \text{ y } y \text{ es } B' \\
 \text{premisa 2 (regla 1):} & \text{si } x \text{ es } A_1 \text{ y } y \text{ es } B_1, \text{ entonces } Z \text{ es } C_1; \\
 \text{OR premisa 3 (regla 2):} & \text{si } x \text{ es } A_2 \text{ y } y \text{ es } B_2, \text{ entonces } Z \text{ es } C_2; \\
 \vdots & \vdots \\
 \text{OR premisa (m+1) (regla m):} & \text{si } x \text{ es } A_m \text{ y } y \text{ es } B_m, \text{ entonces } Z \text{ es } C_m;
 \end{array} \quad (3.39)$$


---


$$Z \text{ es } C'$$

donde  $x, y, z$  son variables lingüísticas representando dos variables de entrada y una variable de salida, respectivamente;  $A, B, y C_i$ , son valores (conjuntos difusos) de las variables  $x, y, z$  en los universos de discurso  $U, V, y W$ , respectivamente; con  $i = 1, 2, \dots, m$ ; y con todas las reglas ligadas con el conector OR constituyendo una base de reglas. La FM del conjunto difuso  $C'$  de salida esta dada por:

$$C' = (w_1 \wedge C_1) \vee (w_2 \wedge C_2) \vee \dots \vee (w_m \wedge C_m) \quad (3.40)$$

o de manera equivalente:

$$\mu_{C'}(z) = \sum_{i=1}^m w_i \wedge \mu_{C_i}(z) \quad (3.41)$$

donde  $w_i$  indica el grado de activación de la regla  $R_i$ .  $\mu_{C_i}(z)$  es la FM de la conclusión difusa  $C_i$  ( $i=1, 2, \dots, m$ ;  $m =$  número de reglas); el símbolo  $\Sigma$  indica unión y no suma aritmética.

La figura 3.15 muestra un ejemplo de la obtención de la conclusión difusa  $C'$  para el caso de 2 reglas, con 2 entradas ( $X_0$  y  $Y_0$ ), fuzzificación tipo singleton y utilizando la composición *max-min*. La figura 3.16 muestra un ejemplo de la obtención de la conclusión difusa  $C'$ , también para 2 reglas, con 2 entradas ( $X_0$  y  $Y_0$ ), fuzzificación tipo singleton, pero ahora utilizando la composición *max-product*.

Las figuras (3.15) y (3.16) muestran los razonamientos difusos definidos previamente. Sin embargo, en la practica un sistema de inferencias difuso puede tener ciertos mecanismos de razonamiento que no necesariamente siguen, de manera estricta, la definición de la regla composicional de inferencias. Por ejemplo, para obtener el grado de activación de una regla se puede utilizar el producto o el mínimo. Otra variación es la utilización de la suma aritmética en lugar del máximo en el razonamiento correlación producto. A esta composición (o razonamiento difuso) se le denomina *sum-product* y al SCD que lo utiliza se le nombra "*sistema difuso aditivo*" (SDA) [10]. La diferencia de este último razonamiento con respecto a los dos anteriores es que el conjunto difuso de salida  $C'$  resulta de realizar la suma de los conjuntos parciales  $C'_i$ . Esto es:

$$C' = (w_1 * C_1) + (w_2 * C_2) + \dots + (w_m * C_m) \quad (3.42)$$

o de manera equivalente:

$$\mu_{C'}(z) = \sum_{i=1}^m w_i * \mu_{C_i}(z) \quad (3.43)$$

donde  $w_i$  indica el grado de activación de la regla  $R_i$ .  $\mu_{C_i}(z)$  es la FM de la conclusión difusa  $C_i$  ( $i=1, 2, \dots, m$ ;  $m =$  número de reglas); en este caso el símbolo  $\Sigma$  sí indica suma aritmética; el símbolo  $*$  significa producto.

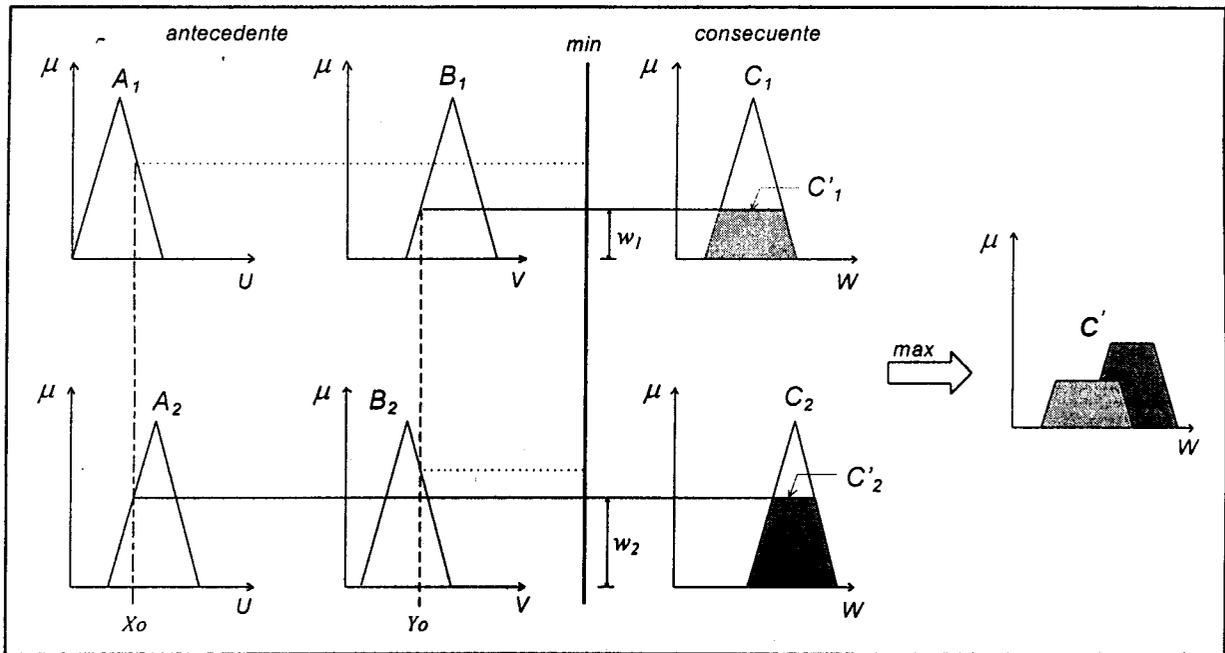


Figura 3. 15 Ejemplo de la obtención de la conclusión difusa  $C'$  utilizando la composición *max-min*.

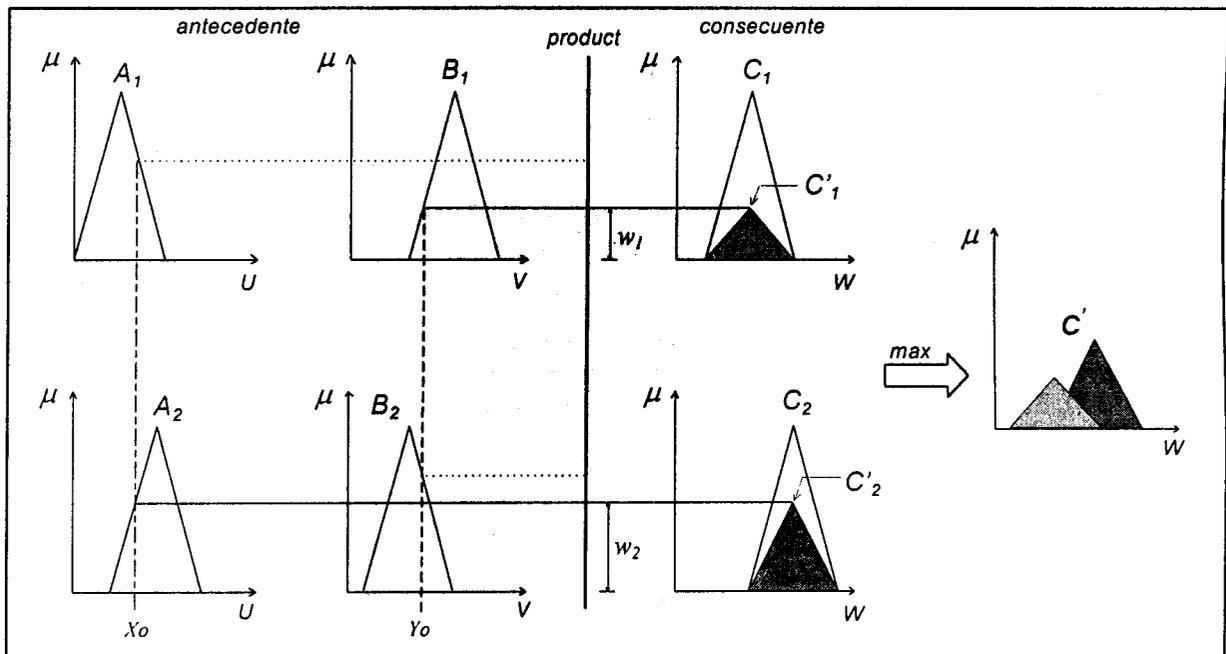


Figura 3. 16 Ejemplo de la obtención de la conclusión difusa  $C'$  utilizando la composición *max-product*.

La principal característica de un SDA es que los traslapes entre los conjuntos  $C'_i$  son tomados en cuenta al calcular  $C'$ , cosa que no sucede en los otros razonamientos. Lo anterior también provoca que la altura del conjunto difuso de salida  $C'$  cambie de tal forma que, si las entradas al SDA cambian lentamente, entonces el conjunto difuso de salida cambiará de la misma forma. Por otro lado, en los sistemas difusos en donde el conjunto de salida  $C'$  se obtiene utilizando el operador *max*, al ignorarse los traslapes también se ignoran los pequeños cambios que se presentan en las entradas; ya que estos pequeños cambios precisamente afectan las zonas de traslape. Otra importante diferencia es que en un SDA a medida que crece el número de conjuntos parciales  $C'_i$  que contribuyen a formar la conclusión difusa  $C'$ , la FM de ésta tiende a una forma gaussiana o a una curva simétrica unimodal [11]. Mientras que en un SCD que utiliza la composición *max-min* o *max-product* a medida que crece el número de conjuntos parciales  $C'_i$ , que contribuyen a formar la conclusión difusa  $C'$ , la FM resultante tenderá a un rectángulo [11]. La figura 3.17 muestra un ejemplo de la obtención de la conclusión difusa  $C'$  para 2 reglas, con 2 entradas ( $X_0$  y  $Y_0$ ), fuzzificación tipo singleton, y utilizando la composición *sum-product*.

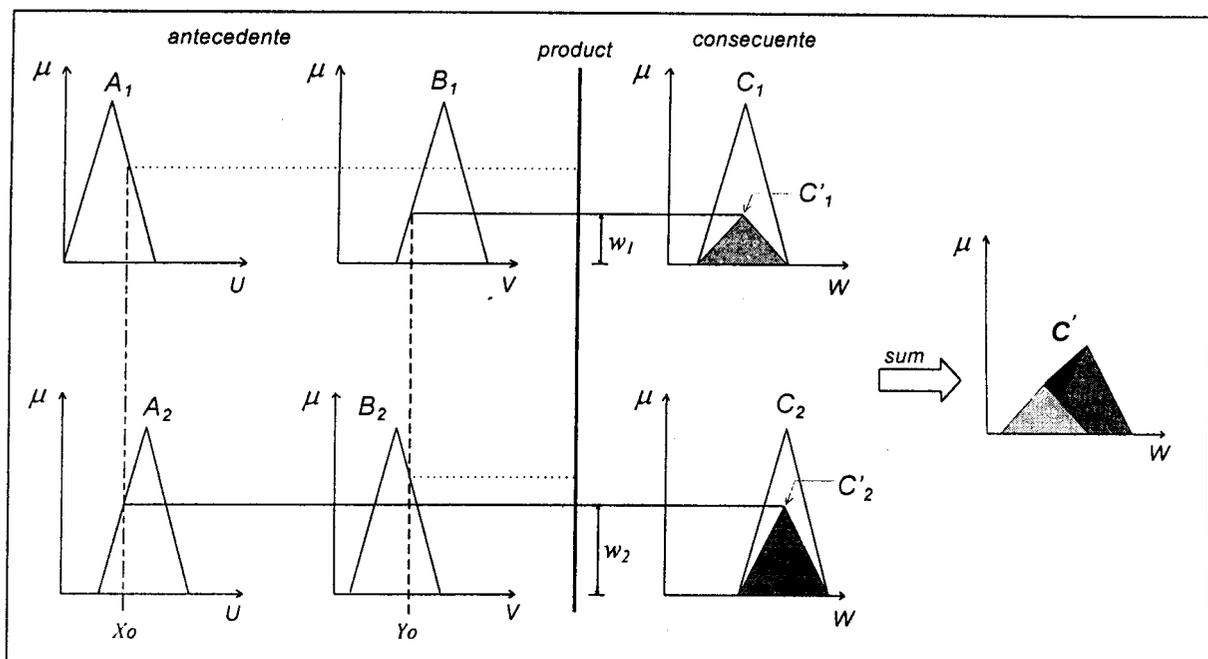


Figura 3. 17 Ejemplo de la obtención de la conclusión difusa  $C'$  utilizando la composición *sum-product*.

### 3. 3. 3 Defuzzificación (Defuzzification)

Hasta aquí, se ha llegado a la obtención de una conclusión  $C'$  que sigue siendo difusa, por lo que si se desea una salida 'nítida' (valor concreto no difuso) del SCD, es necesario realizar el proceso conocido como *defuzzificación*. Este proceso consiste en transformar o compactar la conclusión difusa  $C'$  en un valor específico y concreto; valor que es ejecutable y que sale del SCD para efectuar la acción de control.

Actualmente existen varios métodos para efectuar el proceso de *defuzzificación*, de entre ellos, los más utilizados son tres: a) centro de gravedad o centroide; b) criterio del máximo y c) promedio de influencias. A continuación se describe brevemente en que consiste cada uno de ellos.

#### 3. 3. 3. a Método del Centro de Gravedad o Centroide (Center Of Area method COA)

El mejor método disponible para realizar el proceso de *defuzzificación* es el del centro de gravedad o centroide, ya que éste considera la contribución de cada regla disparada durante el proceso de evaluación de reglas. Matemáticamente, el método del centroide se define por:

$$z_o = \frac{\int_z \mu_{C'}(z) z dz}{\int_z \mu_{C'}(z) dz} \quad (3.44)$$

donde:  $\int_z \mu_{C'}(z) dz$  es el área de la conclusión difusa  $C'$  y  $z_o$  es el centro de gravedad (acción de control concreta) de la función de membresía resultante  $\mu_{C'}(z)$ .

#### 3. 3. 3. b Método del Máximo (Max Criterion Method MAX)

Este método selecciona como salida la conclusión (conjunto difuso) de aquella regla que tenga el máximo grado de activación. De este conjunto, el punto en  $W$  que tiene el máximo valor de pertenencia al mismo, es la salida  $z_o$ . Por ejemplo, si durante el proceso de evaluación de reglas se dispararan 3 reglas  $r_1, r_2$  y  $r_3$  obteniéndose para cada una los grados de influencia  $w_1 = 0.75$ ,  $w_2 = 0.40$ , y  $w_3 = 0.20$  respectivamente; entonces el conjunto de salida que se selecciona como salida es el dado por el consecuente de la regla  $r_1$ , ya que ésta posee el máximo grado de activación con respecto a las otras dos. La salida  $z_o$  es, entonces, el punto en  $W$  que tenga el máximo grado de pertenencia en dicho conjunto. Matemáticamente esto se expresa como:

$$z_o = \max_{z \in W} \mu_{C'}(z) \quad (3.45)$$

El método del máximo es el más sencillo, pero también es el menos usado, debido a que sólo considera aquella regla que posee el máximo grado de activación, ignorando la posible contribución de las demás reglas. Por otra parte, en el caso de que dos reglas posean un mismo valor  $w$ , y éste sea máximo, se crea un conflicto que debe ser resuelto agregando alguna rutina extra.

### 3.3.3. c Promedio de Influencias (Mean of Maximum Method MOM)

Esta técnica genera una acción de control que representa el valor promedio de todas las acciones de control locales dadas por aquellas reglas disparadas durante el proceso de evaluación de reglas. Su expresión matemática es la siguiente:

$$z_o = \sum_{i=1}^n \frac{w_i}{n} \quad (3.46)$$

donde  $w_i$  es el grado de activación asignado a cada regla, y  $n$  es el número de reglas activadas ( $w_i > 0$ ).

En la figura 3.18 se muestra una interpretación gráfica de los distintos métodos para realizar el proceso de defuzzificación para el caso de dos reglas disparadas durante el proceso de evaluación de reglas.

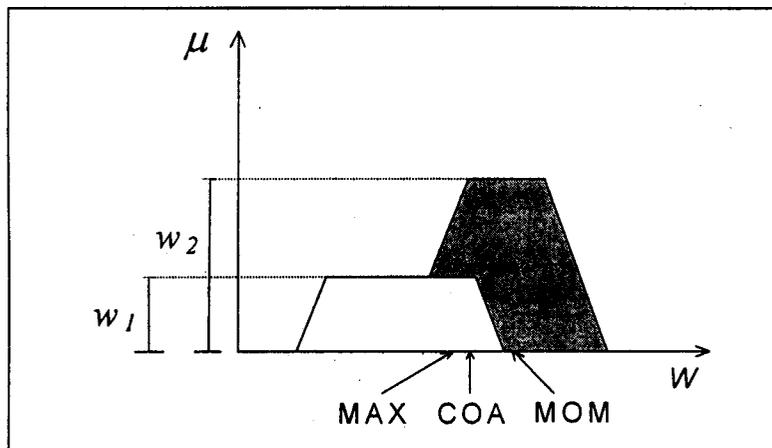


Figura 3.18 Representación gráfica de los métodos de defuzzificación.

En algunos estudios se ha demostrado que cuando se utiliza la estrategia MOM el desempeño de un sistema de control difuso es semejante a un sistema multinivel de relevadores, mientras que si se utiliza la estrategia COA el comportamiento es similar al obtenido con un control convencional PI. Por otro lado un sistema de control difuso que utiliza el método COA presenta menos errores que aquel en donde se utiliza el método MOM; sin embargo, la estrategia MOM es mejor que la MAX.

## 3.4 MODELOS DE SISTEMAS DE CONTROL DIFUSO (SCD)

Tres son los modelos de SCD más utilizados en la actualidad, estos son:

- 1) Modelo Mamdani.
- 2) Modelo Sugeno-Takagi.
- 3) Modelo Tsukamoto.

A continuación se hace una breve descripción de cada uno de estos modelos.

### 3.4.1 Modelo Mamdani

El primer modelo de SCD fue propuesto por Mamdani en 1975 [12]. El modelo Mamdani fue realizado para controlar una pequeña máquina de vapor y un calentador. Esto se realizó mediante la combinación de un conjunto de reglas lingüísticas de la forma "si (*antecedente*), entonces (*consecuente*)" obtenidas de la experiencia de operadores humanos. La característica de este modelo es que tanto en la parte del antecedente como en la parte del consecuente se definen FM comúnmente utilizadas, tales como triangular, trapezoidal o gaussiana. La forma en que se lleva a cabo el razonamiento difuso en este tipo de controladores

es aplicando la regla composicional *max-min*, tal como se muestra en la figura 3.15. La salida  $z_o$  se obtiene aplicando el método del centroide, pero cualquier otro método puede aplicarse.

A la propuesta original de Mamdani se le han hecho modificaciones, por ejemplo, en lugar de utilizar la composición *max-min* también se puede utilizar la composición *max-product*. En este caso el razonamiento es el mostrado en la figura 3.16. Otra modificación fue realizada por Bart Kosko, quien utiliza la regla composicional *sum-product* y a dicho SCD se le denomina sistema difuso aditivo [10], como se vio anteriormente. Este razonamiento se muestra en la figura 3.17.

### 3.4.2 Modelo Sugeno-Takagi

Este modelo fue propuesto por Sugeno, Takagi y Kang [13] en un esfuerzo por desarrollar un sistema capaz de generar, de manera sistemática, reglas difusas a partir de un conjunto dado de entradas-salidas. Una regla típica de este tipo de modelo es :

“si  $x$  es  $A$  y  $y$  es  $B$ , entonces  $z = f(x, y)$ ”

donde  $A$  y  $B$  son conjuntos difusos en el antecedente;  $z = f(x, y)$  es una función nítida en el consecuente. De manera general,  $f(x, y)$  se define como un polinomio de las variables de entrada  $x$  y  $y$ , pero se puede emplear cualquier función que describa de manera apropiada la salida del sistema dentro de la región difusa especificada por el antecedente de cada regla. Cuando  $f(x, y)$  es un polinomio de primer orden el SCD se denomina modelo Sugeno de primer orden, el cual corresponde a la propuesta original descrita por Sugeno en [13]. Si  $f(x, y)$  se define como una constante, entonces el SCD se conoce como modelo Sugeno de orden cero. Este modelo puede verse como un caso especial del modelo Mamdani en el cual el consecuente de cada regla queda especificado mediante un singleton.

Algo importante de hacer notar es el hecho de que la salida de un SCD tipo Sugeno de orden cero es, en la medida en que las FM vecinas en la premisa de cada regla tengan el traslape adecuado, una función suave de sus variables de entrada. Es decir, el traslape de las FM en el consecuente de una regla no tiene efectos decisivos en la suavidad de la interpolación. Es el traslape de las FM en el antecedente de las reglas lo que determina la suavidad del comportamiento entrada-salida.

En la figura 3.19 se muestra el razonamiento difuso para un SCD Sugeno de primer orden. Dado que cada regla tiene una salida nítida, la salida total del sistema se obtiene mediante un promedio ponderado de la conclusión aportada por cada regla disparada, lo cual implica un menor consumo de tiempo en su cálculo.

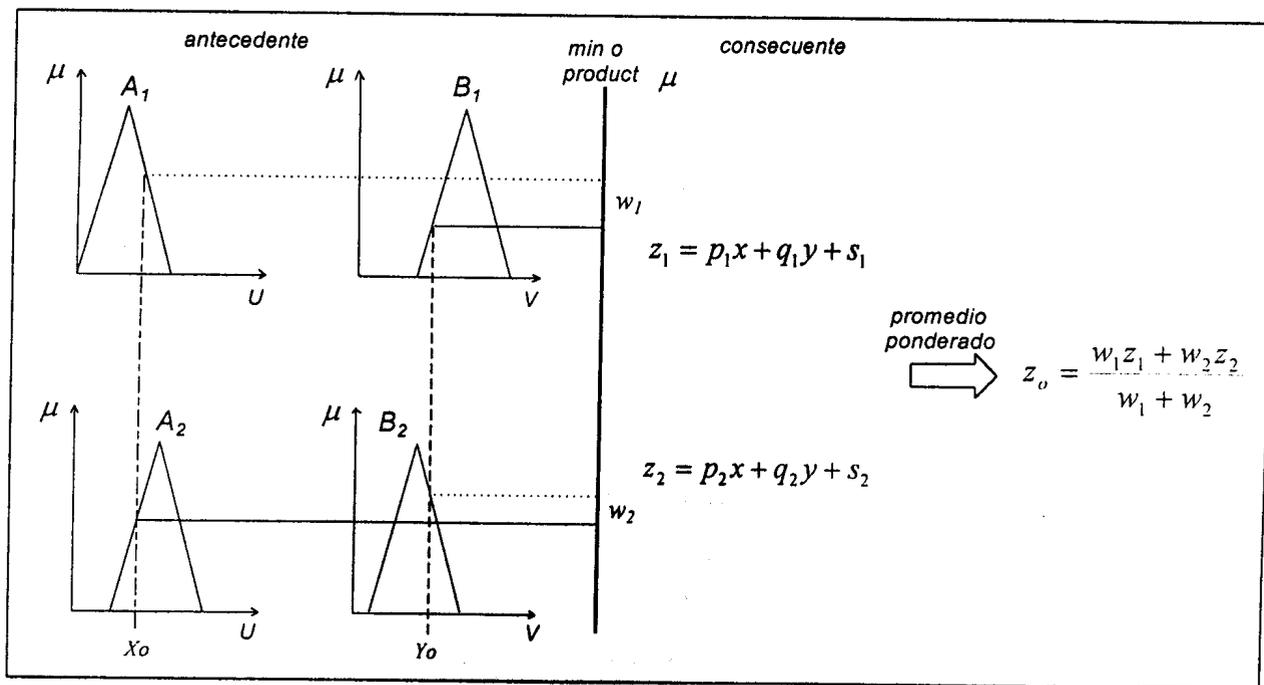


Figura 3.19 Ejemplo del razonamiento difuso para un SCD Sugeno de primer orden

### 3.4.3 Modelo Tsukamoto

En el modelo Tsukamoto [9] el consecuente de cada regla se representa con un conjunto difuso cuya FM es una función monótona, tal como se muestra en la figura 3.20. Como resultado de lo anterior, la salida inferida para cada regla esta definida como un valor nítido inducido por el grado de activación de dicha regla. La salida total del sistema es el promedio ponderado

de las salidas parciales dadas por cada regla. La figura 3.20 muestra el razonamiento en este modelo, para el caso de dos entradas y dos reglas.

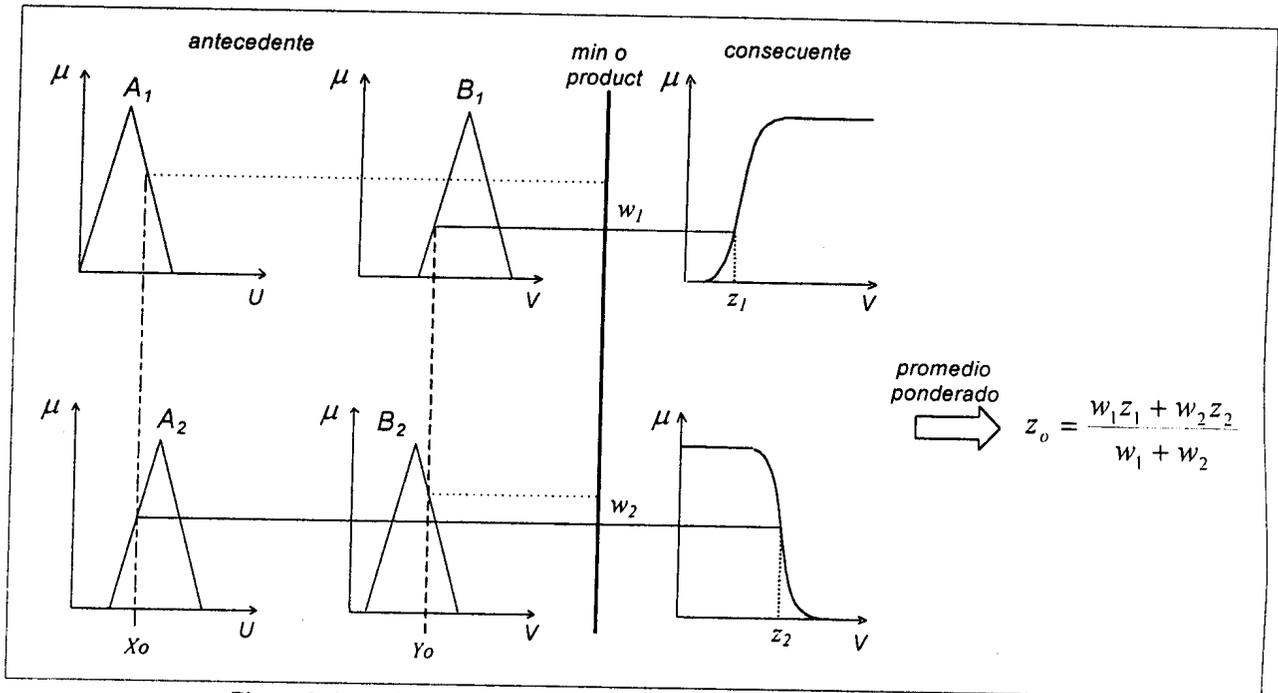


Figura 3.20 Ejemplo del razonamiento difuso para un SCD tipo Tsukamoto.

### 3.5 SISTEMAS DE CONTROL DIFUSO COMO APROXIMADORES UNIVERSALES

Durante las últimas 3 décadas los SCD han tenido gran éxito en la solución de muy variados y diversos problemas de control. Sin embargo, la respuesta a la pregunta del porqué de tal éxito había permanecido sin contestar hasta hace muy poco. Actualmente existen diversos trabajos que demuestran que un SCD que cumple ciertas condiciones es un aproximador universal de funciones [14,15], esta es la razón de su éxito. A continuación se describe, básicamente, el trabajo realizado por Castro [14] en el cual se demuestra como los SCD pertenecientes a una clase amplia de los mismos son capaces de aproximar cualquier función real en un conjunto compacto de arbitraria exactitud.

Considérese un SCD que consta de las tres etapas descritas en la sección anterior: fuzzificación, evaluación de reglas y defuzzificación; y una clase fija de reglas difusas *RUL*. Dada una función arbitraria, real y continua en un universo compacto  $U \subset \mathbb{R}^n$ , y un cierto  $\epsilon > 0$ , es posible encontrar un conjunto de reglas en *RUL* tales que el SCD asociado aproxima  $f$  con la exactitud  $\epsilon$ . Dos clases de SCD realizan dicha aproximación a) SCD con conjuntos difusos en la parte consecuente de las reglas (ejemplo, modelo Mamdani); b) SCD que no tienen conjuntos difusos como consecuentes en las reglas (ejemplo, modelo Sugeno).

Para cada  $a < b \in \mathbb{R}$  sea  $\mu(a,b): \mathbb{R} \rightarrow \mathbb{R}$  una función de membresía tal que  $\mu(a,b)(x) \neq 0$  si y sólo si  $x \in (a,b)$ . Sean  $T$  y  $T'$  dos T-normas,  $I$  una implicación difusa (o relación difusa), y  $S$  una S-norma.

#### 3.5.1 SCD con Consecuentes Difusos son Aproximadores Universales

Sea  $S_I = S_I(T', T, I, S, \mu(a,b))$  la familia de todos los SCD donde:

- i) El método de fuzzificación es el singleton.
- ii) La base de reglas esta compuesta por un número finito de reglas de la forma:

Si  $x_1$  es  $A_1$  y  $\dots$  y  $x_n$  es  $A_n$ , entonces  $y$  es  $B$

donde la función de membresía de cada  $A_{ij}$  es  $\mu(a^1_{ij}, a^2_{ij})$  para alguna  $a^1_{ij} < a^2_{ij} \in \mathbb{R}$ , ejemplo:

$$A_{ij}(x_i) = \mu(a^1_{ij}, a^2_{ij})(x_i)$$

y la función de membresía de cada  $B_j$  es  $\mu(b^1_j, b^2_j)$  para alguna  $b^1_j < b^2_j \in \mathbb{R}$ , ejemplo:

$$B_j(y) = \mu_{A_j}(y)$$

iii) La inferencia se realiza con:

- $T$  como conjunción difusa,
- el modus ponens generalizado construido con  $T'$  e  $I$ ,
- la combinación de todas las inferencias difusas se realiza por medio de  $S$ .

Entonces, la inferencia difusa se realiza de la siguiente forma:

a) Una regla

$$R_j: \text{ Si } x_1 \text{ es } A_{1j} \text{ y } \dots \text{ y } x_n \text{ es } A_{nj}, \text{ entonces } y \text{ es } B_j$$

será aplicada si y sólo si la entrada  $\bar{x} = \bar{x}^o$  activa el antecedente; ejemplo, si  $A_j(\bar{x}) \neq 0$ , siendo  $T(A_{1j}(x_1), A_{2j}(x_2), \dots, A_{nj}(x_n))$ .

b) Si la entrada  $\bar{x}^o$  dispara el antecedente, la inferencia es:

$x_1$  es  $A_1$  y  $x_2$  es  $A_2$  y ... y  $x_n$  es  $A_n$ , entonces  $y$  es  $B$

$$\frac{\bar{x} \text{ es } A'}{y \text{ es } B'}$$

$$B'(y) = \sup \{T'(A'(\bar{x}), I(A(\bar{x}), B(y))) / \bar{x} \in R^n\}$$

$$A(\bar{x}) = T(A_1(x_1), A_2(x_2), \dots, A_n(x_n)),$$

y como entrada se tiene el punto  $\bar{x} = \bar{x}^o$ :

$$A'(x) = \begin{cases} 1 & \text{si } \bar{x} = \bar{x}^o \\ 0 & \text{en otro caso} \end{cases}$$

el resultado es trasladado a:

$$B'(y) = T'(1, I(A(\bar{x}^o), B(y))) = I(A(\bar{x}^o), B(y)).$$

c) En general, la inferencia de la regla  $R_j$  cuando la entrada es  $\bar{x} = \bar{x}^o$  se puede expresar por:

$$B'_j(y) = \begin{cases} 0 & \text{si } A_j(\bar{x}^o) = 0 \\ I(A_j(\bar{x}^o), B_j(y)) & \text{en otro caso} \end{cases}$$

que en el caso de la implicación  $T$ -norma es la expresión general:

$$B'_j(y) = I(A_j(\bar{x}^o), B_j(y)).$$

d) La combinación de todas las inferencias difusas se realiza mediante  $S$ :

$$B'(y) = S(\{B'_j(y)\}_j)$$

iv) El método de defuzzificación satisface la propiedad de producir un punto en el soporte del conjunto difuso original, es decir:

$$y^o = S(\bar{x}^o) \in \text{Supp}(B'),$$

donde  $S \in S_f$  es el mapeo global entrada-salida.

Los únicos parámetros no especificados en esta clase son el número de reglas  $k$  y los coeficientes que determinan las funciones de membresía:  $a^1_{ij}, a^2_{ij}, b^1_j, b^2_j, (i = 1, \dots, n; j = 1, \dots, k)$ .

Teorema 1: Sea  $f: U \subseteq R^n \rightarrow R$  una función continua en un universo de discurso compacto  $U$ . Si  $I(a, 0) = 0$ , si  $a \neq 0$  (una implicación  $R$  o por ejemplo una  $T$ -norma), entonces para cada  $\epsilon > 0$  existe una  $S_\epsilon \in S_f$  tal que:

$$\sup \{ |f(\bar{x}) - S_\epsilon(\bar{x})| \mid \bar{x} \in U \} \leq \epsilon.$$

Lema 1: Bajo las condiciones del teorema 1 existe una  $S_\epsilon \in S_f$  tal que:

$$B'(y) |f(\bar{x}) - y| \leq B'(y) * \epsilon, \text{ para cada } y \in R.$$

Demostración: Sea  $\bar{a} \in U$ . Como  $f$  es continua en  $\bar{a}$ , para cada  $i = 1, \dots, n$  existe una  $\delta^i_{\bar{a}} > 0$  tal que:

$$|x_i - a_i| < \delta^i_{\bar{a}} \quad (i = 1, \dots, n) \Leftrightarrow |f(\bar{x}) - f(\bar{a})| \leq \epsilon/2.$$

Para cada  $\bar{a} \in U$ , sea:

$$O_{\bar{a}} = \{ \bar{x} \mid |x_i - a_i| < \delta^i_{\bar{a}} \quad (i = 1, \dots, n) \}.$$

Entonces,  $O_{\bar{a}}$  es un abierto en  $R^n$  y  $U \subseteq \bigcup_{\bar{a} \in U} O_{\bar{a}}$ , dado que  $\bar{a} \in O_{\bar{a}}$  para cada  $\bar{a} \in U$ . Como  $U$  es compacto, existe una subfamilia finita  $O_{\bar{a}^1}, O_{\bar{a}^2}, \dots, O_{\bar{a}^k}$  tal que:

$$U \subseteq O_{\bar{a}^1} \cup O_{\bar{a}^2} \cup \dots \cup O_{\bar{a}^k}.$$

Sea  $S_\epsilon \in S_f$  definido por:

$$A_{i,j}(x) = \mu(a^1_i - \delta^1_{\bar{a}^j}, a^1_i + \delta^1_{\bar{a}^j})(x) \quad (i = 1, \dots, n; j = 1, \dots, k)$$

y

$$B_j(y) = \mu(f(\bar{a}^j) - \epsilon/2, f(\bar{a}^j) + \epsilon/2)(y) \quad (j = 1, \dots, k).$$

Ahora:

Si  $B'(y) = 0$ , el lema es trivial.

Si  $B'(y) > 0$ , entonces  $B'(y) = S((B'_j(y)))_{j=1, \dots, k}$  existe una  $j_0$  tal que  $B_j(y) > 0$ . Se tiene que:

a) De  $B_j(y) > 0$ , se tiene  $A_{j_0}(\bar{x}^0) \neq 0$ , y  $\bar{x}^0 = T(A_{1j_0}(x_1^0), A_{2j_0}(x_2^0), \dots, A_{nj_0}(x_n^0))$ , de aquí se tiene:

$$\sup \{ |x_i(\bar{x}^0) - a_i| \mid \bar{x}^0 \in U \} < \delta^i_{\bar{a}^{j_0}} \quad (i = 1, \dots, n).$$

Entonces 2 existe una  $S_\epsilon \in S_f$

b) De  $B_j(y) \neq 0$ , se tiene  $B_j(y) = I(A_{j_n}(x^0), B_{j_n}(y)) \neq 0$ . Si  $B_{j_n}(y) = 0$ , entonces  $B'_{j_n}(y) = I(A_{j_n}(x^0), B_{j_n}(y)) = I(A_{j_n}(x^0), 0) = 0$  (de la propiedad de la implicación seleccionada). De aquí,  $B_j(y) \neq 0$ , esto origina que  $B_{j_n}(y) \neq 0$ .

De b) se tiene  $|f(\bar{a}^{j_n}) - y| \leq \epsilon/2$ .

De a) se tiene  $|x_i^0 - a_i^{j_n}| < \delta_{a_i^{j_n}}^i$  ( $i = 1, \dots, n$ ) y  $|f(x^0) - f(\bar{a}^{j_n})| \leq \epsilon/2$ .

Por lo tanto:

$$|f(x^0) - y| \leq |f(x^0) - f(\bar{a}^{j_n})| + |f(\bar{a}^{j_n}) - y| \leq \epsilon/2 + \epsilon/2 = \epsilon$$

Q. E. D.

Demostración del teorema 1: De acuerdo con la hipótesis planteada,  $S_\epsilon(x^0)$  pertenece a el soporte de  $B'$  como  $S_\epsilon(x^0) = \text{defuzz}(B')$ . De aquí  $B'(S_\epsilon(x^0)) \geq 0$ , y entonces de  $B'(y) |f(x^0) - y| \leq B'(y) \cdot \epsilon$ , se puede concluir  $|f(x^0) - S_\epsilon(x^0)| \leq \epsilon$ . Q. E. D.

### 3. 5. 2 SCD con Consecuentes no Difusos son Aproximadores Universales

Sea  $S_2 = S_2(T', T, I, \mu(a,b))$  definida como  $S_1$ , pero intercambiando ii) por ii') y iv) por iv')

ii') La base de reglas esta compuesta por un número finito de reglas de la forma:

Si  $x_1 = a_1^j$  y  $x_n = a_n^j$ , entonces  $y$  es  $w$

donde la funcione de membresía de cada  $A_j$  es  $\mu_{A_j}(x) = 1$  para alguna  $a_1^j < a_2^j \in R$ , y  $w \in R$ .

iv') La defuzzificación es  $\delta = \sum \lambda_i w_i$ , donde  $\lambda_i = \frac{B'(w_i)}{\sum B'(w_i)}$  y  $B'(w_i) = \mu_{B'}(w_i)$  para alguna  $w_i$  que exista una

$$B'(w_i) = \mu_{B'}(w_i) = \mu_{B'}(w_i) \cdot \mu_{A_1}(x_1) \cdots \mu_{A_n}(x_n)$$

y  $\lambda_i = \lambda_i / \sum \lambda_j$ .

En este caso la expresión general de  $B'$  es:

$$B'_j(y) = \begin{cases} \mu_{A_j}(x^0) & \text{si } A_j(x^0) = 0 \text{ ó } y \neq w \\ \mu_{A_j}(x^0) \cdot \mu_{B_j}(y) & \text{en otro caso} \end{cases}$$

Los únicos parámetros no especificados en esta clase son el número de reglas  $k$ , tales que  $j = 1, \dots, k$ ; así como los parámetros:  $a_1^j, a_2^j$ , que describen las funciones de membresía; y  $w_j$  ( $i = 1, \dots, n; j = 1, \dots, k$ ).

Teorema 2: Sea  $f: U \subseteq R^n \rightarrow R$  una función continua en un universo de discurso compacto  $U$ . Para cada  $\epsilon > 0$  existe una  $S_\epsilon \in S_2$  tal que:

$$|f(x) - S_\epsilon(x)| / x \in U \leq \epsilon$$

Lema 2: Bajo las condiciones del teorema 2 existe una  $S_\epsilon \in S_1$  tal que:

$$B'(y)|f(\bar{x}) - y| \leq B'(y)* \in \epsilon, \text{ para cada } y \in R.$$

Demostración: Sea  $\bar{a} \in U$ . Como  $f$  es continua en  $\bar{a}$ , para cada  $i = 1, \dots, n$  existe una  $\delta_{\bar{a}}^i > 0$  tal que:

$$|x_i - a_i| < \delta_{\bar{a}}^i \quad (i = 1, \dots, n) \Leftrightarrow |f(\bar{x}) - f(\bar{a})| \leq \epsilon.$$

Para cada  $\bar{a} \in U$ , sea:

$$O_{\bar{a}} = \left\{ \bar{x} / |x_i - a_i| < \delta_{\bar{a}}^i \quad (i = 1, \dots, n) \right\}.$$

Entonces,  $O_{\bar{a}}$  es un abierto en  $R^n$  y  $U \subseteq \bigcup_{\bar{a} \in U} O_{\bar{a}}$ , dado que  $\bar{a} \in O_{\bar{a}}$  para cada  $\bar{a} \in U$ . Como  $U$  es compacto, existe una subfamilia finita  $O_{\bar{a}^1}, O_{\bar{a}^2}, \dots, O_{\bar{a}^k}$  tal que:

$$U \subseteq O_{\bar{a}^1} \cup O_{\bar{a}^2} \cup \dots \cup O_{\bar{a}^k}.$$

Sea  $S_\epsilon \in S_2$  definido por:

$$A_{i,j}(x) = \mu(a_i^j - \delta_{\bar{a}^j}^i, a_i^j + \delta_{\bar{a}^j}^i)(x) \quad (i = 1, \dots, n; j = 1, \dots, k)$$

y

$$w_j(y) = f(\bar{a}^j) \quad (j = 1, \dots, k).$$

Ahora:

Si  $B'(y) = 0$ , el lema es trivial.

Si  $B'(y) > 0$ , entonces  $B'(y) = S((B'_j(y))_{j=1, \dots, k}) > 0$ , de aquí existe una  $j_0$  tal que  $B_{j_0}(y) > 0$ . Se tiene que:

a) De  $B_{j_0}(y) \neq 0$ , se tiene  $A_{j_0}(\bar{x}^0) \neq 0$ , y como  $A_{j_0}(\bar{x}^0) = T(A_{1j_0}(x_1^0), A_{2j_0}(x_2^0), \dots, A_{nj_0}(x_n^0))$ , de aquí se tiene:

$$A_{ij_0}(x_i^0) \neq 0 \quad (i = 1, \dots, n).$$

c) De  $B_{j_0}(y) \neq 0$ , se tiene  $y = w_{j_0} = f(\bar{a}^{j_0})$ . De a) se tiene que  $|x_i^0 - a_i^{j_0}| < \delta_{\bar{a}^{j_0}}^i \quad (i = 1, \dots, n)$  y  $|f(\bar{x}^0) - f(\bar{a}^{j_0})| \leq \epsilon$ . De aquí se tiene:

$$|f(\bar{x}^0) - y| \leq |f(\bar{x}^0) - f(\bar{a}^{j_0})| \leq \epsilon$$

Q. E. D.

Prueba del teorema 2:

$$|f(\bar{x}^0) - S_\epsilon(\bar{x}^0)| \leq \frac{\sum B'(y)|f(\bar{x}^0) - y|}{\sum B'(y)} \leq \frac{\epsilon * \sum B'(y)}{\sum B'(y)} \leq \epsilon.$$

Q. E. D.

## BIBLIOGRAFÍA

- [1] Padilla, Lares Francisco, *Diseño y configuración de control de procesos*, DEPTAC, México, 1992.
- [2] Creus; Antonio; *Instrumentación Industrial*, Alfaomega-Marcombo, 4ª Edición, México, 1992.
- [3] Earl, Cox; Fuzzy Fundamentals, *IEEE Spectrum*, October 1992, pp. 58-61.
- [4] Lee, Chuen Chien; Fuzzy Logic in Control Systems: Fuzzy Logic Controller, Part I and II, *IEEE Trans. on Syst. Man and Cybern.*, Vol. SMC-20, No. 2, 1990, pp. 404-435.
- [5] Procyk T. J., Mamdani E. H., A Linguistic Self-Organising Process Controller, *Automatica*, Vol.155, 1979, pp. 15-30.
- [6] Zadeh., L. A., Outline of a New Approach to the analysis of Complex Systems and Decision Processes, *IEEE Transactions on Syst, Man and Cyber*, Vol. SMC-3, No. 1, pp. 28-44, Jan 1973.
- [7] Murphy, Pat. Fuzzy logyc and its application in control systems, *ISA*, 1991, Paper No. 91-0474; pp. 1231-1245.
- [8] Harris, C. J. and Brown, M., *Neurofuzzy Adaptive Modelling and Control*, Prentice Hall International, UK, 1998.
- [9] Jyn-Shing, R. Jang; Neuro-fuzzy Modeling and Control, *Proceedings of the IEEE*, Vol. 83, No. 3, pp. 387-406, March, 1995.
- [10] Bart Kosko; *Neural Networks and Fuzzy Systems*, Prentice Hall, USA, 1992.
- [11] Bart Kosko; Fuzzy Systems as Universal Approximators, *IEEE Trans. on Computers*, Vol. 43, No. 11, Nov., 1994, pp. 1329-1333.
- [12] Mamdani, E. H. ans S. Assilian, An experiment in linguistic synthesis with a fuzzy logic controller, *Int. J. Man-Machine Studies*, vol. 7, No. 1, 1975, pp. 1-13.
- [13] Takagi, T. and Sugeno, M., Fuzzy identification of systems and its applicatins to modelling and control, *IEEE Trans. on Syst. Man and Cybern.*, Vol. 15, 1985, pp. 116-132.
- [14] Castro, J. L., Fuzzy Logic Controllers Are Universal Approximators, *IEEE Trans. on Syst. Man and Cybern.*, Vol. 25, No. 4, April 1995, pp. 629-635.
- [15] Wang, Li-Xin; Fuzzy Systems Are Universal Approximators, *IEEE*, 1992.